The Efficient Production Of Law
The Political Economy of the Sources of Law

by

Daniele Bertolini

A thesis submitted in conformity with the requirements for the degree of Doctor of Juridical Science
Faculty of Law
University of Toronto

© Copyright by Daniele Bertolini 2014
Abstract

Law and economics scholarship has been predominantly concerned with the content of legal rules rather than the process by which rules are created. The analytical separation of law by its formative process has resulted in an almost exclusive focus on the allocative efficiency of legal entitlements. According to this view, legal rights are treated as “commodities” that people (absent transaction costs and wealth effects) can freely buy and sell, such that the rights are allocated to their highest valued use. In this thesis, I maintain that this conventional approach needs to be integrated with a complementary, process-oriented analysis, capable of accounting for the causal relationship between the efficiency of legal rules and the efficiency of the lawmaking process. The central hypothesis of the present research is that the efficiency of the law-making process is not neutral with respect to the efficiency of the rules: the more efficient the process is in dealing with the pervasive information and public choice problems inherent in the production of legal rules, the better the outcome will be from an efficiency standpoint.

Once this logic is recognized, the problem of legal efficiency becomes one of identifying the comparative advantages and disadvantages of alternative sources of law. Based on
these premises, this thesis has two ambitious purposes: (i) to develop a methodology for
the analysis of alternative lawmaking institutions, based on the idea of “process
efficiency” and a unified taxonomy of lawmaking costs, and (ii) to provide a comparative
analysis of alternative lawmaking processes (i.e., politics, bureaucracy, adjudication and
spontaneous lawmaking), with the purpose of identifying their relative advantages and
disadvantages in various regulated environments. The ultimate goal is to advance the
understanding of the relationship between lawmaking mechanisms and the efficiency
properties of legal rules and, hence, to offer a more refined toolbox for identifying
efficiency improvements in the organization of the sources of law.
To the memory of my father, Giuseppe Bertolini
Acknowledgments

I would like to express my deepest appreciation and sincerest gratitude to my supervisor, Professor Michael Trebilcock, for his excellent guidance, for his constant support, and for providing me with an ideal atmosphere for completing my research at the University of Toronto Faculty of Law. During the challenging times of writing this thesis, he gave me the support and freedom I needed to continue. I could not have imagined having a better supervisor for my doctoral study.

I would like to sincerely thank the other members of my thesis committee. I have greatly benefited from Professor Giorgio Rampa’s detailed comments on the use of economic formalization. I thank him for his patience in discussing economic concepts with me and for being tolerant of my excess of enthusiasm in transferring microeconomic models into the legal context. I am grateful to Professor Bruce Chapman for his insightful comments, his generous help, and his kind encouragement. My discussions with him have always been illuminating and have been greatly beneficial to my work. I would also like to thank Professor Edward Iacobucci for his hard questions and sharp objections, which often led me to reconsider (and improve) portions of my thesis.

I owe my sincere gratitude to Professors Nuno Garoupa and Dhammika Darphala of the College of Law at the University of Illinois at Urbana-Champaign, who have always supported me in pursuing my research and academic goals and encouraged me to continue my research activity in the field of law and economics. I owe my gratitude also to Dean Charlotte Ku, who first gave me the opportunity to undertake the study of law and economics at the College of Law at Urbana-Champaign.

A heartfelt thank to Professor Rofolfo Jannaccone Pazzi of the University of Pavia Faculty of Law, under whose supervision I conducted my first study at the intersection of law and economics many years ago. I still remember the day that he enthusiastically agreed to supervise my research project, entitled “An Economic Approach to the Legal
Value Of Academic Degrees”—a hotly debated topic, a widely criticized approach, and a rather unpopular thesis within the Italian academic environment. Everything started with his support, and with the many conversations we used to have in his office on the first floor of the Dipartimento di Economia Pubblica e Territoriale. Most of all, I thank him for his deep and sincere friendship, which has lasted over the years and continued to be a source of great support throughout my doctoral study.

I would like to thank and express my immense gratitude to my mum, Anna Maria Ariatti, and my sister, Sara Bertolini, for their unconditional support. They have always been there for me, standing by me through the most challenging times. Without their support, this research thesis would not have been possible.

Thanks to Youngming Song for his friendship over the last three years and for being always ready to interrupt his work and join me for our many Italian coffee breaks, without which—and he knows it—my productivity falls inexorably. Thanks also to Nir Harrel for enriching my academic experience with his warm and sincere friendship.

Finally, thanks to Nataly, who brightens my days and whose presence fills me with joy.
Law-making is much more a theoretical process than an act of will.

Table of Contents

ABSTRACT ................................................................................................................................. ii

ACKNOWLEDGEMENTS ........................................................................................................... v

INTRODUCTION ........................................................................................................................ 1
  1. PROLOGUE .......................................................................................................................... 1
  2. ANALYTICAL AIMS ........................................................................................................... 3
  3. RELEVANCE ........................................................................................................................ 6
  4. PREVIOUS CONTRIBUTIONS .............................................................................................. 7
      4.1. Komesar’s Participation-Centered Approach ............................................................... 8
      4.2. Parisi’s Investment-Theory Approach ......................................................................... 9
  5. ORGANIZATION OF THE INQUIRY ................................................................................. 10
      5.1. Definitions ..................................................................................................................... 10
      5.2. Analytical Core ............................................................................................................. 13
          5.2.1. Output Efficiency ................................................................................................. 14
          5.2.2. Process Efficiency ................................................................................................. 14
          5.2.3. Law-making Costs ................................................................................................. 15
          5.2.4. Supply of and Demand for Law .......................................................................... 15
          5.2.5. The Variables Affecting Process-Efficiency ......................................................... 17
      5.3. Hypotheses .................................................................................................................... 19
  6. LIMITATIONS ....................................................................................................................... 21
  7. SUMMARY OF CONTENTS .................................................................................................... 22

PART I
METHODOLOGY

CHAPTER 1 EFFICIENCY AS A LEGAL CONCERN ................................................................. 24
  INTRODUCTION ...................................................................................................................... 24
  1.1. LOGICAL FEATURES OF LEGAL EFFICIENCY .......................................................... 26
      1.1.1. Efficiency as a Relative Concept ............................................................................. 27
          1.1.1.1. Efficiency as a Function of Rights .................................................................... 29
          1.1.1.2. Efficiency and Distribution ............................................................................ 30
          1.1.1.3. Means-Ends Relationship ............................................................................. 33
      1.1.2. Efficiency as a Comparative Criterion ..................................................................... 35
      1.1.3. Theoretically Sound Definition of Efficiency: Avoiding Logical Circularity .......... 36
  1.2. THREE NOTIONS OF ECONOMIC EFFICIENCY ......................................................... 38
      1.2.1. Allocative Efficiency ............................................................................................... 39
          1.2.1.1. Allocative Efficiency and Economic Surplus .................................................... 43
          1.2.1.2. Allocative Efficiency and Single Institutional Analysis .................................... 43
      1.2.2. Productive Efficiency ............................................................................................... 47
      1.2.3. Dynamic Efficiency ................................................................................................ 49
  1.3. LIMITATIONS OF OUTPUT-ORIENTED APPROACH ................................................ 50
      1.3.1. Pareto Efficiency ...................................................................................................... 50
          1.3.1.1. Logical Criticisms .............................................................................................. 52
PART II
FUNDAMENTAL LAW-MAKING DIMENSIONS

CHAPTER 3 EX-ANTE VERSUS EX-POST LAW-MAKING .............................. 144
INTRODUCTION: ASSUMPTIONS ................................................................ 145
3.1. PRODUCTIVE EFFICIENCY .................................................................. 147
  3.1.1. The Demand for Ex Ante Legal Specificity ...................................... 147
    3.1.1.1. The Compliance-Information Costs ........................................... 149
    3.1.1.2. The Expected Benefit from Compliance .................................... 152
    3.1.1.3. Probability and Frequency per Actor ......................................... 157
    3.1.1.4. Conclusions on Individual Legal Demand ................................. 159
  3.1.2. The Supply of Ex-Ante Legal Specificity ........................................ 159
    3.1.2.1. Cost-Functions of the Lawmaker .............................................. 161
    3.1.2.2. The Cost-Minimizing Supply of Law ......................................... 171
  3.1.3. The Maximization of Legal Compliance .......................................... 175
  3.1.4. Legal Homogeneity ...................................................................... 181
    3.1.4.1. Legal Homogeneity as Uniform Demand for Legal Protection .......... 181
    3.1.4.2. Legal Homogeneity as Productivity of Ex Ante Law-making ......... 185
    3.1.4.3. Measuring the Costs of Legal Heterogeneity ............................. 189
    3.1.4.4. Conclusions on Legal Heterogeneity ....................................... 191
  3.1.5. Legal Frequency .......................................................................... 191
    3.1.5.1. Aggregate Frequency and Frequency per actor .......................... 192
    3.1.5.2. Frequency and Scope of Application of the Law ......................... 194
    3.1.5.3. The Magnitude of the Interests at Stake .................................... 196
    3.1.5.4. The Number of Protected Third Parties .................................... 198
    3.1.5.5. Conclusions on Legal Frequency ............................................ 200
  3.1.6. Conclusions on Production Efficiency ........................................... 200
3.2. AGENCY EFFICIENCY ....................................................................... 203
  3.2.1. Ex-Ante Lawmaking and Commitment Problem ............................ 203
  3.2.2. Ex-Ante Lawmaking and Rent-Seeking ....................................... 206
3.3. SOCIAL CHOICE EFFICIENCY ......................................................... 207
3.4. ADAPTIVE EFFICIENCY ................................................................. 208
CONCLUSIONS ...................................................................................... 209

CHAPTER 4 CENTRALIZED VERSUS DECENTRALIZED LAW-MAKING 211
INTRODUCTION ..................................................................................... 212
4.1. EX ANTE VS. CENTRALIZED LAW-MAKING .................................... 213
4.2. PRODUCTIVE EFFICIENCY ............................................................. 215
  4.2.1. The Advantages of Centralization ................................................ 215
    4.2.1.1. Economies of Scale .............................................................. 215
    4.2.1.2. Uniformity and Legal Consistency ......................................... 216
    4.2.1.3. Regulating Externalities ......................................................... 219
  4.2.2. The Advantages of Decentralization ............................................ 219
    4.2.2.1. The Incentive Effect ............................................................ 221
    4.2.2.2. The Knowledge-Creating Effect ............................................. 224
    4.2.2.3. Maladaptation Costs ............................................................ 227
  4.2.3. Heterogeneity and Levels of Centralization .................................... 229
4.3. SOCIAL CHOICE EFFICIENCY (REMINDER) ................................... 239
4.4. AGENCY EFFICIENCY ................................................................. 240
5.2.5. Observed Institutional Stability ........................................... 253
5.2.5.1. Vote-trading .......................................................... 253

4.6.2. The Chaos Theorem ..................................................... 245
4.6.2.1. Arrow’s Theorem .................................................... 245
4.6.2.2. The Condorcet Criterion .......................................... 250
4.6.2.3. Intransitive Cycle .................................................. 250

5.2.4.1. The Exit-Downgrading Effect ..................................... 240
5.2.4.2. The Expected Return from Pressure ............................ 245
5.2.4.3. Competition among Principals ................................... 247
5.2.4.4. Competition among Agents ...................................... 249

5.2.3.2.3. Intransitive Cycle ................................................ 250
5.2.3.2.2. The Efficiency of the Condorcet Criterion ................ 250
5.2.3.2.1. The Condorcet Winner ........................................ 250
5.2.3.1.2. Efficiency of the Median Voter Platform ................. 250
5.2.3.1.1. The Median Voter Equilibrium ............................... 246
5.2.3.1. One-Stage Voting ................................................... 246
5.2.3.2. Pairwise Voting ..................................................... 246
5.2.3.1.1. Voter Uncertainty ............................................... 246
5.2.3.1.2. Voter Rational Ignorance ..................................... 247
5.2.3.1.3. Voter Rational Irrationality ................................... 249
5.2.3.1. The Demand Side .................................................. 246
5.2.2. People Subject to Law .................................................. 274
5.2.2.1. Voter Uncertainty ................................................... 274
5.2.2.2. Voter Rational Ignorance ......................................... 275
5.2.2.3. Voter Rational Irrationality ...................................... 276

5.2.2. Competition among Principals ........................................ 247
4.6.2. Competition among Principals ........................................ 250
4.6.1. Competition among Agents ........................................... 250

4.4.2.2. Competition among Principals .................................... 286
4.4.2.1. The Expected Return from Pressure ............................ 287
4.4.2. Competition among Agents ........................................... 288
4.4.1.2. Competition among Agents ....................................... 289
4.4.1. The Supply Side ....................................................... 289
4.4.1.1. The Exit-Downgrading Effect .................................... 290
4.4.1. The Moral Hazard of the Centralized Lawmaker ................ 290
4.4.1.1. Voters ................................................................. 290
5.1.1.1. Voters ................................................................. 290
5.1.1. The Demand Side ...................................................... 290
5.1.1.1.1. Voter Uncertainty ............................................... 290
5.1.1.1.2. Voter Rational Ignorance ..................................... 291
5.1.1.1.3. Voter Rational Irrationality ................................... 292

4.5. ADAPTIVE EFFICIENCY ................................................ 253
4.5.1. The Coordination of Compliance Decisions ....................... 254
4.5.2. Conclusions on Adaptive Efficiency ................................. 259

4.4.1.1.1. Voter Rational Ignorance ..................................... 292
4.4.1.1.2. Voter Rational Ignorance ..................................... 293
4.4.1.1.3. Voter Rational Irrationality ................................... 294

4.6. THE FUNDAMENTAL DIMENSIONS OF LAW-MAKING. INTEGRATING THE ANALYSIS. 260
4.6.1. Institutional Dimensions and the Four Component of Process 260
4.6.2. The Interaction between Institutional and Structural Dimensions 261

4.7. HYPOTHESES TO BE VERIFIED THROUGH COMPARATIVE INSTITUTIONAL ANALYSIS 264

PART III
COMPARATIVE INSTITUTIONAL ANALYSIS

CHAPTER 5 POLITICS ........................................................................ 268
INTRODUCTION. FROM THE RULE OF “LAW” TO THE RULE OF “LEGISLATORS” .... 269
5.1. PRODUCTIVE EFFICIENCY ................................................ 272
5.1.1. The Demand Side ....................................................... 273
5.1.1.1. Voters ................................................................. 273
5.1.1.1.1. Voter Uncertainty ............................................... 273
5.1.1.1.2. Voter Rational Ignorance ..................................... 274
5.1.1.1.3. Voter Rational Irrationality ................................... 275
5.1.1.2. People Subject to Law .............................................. 276
5.1.2. The Supply Side ....................................................... 288
Conclusions ............................................................................. 289

5.2. SOCIAL CHOICE EFFICIENCY .......................................... 291
5.2.1. Unanimity versus Majority Rule ..................................... 291
5.2.2. “Intensity,” “Externality” and “Cycling” Problems ............. 292
5.2.3. Majority Equilibrium ................................................... 296
5.2.3.1. One-Stage Voting ................................................... 296
5.2.3.1.1. The Median Voter Equilibrium ............................... 296
5.2.3.1.2. Efficiency of the Median Voter Platform ................. 298
5.2.3.2. Pairwise Voting ..................................................... 299
5.2.3.2.1. The Condorcet Winner ........................................ 299
5.2.3.2.2. The Efficiency of the Condorcet Criterion ............... 300
5.2.3.2.3. Intransitive Cycle ................................................ 300
5.2.4. No Majority Equilibrium .............................................. 307
5.2.4.1. Arrow’s Theorem .................................................... 307
5.2.4.2. The Chaos Theorem ................................................ 308
5.2.5. Observed Institutional Stability ....................................... 310
5.2.5.1. Vote-trading .......................................................... 310
6.1.1.3. X-Inefficiency .................................................................................. 403
   6.1.1.3.1. Labour/Capital Ratio ............................................................... 403
   6.1.1.3.2. Laziness .................................................................................. 404
   6.1.1.3.3. Risk Aversion and X-Inefficiency ........................................... 406
6.1.2. Output Ineffectiveness ...................................................................... 410
   6.1.2.1. Risk-Aversion and Output Effectiveness ...................................... 410
   6.1.2.2. The Absence of Price Signals ...................................................... 412
   6.1.2.3. Decision-Making Biases ............................................................... 415
6.2. AGENCY EFFICIENCY ........................................................................... 418
   6.2.1. Legislators as Faithful Agents ......................................................... 419
      6.2.1.1. Ex Post Control ...................................................................... 423
      6.2.1.2. Ex Ante Control .................................................................... 427
   6.2.1.3. The Determinants of Delegation .................................................. 430
      6.2.1.3.1. Regulatory Uncertainty .......................................................... 431
      6.2.1.3.2. Conflict of Preferences between Politicians and Bureaucrats . 433
      6.2.1.3.3. Asymmetry of Information .................................................... 436
   6.2.2. Legislators as Strategic Agents ......................................................... 439
      6.2.2.1. Blame-Shifting Delegation ........................................................ 439
      6.2.2.2. Uncertainty-Reducing Delegation .............................................. 441
      6.2.2.3. Credibility-Enhancing Delegation ............................................ 445
CONCLUSIONS ............................................................................................. 446

CHAPTER 7 ADJUDICATION ........................................................................ 452
INTRODUCTION ADJUDICATION AS A TRIADIC LAW-MAKING PROCESS .... 452
7.1. PRODUCTIVE EFFICIENCY ................................................................ 455
   7.1.1. The Demand Side ......................................................................... 455
      7.1.1.1. The Stakes .............................................................................. 456
      7.1.1.1.1. Inefficient rules are more frequently litigated than efficient rules. 457
      7.1.1.1.2. Inefficient rules are more frequently litigated when parties have future stakes in the dispute, ........................................... 460
      7.1.1.1.3. Inefficient Rules Are More Intensively Litigated than Efficient Rules Are .............................................................. 464
      7.1.1.2. The Information Theory ............................................................ 465
      7.1.1.2.1. Symmetric Information ......................................................... 466
      7.1.1.2.2. Asymmetric Information ....................................................... 467
      7.1.1.3. The Public Good Problem ......................................................... 469
   7.1.2. Conclusions on Productive Efficiency ............................................ 472
   7.1.2. The Supply Side ............................................................................ 472
      7.1.2.1. The Information Problem ......................................................... 473
      7.1.2.1.1. The Austrian Critiques .......................................................... 473
      7.1.2.1.2. Imperfect Decision-Making .................................................... 475
      7.1.2.2. The Doctrine of Precedent ......................................................... 477
      7.1.2.2.1. The Reason-Based Decision-Making ...................................... 477
      7.1.2.2.2. The Exclusionary Reasons of Precedents versus Statutes .... 480
      7.1.2.2.3. Advantages of the Doctrine of Precedent ............................... 481
      7.1.2.2.4. The Disadvantages of the Doctrine of Precedent .......... 484
      7.1.2.3. The Degree of Polycentrism ....................................................... 486
      7.1.2.3.1. “Grown Orders” versus “Made Orders” .................................. 487
      7.1.2.3.2. Judge-Made Law as a Grown Order ....................................... 488
      7.1.2.3.3. Judge-Made Law as a Made Order ........................................ 489

xiii
9. THE EFFICIENT PRODUCTION OF LEGAL RULES. INTEGRATING THE ANALYSIS

9.1. CORE THESIS

9.1.1. Law Cannot Be Separated By Its Sources

9.1.2. Law Is Not Allocatively Efficient: Scholars Should Focus More on “Inefficient” Law-making

9.1.3. Process-Efficiency Analysis Complements Output-Oriented Analysis

9.1.4. The Need for a General, Systematic Theory of Law-Making

9.2. METHODOLOGICAL CONTRIBUTION

9.2.1. Part I: Path to a Definition of Legal Efficiency

9.3. COMPARATIVE INSTITUTIONAL ANALYSIS

9.3.1. Part II: Fundamental Law-Making Dimensions

9.3.1.1. Hypothesis To Be Tested in Part II

9.3.2. Part III: Law-Making Institutions

9.3.3. Parts II and III: The Structural Variables

9.4. NORMATIVE CONTRIBUTION

9.4.1. Decentralizing and Depoliticizing the Production of Law

9.5. DIRECTIONS FOR FUTURE RESEARCH

9.5.1. The Historical Determinants of the Law-making Process

9.5.2. The Institutional Determinants of People’s Preferences

9.5.3. The Dynamic Interaction among Sources of Law

9.5.4. Biases and Heuristics in Law-Making
List of Figures

FIGURE 1 The Outcome-Process Relationship ................................................................. 17
FIGURE 2 Production Possibilities Frontier ........................................................................ 40
FIGURE 3 Utility Possibilities Frontier ................................................................................ 42
FIGURE 4: The Rancher-Farmer Example .......................................................................... 58
FIGURE 5: Edgeworth Box with changing rules of the game .......................................... 81
FIGURE 6 Cost-minimization Problem in Neoclassical Theory of the Firm .................... 106
FIGURE 7 U-Shaped Cost Curves ..................................................................................... 164
FIGURE 8 The Marginal and Average Production-Information Costs Curves ................. 166
FIGURE 9 Four Types of the Demand for Law .................................................................. 172
FIGURE 10 The Cost-Minimizing Supply of Law ............................................................... 174
FIGURE 11 Aggregate Frequency and Frequency per Actor ............................................. 194
FIGURE 12: Ex Ante-Ex Post Choice and the Types of Law Demand .............................. 202
FIGURE 13 Law-making Techniques for Increasing Levels of Legal Heterogeneity ......... 232
FIGURE 14: The Circular Relationship between Internal Rent-seeking and Centralization 243
FIGURE 15 The Four Fundamental Dimensions of Law-making .................................... 263
FIGURE 16: Single Peaked Voter Utility Functions .......................................................... 299
FIGURE 17. The Intensity Problem .................................................................................... 301
FIGURE 18 Coasean Bargaining for Votes in a Multidimensional Issue Space ............. 321
FIGURE 19. One-Dimensional Issue, Single-Peaked Preferences, Three Voters, Three-Alternatives, Pairwise Voting, and Agenda Setter’s Control ................................................................. 329
FIGURE 20 Available CW, Non-Condoorcet Rule, Agenda Manipulation ...................... 331
FIGURE 21 Condorcet Winner Coincides With Agenda Setter’s Preferred Outcome .......... 332
FIGURE 22 Alternative Models of Bureaucratic Behaviour ............................................. 409
FIGURE 23. The Preferred Delegation Degree, D*, Given Marginal Benefit And Cost ...... 423
FIGURE 24 The Effect Of Increased Uncertainty ............................................................... 431
FIGURE 25 The Effect of Increased Degree of Conflict .................................................... 436
FIGURE 26 Judicial Evolutionary Law-making ................................................................. 454
FIGURE 27 Settlement-Trial Decision .............................................................................. 467
FIGURE 28 Grow Orders versus Made Orders ................................................................. 493
FIGURE 29 Demand and Supply Curves in the Political and Adjudication Arena with Access Costs Thresholds ........................................................................................................ 497
FIGURE 30 Demand and Supply Curves in the Political and Adjudication Arena with Output Thresholds ............................................................................................................... 503
FIGURE 31 The Optimal Degree of Strictness of the Doctrine of Precedent .................... 523
FIGURE 32 Increase in the Rate of Change ....................................................................... 524
INTRODUCTION

1. Prologue

Law and economics scholarship has been predominantly concerned with the content of legal rules rather than the process by which rules are created. The analytical separation of law by its formative process has resulted in an almost exclusive focus on the allocative efficiency of legal entitlements. According to this view, legal rights are treated as “commodities” that people (absent transaction costs and wealth effects) can freely buy and sell, such that the rights are allocated to their highest valued use. This approach is congenial to economists’ prevailing attitude to predict efficient equilibria resulting from competitive market interactions. It is indeed convenient on technical grounds to forego all the complications associated with the highly complex dynamics that are at work in law-making processes and focus on the allocation of legal rights based on implicit highly simplifying institutional assumptions. However, although convenient this approach does not necessarily explain much.

This study stems from the belief, acquired during years of professional experience in a variety of legal settings, that the nature of substantive legal rules cannot be fully understood from an economic standpoint without careful inquiry into the mechanisms of law creation. Unlike what is assumed by standard economic models, the legal setting shows more complex relationships between the production process (i.e., law-making process) and the characteristics of the outcome (i.e., legal rules). Therefore, focusing on the content of legal rules while assuming that the law-making process is exogenous to the analysis of legal efficiency is not a promising methodology to capture the determinants of legal efficiency. In essence, law cannot be separated by its sources.

In light of the foregoing considerations, it is useful to distinguish at the outset between the following two distinct interpretations of legal efficiency: “output efficiency” and “process efficiency.” The first, which corresponds to the use of the term “efficiency” in the standard law and economics literature, views the efficiency as a function of the allocation of legal rights; the second, which is the major focus of this study, views efficiency of legal
rules as a function of the interaction between the institutional features of the law-making process and the characteristics of the regulated environments. This study examines the complex relationship between these two kinds of efficiency; furthermore, it addresses the central issue of whether and how the efficiency of the law-making process is related to the efficiency of the law-making outcome.

In general terms, the central hypothesis of my research agenda is that the efficiency of the law-making process is not neutral with respect to the efficiency of the rules: Rather, the more efficient the process is in dealing with the pervasive information and public choice problems inherent in the production of legal rules, the better the outcome will be from an efficiency standpoint. Once this logic is recognized, the problem of legal efficiency becomes one of identifying the comparative advantages and disadvantages of alternative sources of law relative to the characteristics of regulated environments. The major concern of process efficiency analysis is to investigate the dynamic interaction between law-making processes and regulated environments, and to examine the consequences of this interaction in terms of the efficiency properties of the outcomes, rather than to identify the “optimal” allocation of legal entitlements insulated from its formation process.

From the foregoing perspective, this thesis aims to enhance to the method and theory of law and economics by providing four distinct contributions:

(i) it discusses the problematic nature of efficiency in the legal context and identifies the limitations of the conventional output-oriented analysis (Chapter 1);
(ii) it develops a general and unified methodological framework for the efficiency analysis of law-making processes based on the idea of process efficiency and a related unified taxonomy of law-making costs (Chapter 2);
(iii) it identifies a general set of institutional and structural variables predicted to be most relevant in affecting the efficiency of the law (Chapters 3 and 4);
(iv) it compares alternative law-making institutions with the purpose of identifying the relative advantages and disadvantages of alternative sources of law (Chapters 5–8).
The following introductory pages define more clearly the goals of my research agenda and its style of analysis, provide some useful preliminary definitions, and anticipate a brief overview of the analytical path followed throughout the discussion.

2. Analytical Aims

This thesis has two ambitious purposes. The first is strictly methodological and can be synthesized in the idea that the process-outcome relationship should be regarded as the central focus of the law and economics discipline. From this perspective, this study investigates the routinely overlooked connection between process and outcome. The second analytical goal is to develop a comparative institutional analysis of alternative sources of law. From this perspective, I attempt to provide a full description of the mechanisms that most affect the efficiency of law-making institutions. This second strand of analysis prepares the groundwork for a normative discourse that I call “the political economy of the sources of law,” the practical goal of which is to provide policy-makers with guidelines on how to better organize the production of legal rules in various area of law.

The major contribution of this study is to provide a general unified framework for the analysis of the efficient production of law. The variables identified are numerous, and the salience of any particular explanatory factor in each context differs significantly across situations and institutions. This therefore makes it impossible to provide unique predictions with high level of confidence. Many see the absence of a unique prediction as a defect; I do not share this view. I am dealing here are mainly with general principles. I do not offer particular solutions for particular problems. I am convinced, however, that such solutions can be found much more easily in accordance with the general principles identified here than by applying the conventional outcome-oriented approach.

Before proceeding, I must offer a more precise qualification of what I mean by a “general unified analytical framework” of the law-making process. In order to better clarify this
point, it is useful to draw a distinction between the following three terms: “framework,” “theory” and “model.” Ostrom has proposed a conceptual distinction between these terms that illuminates the issue.\(^1\) The following three statements provide a useful reference point.

A *framework* helps to identify the elements (and the relationships among these elements) that one needs to consider for institutional analysis. Frameworks organize diagnostic and prescriptive inquiry. They provide the most general set of variables that should be used to analyze all types of settings relevant for the framework. […] *They attempt to identify the universal elements that any relevant theory would need to include.* […] *The elements contained in a framework help the analyst generate the questions that need to be addressed when first conducting an analysis.*\(^2\)

*Theories* enable the analyst to specify which components of a framework are relevant for certain kinds of questions and to make broad working assumptions about these elements. Thus, theories focus on parts of a framework and make specific assumptions that are necessary for the analyst to diagnose a phenomenon, explain its process and predict outcomes. […] *Several theories are usually compatible with any framework.*\(^3\)

*Models* make precise assumptions about a limited set of parameters and variables. [They] are used to explore the consequences of these assumptions systematically on a limited set of outcomes. \(^4\)

Ostrom’s proposed terminology helps to clarify the ultimate goal of my line of inquiry. A framework identifies a very general set of variables that should be considered in analyzing all types of law-making processes; it organizes the positive and prescriptive inquiry into the nature of the law-making process. My proposed framework is *general* because it identifies the “universal building blocks”\(^5\) underlying the diversity of law-making institutions; it is *unified* because it is compatible with many different theoretical approaches that are dispersed in various areas of social science literature (e.g., public choice analysis, transaction cost economics, positive political theory, evolutionary

---

\(^2\) *Ibid.* [Emphasis is mine].
\(^3\) *Ibid.* [Emphasis is mine].
\(^4\) *Ibid.* [Emphasis is mine].
\(^5\) *Ibid* at 5.
This general unitary framework of the law-making process prepares the groundwork for the future construction of a general systematic theory of law-making.

I emphasize that the literature about the law-making process, although rich and enlightening in many respects, has not yet produced a general systematic theory of law-making that enables the analyst to explain and assess the organization of the sources of law. The analysis developed in this thesis is certainly not mature enough to be regarded as a systematic theory of law-making; nonetheless, it is comprehensive enough to propose a preliminary framework upon which a more systematic comparative institutional analysis could be developed in the future. To date, no one in the literature (with a few notable exceptions that I will discuss shortly) has attempted to propose a general theory of law-making; it is likely that the failure to provide such a general theory is due to the lack of a commonly agreed unified analytical framework. In this respect, the unified analytical framework proposed here lays the groundwork for a systematic theory of law-making.

I do not mean to exhaust the issue of process efficiency or to provide the only valid explanatory framework of legal efficiency. However, I do hope to advance academic thinking beyond the generic recognition of the importance of the law-making process. Many might be pessimistic about the law-making process being susceptible to a satisfying general analysis. I prefer to be optimistic and think that it is possible to capture some core tendencies of the production of law. I am aware that the approach suggested here might not be able to explain many cases and also leaves many important instances out of the picture. However, I think that it is worth making the effort of developing a framework for a general and more systematic theory of law-making.

This study is addressed to the following three categories of readers: (i) students of law, (ii) law and economics scholars and (iii) legal scholars and professional lawyers (not listed them here in order of importance).
3. Relevance

The practical relevance of the theoretical work developed in this thesis is summarized by the following two considerations. The first is related to the role that economic and legal-economic scholarship plays in policy-making debates. Traditionally, economists seek to recommend specific policies to legislators and other policymakers that they believe to be efficient (e.g., “eliminate tariffs,” “liberalize protected markets” and so forth); likewise, law and economics scholars try to suggest specific legal rules they believe to be efficient. Despite these efforts by academic scholarship, policymakers often ignore, if not deliberately resist, such advice and adopt inefficient policies. This circumstance ironically reflects the uncomfortable position of legal-economics scholars with respect to the policy-making debate. In fact, as recognized by early public choice contributions, political actors’ preferences and incentives often conflict with the pursuit of socially efficient policies. This implies a paradoxical situation for scholars as policy advisors. If policymakers face personal incentives to adopt socially inefficient policies, then scholars’ policy recommendations are predictably superfluous and are reduced to mere flatus vocis. If policymakers face personal incentives to adopt socially efficient policies, then scholars’ policy recommendations are useless, as policymakers will adopt them even without scholars’ advice. In either case, policy recommendations have no practical relevance. The preceding consideration suggests that in order to avoid the pointlessness of recommendations that are unexceptionable in theory but unworkable in practice, law and economics scholars should pay closer attention to the choice of the policymaker that is most suitable for producing legal rules in a given area of law.

The second reason underlying the practical relevance of my proposed line of inquiry is that it is intended to complement the standard efficiency approach to legal rules. This is usually based on the assumption that individuals do not value the process by which rules are achieved. But if both process and rules are independent sources of utility, their analytical separation cannot be sustained at a conceptual level, and the efficiency calculation that does not incorporate them both is likely to be incomplete and misleading.
In short, the law-making process must be included in the efficiency calculus, despite the difficulties involved in measuring its costs and benefits.

4. Previous Contributions

Legal-economic scholarship on the law-making process is far from being at a point of maturity. I identify two main shortcomings. First, there is no consolidated notion of “process” efficiency. To date little attention has been paid to identifying the normative criteria for assessing the efficiency of the law-making process. As a result, there is a large disparity in maturity between the theoretical discussion on outcome efficiency and process efficiency. While outcome efficiency analysis relies on a well-established methodology and on a generally agreed analytical framework, the analysis of the law-making process is dispersed in different areas of the legal-economic literature; there is neither a systematic explanatory model of the law-making process, nor general agreement among scholars on the most desirable analytical framework to employ. Second, the causal relationship between law-making costs and institutional and structural variables is under-investigated. This is an obstacle toward identifying the efficient allocation of law-making powers among alternative sources of law. In addition, since there is no uniform taxonomy of law-making costs, there is also no uniform framework (nor theory) of law-making.

My thesis builds upon two major contributions on the economics of law-making that require a preliminary comparison to my proposed framework.6

6 Other contributions explicitly recognize the need for closer attention by legal-economic scholars to the law-making process. See, for example, Gillian K Hadfield, “Bias in the Evolution of Legal Rules” (1991) 80 Geo. L.J. 583 [Hadfield, “Bias”] at 585 (“If we desire efficiency […] we must reexamine the structure of the legal process”); Giuseppe Dari-Mattiacci and Bruno Deffains, “Uncertainty Of Law And The Legal Process” (2007) Journal of Institutional and Theoretical Economics, 627 [Dari-Mattiacci and Deffains, “Uncertainty”] at 7 (“Conceptually, we make a distinction between the efficiency of the products of the lawmaking process and the efficiency of the legal process itself in providing a certain, complete and predictable set of rules at the lowest cost for society […] We suspect that enhancing the study of the legal process will also advance our understanding of the rules thereby produced”).
4.1. Komesar’s Participation-Centered Approach

Komesar has proposed a comparative analysis of alternative institutions (i.e., politics, courts and markets) that illuminates the limitations of the prevailing scholarship based on single institutional analysis. Komesar’s approach is “participation-centered”: it “identifies the actions of the mass of participants as the factor that in general best accounts for the variation in how institutions function”; it focuses on the cost and benefit functions of those who participate in the decision-making process. Komesar’s general analytical framework illuminates the close connection that exists between social goals and institutional choices; in this respect, it constitutes an important reference point for all those who engage in a comparative institutional discourse.

The analysis developed here differs from Komesar’s work in two important respects. First, Komesar’s framework is based on the assumption that the insights offered by comparative institutional analysis “are true for any analysis and any social goal.” Comparative institutional analysis is “essential for understanding and reforming law and public policy no matter what the social goal or end value contemplated.” Furthermore, “What can be said about efficiency can be said about any goal or end value.” Unlike Komesar, I focus on process values that are specific to the law-making process. Rather than considering efficiency as a possible target among many others, my purpose is to place the efficiency of the law in its own contest: the law-making process. To do so, I investigate the specific relationship between the efficiency characteristics of alternative (and competing) legal processes and the efficiency properties of resulting legal outcomes. Furthermore, Komesar does not precisely identify the costs that are associated with the law-making process. He focuses on individuals’ costs of participating in the process, while my chief concern is how institutional and structural variables affect the costs of producing legal rules.

---

8 Ibid. at 7.
9 Ibid. at 30 [emphasis is mine].
10 Ibid. at 30.
11 Ibid. at 30
Second, Komersar’s work is more descriptive than analytical. That is, it identifies the comparative advantages of alternative institutions but does not always identify the causal mechanism explaining them with sufficient precision. For example, Komesar notes many times the importance of the costs of information sustained by individuals to participate in the political process; however, he does not explain what causes these information costs, nor does he explain the determinants of individuals’ utility function. There is no mention in Komesar’s work of the fundamental dimensions of law-making that decisively affect the incentives and constraints of the actors involved in law-making. By comparison, my suggested approach pays close attention to the institutional mechanisms generating the information costs of both the lawmaker and the people subject to law (e.g., ex-ante versus ex post dimension, degree of centralization of law-making, inter-jurisdictional competition and so on). In many cases, while my conclusions are not in conflict with Komesar’s work, I attempt to provide an in-depth discussion of the explanatory mechanism justifying them.

In conclusion, I share with Komesar the fundamental idea that the choice of the efficient outcome is necessarily grounded on some assumptions about the relative merits of alternative institutions. But I focus more specifically than he does on the law-making costs, the institutional and structural variables affecting them and the explanatory mechanisms of the comparative advantages of alternative institutions.

4.2. Parisi’s Investment-Theory Approach

In their recent monograph *Economics of Law-making*, Parisi and Fon examine “the relative advantages and the respective limits of alternative sources of law.”12 One important set of tools they use to investigate the structure of law-making costs is provided by modern investment theory. From this methodological perspective, the production of legal rules is conceptualized as an economic investment in which the lawmaker sustains present costs in view of future uncertain benefits. This analytical perspective provides important insights on law-making through “politics,” “adjudication” and “practice.” I will

---

refer to Parisi and Fon’s contribution many times throughout the discussion; here I briefly identify some differences between their analytical perspective and mine.

First, Parisi and Fon focus on *specific* features of alternative law-making processes; although their work provides many insights useful for comparative institutional analysis, their contribution does not aim at providing a general comprehensive framework for the analysis of the law-making process. By contrast, as mentioned earlier, the main goal of this thesis is to provide a general unified framework for understanding and assessing the law-making process. Second, while Parisi and Fon conceptualize the law-making process in terms of resource investment, I view the structure of the law-making process as a response to a chain of collective action problems and related transaction costs. Therefore, my methodology is predominantly based on transaction-costs economics, agency-theory and evolutionary economics, rather than on investment theory. Third, Parisi and Fon seem to view efficiency as a cost-minimization problem: they identify certain costs associated with law-making and investigate the institutional and environmental conditions that minimize them. From their perspective, the problem of outcome efficiency is separated by this problem of cost-minimization. By comparison, according to my suggested approach, the efficient law-making process is chosen not only based on its cost-minimizing properties, but also for its ability to enhance the efficiency properties of the legal outcome. This latter point will become clearer as the discussion proceeds.

5. Organization of the Inquiry

5.1. Definitions

The preliminary obstacle to this research agenda involves problems of definitions and terminology.

Efficiency

Reducing the complexity and variety of forms of rule-producing processes to a single, abstract value is a heroic project. I wish to reassure the reader at the outset that this is not my purpose. The distinction between efficiency and alternative (and often competing)
values is only an aid to analytical tractability. In fact, many different values are involved in the operation of the law-making process, and it is not sustainable in practice to single out one single value. Thus, I do not suggest that economic analysis provides the sole or dominant method for understanding the institutions under review. I do believe, however, that economic analysis helps explain several important features about law and legal decision-making institutions. In particular, it illuminates the trade-offs that are at work among different value-laden goals and provides criteria for choosing among alternative institutional arrangements with different consequences for the efficiency of the legal outcome.

Approaches
This thesis draws from different approaches to legal-economic analysis. For brevity, I group them into two groups. The first approach has come to dominate law and economics scholarship since the early development of the discipline; I synonymously call it “neoclassical,” “market-based,” “traditional,” “conventional,” “mainstream” and the like. It is built upon the neoclassical conceptual framework of the competitive market, and is organized around the key concept of allocative efficiency. In essence, it calls for cost-benefit analysis as the foundation for the assignment of legal entitlements. I will discuss some strengths and limitations of this approach in Chapter 1.

A heterogeneous group of alternative approaches has been developed over time based on the methodology of institutional economics.13 I refer to this array of approaches by using the sweeping label of “institutionalism.”14 It emphasises the mutual interdependence of economic and legal processes, and focuses on the institutional environment in which economic activities take place, rather than on the final allocative results attained through trade and exchange.15

13 Although the array of relative contributions is considerably heterogeneous, for the purpose of my analysis I group them under the label of “neo-institutional law and economics.”
14 It is possible to identify the following three distinct traditions of analysis: rational choice institutionalism, sociological institutionalism and evolutionary institutionalism.
In general, I view these two approaches as complementary, rather than competing; taken together, they help to unlock the black box of the legal-economic relationship and shed light on how legal institutions affect the economic performance of individuals and society\textsuperscript{16}. From this perspective, my research agenda may be viewed as an attempt to integrate these two approaches by showing the analytical advantages of combining the two perspectives.

I refer to my proposed methodology as a “structural” approach to the law-making process. It is grounded in the institutional tradition of analysis and, as such, it gives primary emphasis to the specific features of the institutional environment influencing the behavior of individuals involved in the law-making process. The structural approach (i) is micro-founded in that it explains the law formation process by connecting the decisions of individual actors involved in the legal process to the macro-phenomenon of the production of law; (ii) relies on rational choice and evolutionary theory; (iii) is premised on the key organizational concept of process efficiency and (iv) combines alternative styles of analysis, including “public choice analysis,” “transaction costs economics” and evolutionary law and economics.

Law-Making

This study focuses on the creation of legal rules. I adopt a broad definition of legal rules (and law) that encompasses all formal or informal rules that give rise to regularity in the behavior of individuals. From this broad perspective, I include in the scope of my analysis all the institutions producing rules affecting individual behaviour. I do not restrict the analysis to authoritative public law-making institutions (such as legislatures and courts); rather, I extend the notion of law-making to both private and unintentional processes, such as, for example, the emergence of customs and social norms.

Institutional and Structural Variables

\textsuperscript{16} I will discuss the methodological features of the structural approach in Chapter 1.3 and in Chapter 3.
The central claim of my research agenda is that the efficiency of the law is a function of institutional and structural variables affecting the law-making process. The *institutional* variables are characteristics of the law-making process that are predicted to be most important in affecting law-making costs. The *structural* variables are the characteristics of the regulated environment that are predicted to be most important in determining the relative advantage of alternative sources of law.

The crucial difference between institutional and structural variables is that while the former are choice variables, the latter are exogenously given environmental conditions.

**People Subject to Law**

When using the expression “people subject to law,” I refer to the following two sets of individuals:

(i) those whose behaviour falls into the scope of application of the law (hereinafter “target actors”);

(ii) those whose interests are protected by a given set of legal rules (hereinafter “law beneficiaries”).

A crucial point is that the relationship between the people subject to law and the lawmaker is a *circular* one. The people subject to law are concurrently the individuals that participate in the production of law (e.g., participating in a bargaining process by litigating before a judge or taking part in the political-electoral process) and those who are subject to the legal rules. That is, their role is twofold because they concurrently influence the *input* and are affected by the *outcome* of the law-making process. This gives rise to a mechanism of *circular* causation that is one of the major sources of the complexity of the economics of law-making.\(^{17}\)

\[\text{\textsuperscript{17}}\text{The concept of circular causation has been introduced by Gunnar Myrdal, } An\ American\ Dilemma: The\ Negro\ Problem\ and\ Modern\ Democracy (New\ York,\ Harper: 1944).\]

**5.2. Analytical Core**

I now briefly synthesize the logical path followed throughout this study.
5.2.1. Output Efficiency

An “output bias” underlies the conventional method of legal-economic inquiry that often results in incomplete scholarship. The incompleteness manifests itself at two levels. First, output efficiency analysis does not include law-making costs in the efficiency calculus. If the law-making costs exceed the magnitude of the surplus achieved, either the efficient legal rule is unattainable in practice or the production of the efficient rule generates a net utility loss for the society. In both cases, the availability of a theoretically efficient outcome does not result in any improvement in terms of social efficiency. Second, the exclusive focus on the output loses sight of the variables affecting the legal outcome that are located outside the analytical domain of the outcome itself. Third, the output-oriented approach is susceptible to the following criticisms: (i) it suffers from logical circularity and/or logical incompleteness; (ii) does not guarantee the identification of increases in social welfare; (iii) it does not account for losers, thereby exhibiting a status quo bias; and (iv) does not account for the dynamics of the legal change.

5.2.2. Process Efficiency.

It follows from the preceding considerations that a full understanding and a logically consistent assessment of legal rules requires the process-outcome relationship to be regarded as the fundamental unit of analysis of the law and economics discipline. This has a fundamental methodological implication. Since the concept of allocative efficiency is not suited for the analysis of the law-making process,18 a notion of economic efficiency is needed that is suitable to the analysis of the law-making process. Thus, the first methodological step is to develop a notion of process efficiency. According to my definition, a law-making process is efficient if there is no other alternative institution that does better, across the circumstances in which it actually operates, in the production of social welfare-enhancing legal rules. This requires the law-making process to efficiently solve the following four problems that are invariably associated with the production of legal rules: (a) information, (b) agency, (c) social choice and (d) the adaptation problem.

---

18 See Alfred Allan Schmid, Conflict And Cooperation. Institutional and Behavioral Economics at 82 (“To use [allocative] efficiency as a basis for institutional choice is to be caught in a ‘value circularity problem’”).
As we will see, these four problems plague every law-making mechanism, thereby preventing the production of efficient legal rules.

To summarize, an efficient law-making process has comparative advantages over its alternatives in minimizing the costs associated which each of these four sources of inefficiency in the production of law. A significant problem is how to trade-off these four criteria of efficiency. The only way of dealing with this problem is to assess on a case-by-case basis the relative salience of each of the problem identified above (e.g., in some cases the information problem will be more salient than the others; in other cases the adaptive problem will be the most salient, and so on).

5.2.3. Law-making Costs

To make predictions from the theory, one needs to identify the costs associated with the law-making process and the variables that cause these costs to vary. The identification of costs is logically related to the idea of efficiency that one adopts. I thus proceed to identify a unified taxonomy of the law-making costs based on the analysis of the four problems identified above. The taxonomy is intended as a tool for organizing the inquiry into alternative law-making processes and identifying the mechanisms generating inefficiency in the production of law. Importantly, *I do not identify a unique cost-minimizing point of the law-making process*. As we will see, the law-making cost-functions depend on a high number of variables whose salience varies across institutions and situations. The specific relevance of the variables in one context differs from the specific relevance of variables in another context. For this reason, my suggested taxonomy should be viewed as a framework of analysis that helps to identify the fundamental stages of analysis in explaining and assessing the relative merits of alternative law-making processes on a case-by-case basis.

5.2.4. Supply of and Demand for Law

Once the notion of process efficiency has been identified and the costs of law-making have been clearly defined, the final step to build a methodology is to identify the variables that are predicted to be most important in affecting the structure of the law-making costs, thereby determining the relative efficiency of alternative sources of law. To
accomplish this task I organize the line of inquiry along the lines of a supply of and demand for law. On the one hand, the supply side summarizes the features of the law-making process that affect the structure of the law-making costs (i.e., the institutional variables). On the other hand, the demand side summarizes the features of the situation from which the need for law arises. This includes both (i) the exogenous conditions of the regulated environments and (ii) the preferences, incentives and constraints of the people subject to law. The central hypothesis of my research agenda is that the efficiency of law depends upon the interplay between the demand for and supply of law. More precisely, it is the interaction between institutional and structural variables that most affects the efficient properties of the outcome.

Figure 1 summarizes the theoretical framework for the analysis of the process-outcome relationship that constitutes the methodological underpinning of my research project.
5.2.5. The Variables Affecting Process-Efficiency

Once I have developed the tools for investigating the process-outcome relationship, I organize my research agenda along two gradually decreasing levels of abstraction. This choice is not just a matter of convenience; it reflects the methodological assumption that the best way to construct a sound explanation of social phenomena is to proceed gradually from higher to lower levels of abstraction.\(^\text{19}\)

At the highest level of abstraction I develop a four-dimensional model of law-making that focuses almost exclusively on four variables (two institutional and two structural variables), which I call “fundamental dimensions of law-making.” On the supply side, the

\(^{19}\) In general terms, the employment of a higher level of abstraction represents a more powerful analytical instrument, and a lower level of abstraction allows a more reliable empirical inherence. See on this point Siegwart Lindenberg, “Constitutionalism versus Relationalism: Two Versions of Rational Choice Sociology” in Jon Clark, *James S. Coleman* (London: The Falmer Press, 1996) chapter 18; Siegwart Lindenberg, “Social Rationality versus Rational Egoism” in *Handbook of Sociological Theory* (New York: Plenum, 2002).
first institutional dimension of law-making is related to the time perspective of the production of legal rules with respect to the object of regulation. From this view, we can distinguish between ex-ante and ex post law-making. The ex-ante approach is forward looking, lays down rules in advance of the regulated activity and proceeds by considering abstract classes of cases producing legal consequences. The ex post approach is backward looking, occurs at a time in which the regulated activity has already taken place and is necessarily case-specific. The second institutional dimension of the law-making process is the degree of centralization. On the basis of this factor we can distinguish between centralized and decentralized law-making processes. As to the demand side, I assume that the degree of frequency and homogeneity of the behaviour subjected to regulation are the fundamental structural variables that determine the optimal mix between ex-ante and ex post regulation and centralized versus decentralized law-making.

The four-dimensional framework of law-making enables me to formulate some general hypotheses on how law-making costs are affected by the characteristics of the law-making process and the features of the regulated environments. These hypotheses are tested at a lower level of abstraction through the analysis of the following four alternative models of law-making institutions: politics, bureaucracy, adjudication and spontaneous law-making. Each of these law-making institutions represents a combination of ex-ante versus ex post and centralized versus decentralized dimensions with additional institutional features. Politics is an ex-ante centralized process based on the mechanism of political representation. Bureaucracy is an ex-ante centralized process in which the lawmaker is, to varying degrees, insulated from the political decision-making. Adjudication is an ex post law-making process characterized by varying degrees of centralization, the doctrine of precedent and the settlement-trial decisions of private parties. Finally, spontaneous law-making is based on repeated private bargaining, which is a form of ex-ante decentralized law-making. The analysis of politics, bureaucracy, adjudication and spontaneous law-making allows the identification of the interactions between the fundamental dimension of law-making and additional institutional and structural variables in determining the efficiency properties of legal rules.
5.3. Hypotheses

As already emphasized, the central hypothesis of my research agenda is methodological. I aim to demonstrate that the efficiency of the law is better explained as a function of institutional and structural variables affecting the law formation mechanisms rather than the allocation of legal entitlements insulated from the law-making process.

Once I have clarified the methodological standpoint of the thesis, I identify some hypotheses suggested by the four-dimensional law-making framework. I list them through the following summary statements whose meaning and relevance will become clearer as the research proceeds.

(i) Under conditions of high homogeneity and frequency of the demand for law, the ex-ante centralization of law-making allows for significant economies of scale effects that significantly reduce the average (production and compliance) information costs.

(ii) The ex-ante centralization of law-making entails a significant increase in the agency costs of law-making.

(iii) The various combinations of ex-ante versus ex post and centralized versus decentralized law-making differentially affect the structure of the social choice problem confronted by the lawmaker. Ex-ante law-making (independent of the degree of centralization) has relative advantages over ex post processes with respect to the problem of empty core bargaining (i.e., intransitive cycling). Centralization tends to exacerbate the social choice problem.

(iv) The ex-ante centralization of law-making exerts a twofold effect on law-making costs. It reduces the magnitude of adaptive and information costs associated with legal change on the one hand and increases the magnitude of resistance and maladaptation costs on the other.
These hypotheses will be tested against the results of comparative analysis of alternative law-making institutions. In particular, the analysis of politics, bureaucracy, adjudication and spontaneous law-making suggests the following hypotheses that capture some core tendencies in the production of legal rules.

(i) *Ex-ante centralization allows for significant economies of scale effects. However, in politics and bureaucracy this advantage should be weighed against the costs associated with, respectively, (a) the high irrationality and limited information capabilities of the political decision-making process, and the tendency toward (b) overregulation (c) x-inefficiency and (d) output ineffectiveness of bureaucratic processes.*

(ii) *Pure decentralized spontaneous law-making entails relative advantages in terms of productive efficiency under the restrictive conditions of close-knittedness.*

(iii) *Politics and bureaucracy exacerbate the increase in agency costs associated with ex-ante centralization. Adjudication in centralized “made orders” entails higher agency costs than adjudication in decentralized “grown orders.” Spontaneous law-making decreases the agency costs significantly (although it is not immune from the agency problem).*

(iv) *Every law-making mechanism can be subject to the problem of intransitive cycling. Law-making is always path-dependant and structurally induced. However, the social problem affects law-making institutions differentially. The “reason-based” nature of judicial decision-making is a source of comparative advantages of adjudication versus politics with respect to social choice efficiency.*

(v) *The comparative analysis of alternative law-making institutions confirms that the ex-ante centralization of law-making reduces the adaptive and information costs associated with legal change; at the same time, however, it increases the resistance and maladaptation costs.*
6. Limitations

As already emphasized, this study is admittedly preliminary in nature. This implies that the discussion does not consider a large number of issues that will have to be considered at a further stage of the investigation. In this respect, it is useful to emphasize at the outset the major limitations of my proposed line of inquiry. First, *I assume that individuals’ preferences and expectations are exogenous to the analysis.* I do not address the issue of how the choice among alternative lawmaking institutions affects the formation process of individuals’ preferences and expectations over policy outcomes. Second, in order to be able to focus on a limited number of institutional variables, *I discuss four over-simplified models of lawmaking.* A full assessment of the relative advantages of alternative sources of law in real-world situations would require taking into account the jurisdiction-specific nature of lawmaking institutions (e.g., common-law versus civil law systems, parliamentary versus presidential constitutional regimes, and so on). Third, *I do not explicitly discuss the role of lawyers in the production of law.* I implicitly recognize the role played by them in the discussion of private legal orderings; but a comprehensive analysis of the production of law will require an in-depth investigation of lawyers’ role in determining the direction of the legal change. Fourth, *I analyze each lawmaking institution in isolation from other sources of law.* That is, I do not consider the dynamic interaction among lawmaking institutions (e.g., judicial interpretation of statutory rules, judicial review of bureaucratic action, the impact of the statutorification of law on other sources of law, and so on). However, the study of lawmaking should take into consideration that in real-world situations the production of law is strongly affected by the way alternative lawmaking institutions interact with each other. Fifth, *when I analyze the bureaucratic process I limit the discussion to information and agency efficiency, and do not consider social-choice and adaptive efficiency.* These two latter problems will deserve close attention at a later stage of the analysis. Finally, it must be recognized that many institutional and structural variables considered in this study affect the cognitive processes of the actors involved in the lawmaking process. In particular, different kinds of cognitive biases and heuristics affect alternative decision-
making processes differentially depending upon the structure of the lawmaking institutions. For the sake of brevity, I leave this important issue to future research.

7. Summary of Contents

The thesis is organized as follows. Part I is purely methodological. It discusses the reasons for a shift of focus from outcome efficiency to process efficiency. First, it examines the logical features of the concept of efficiency and the related theoretical difficulties of conventional outcome-oriented analysis (chapter 1). Second, it develops the notion of process efficiency, analyzes its relationship with the substantive efficiency of legal rules and constructs a uniform taxonomy of law-making costs (chapter 2). Part II develops a four-dimensional framework of the law-making process that I have briefly outlined above (chapters 3-4). Part III develops a comparative analysis of alternative law-making institutions (i.e., politics, bureaucracy, adjudication and spontaneous law-making); in so doing it tests the hypothesis formulated in part II and identifies core tendencies in the production of legal rules (chapters 5-8). The concluding chapter provides an integrated analysis illustrating the main results of the study (chapter 9).
Part I

METHODOLOGY

Chapter 1 discusses the problematic nature of economic efficiency in the legal context. Since economic efficiency results from the choices made at the institutional level, the decision-making processes at work in the ‘legal–economic nexus’ should be investigated in order to identify its main determinants. Based on this premise, the subsequent discussion focuses on the predominant normative efficiency standards used in law and economics (i.e. Pareto optimality, Kaldor-Kicks efficiency, wealth maximization and social welfare maximization). The analysis of these criteria demonstrates that their major limitations in conventional law and economic analysis depend on the exclusive focus on the allocative outcome.

Chapter 2 illustrates the nature of process efficiency analysis. It demonstrates that output efficiency and process efficiency are related to each other and that their relationship should be understood from the standpoint of social welfare maximization. This chapter shows that the neoclassical theory of production does not adequately explain the production of legal rules; the chapter therefore emphasizes the need for a more appropriate methodology of inquiry into the process–outcome relationship. Based on this premise, the last two sections in the chapter develop the notion of process efficiency and propose a unified taxonomy of lawmaking costs, both of which comprise the methodological tools used in the subsequent analysis of the lawmaking process.
Chapter 1
EFFICIENCY AS A LEGAL CONCERN

Trying to measure efficiency is like trying to pull oneself out of quicksand without a rope. There is no firm ground.

Deborah Stone, Policy Paradox: the Art of Political Decision Making, 2001

This chapter examines the difficulties of importing the concept of economic efficiency into legal discourse and provides an introductory definition of the three notions of economic efficiency widely used in economics—allocative, productive, and dynamic efficiency. A discussion of these points is important for two reasons. First, it clarifies the logical foundations of the central claim that the output-oriented approach needs to be integrated with process-oriented analysis. Second, it provides preliminary indications as to how to construct the proposed complementary approach in a theoretically consistent manner.

Introduction

I define traditional law and economics scholarship as either “neoclassical” or “output-oriented.” I use these labels to emphasize two characteristics of mainstream law and economics that in my view constitute its major criticism. First, the conventional approach is output-oriented in the sense that scholars devote great attention to the output of the law-making process, while often losing sight of the legal-economic consequences arising from the structure of the law-making process itself. I also indicate this attitude with the term “output bias” to emphasize its distortive effect on the results of efficiency analysis. Second, conventional law and economics is neoclassical in the sense that it imports the methodology of neoclassical economic theory into the legal context. A serious weakness of this methodological approach is the adoption of the positivistic attitude underlying
neoclassical efficiency without careful adaptation to the characteristics of the legal environment. While the assumptions of the neoclassical model reflect the idea of economics as a value-free discipline, the legal context is characterized by the production and exchange of resources that are inherently normative in nature (i.e., legal entitlements).

To remedy to the preceding two shortcomings, I propose a structural approach to law and economics. I call the proposed approach “structural” to emphasize the focus on the institutional structures and regulatory environments within which legal rules are produced, as opposed to mainstream output-oriented scholarship. In the context of the structural approach, the notion of “process efficiency” captures the causal connections among structures of the law-making process, characteristics of regulated environments, and efficiency of substantive legal rules. I call “substantive efficiency,”'efficiency in the allocation of legal entitlements as traditionally investigated by conventional law and economics. In addition, I refer to the constitutive elements of process efficiency as the “procedural determinants” of substantive efficiency; thus, I emphasize the causal connection between the institutional structures of law-making and the efficiency properties of substantive legal rules.

In this chapter, I raise two specific criticisms of the neoclassical output-oriented approach. First, it exposes legal–economic discourse to the risk of logical circularity. To avoid the risk of tautological propositions, economic analysis of law should carefully identify the normative specifications implicit in the notion of efficiency. A discussion of this point allows for concurrent clarification of the logical foundation of my proposed approach. Second, the outcome-oriented approach does not properly consider the implications in terms of productive and dynamic efficiency of alternative legal regimes. I contend that the focus on allocative efficiency without a complementary analysis of productive and dynamic efficiency leads to unreliable theoretical conclusions. I do not intend to deny the usefulness of conventional scholarship, but to complement mainstream output-oriented analysis with a complementary process-oriented approach. Ultimately, the goal of this research is to provide the methodological tools for a comprehensive legal
and economic approach based on the “process-output” dichotomy as the fundamental unit of analysis.

The discussion in this chapter proceeds as follows. Section 1.1 discusses the logical features of the concept of efficiency. In particular, it clarifies that efficiency is a relative concept and a comparative criterion. These latter two characteristics constitute possible sources of logical circularity that may undermine the theoretical consistency of the efficiency discourse. Section 1.2 briefly defines the three notions of efficiency widely used in economic discourse. In particular, subsection 1.2.1 defines the concept of allocative efficiency, which represents the standard criterion employed by positive law and economics. Subsections 1.2.2 and 1.2.3 briefly clarify the concepts of productive and dynamic efficiency. Sections 1 and 2 are related. As will become clearer subsequently, the shift of focus from the outcome-oriented approach to the process approach entails a parallel shift of focus from allocative to productive and dynamic efficiency. The notion of process efficiency, which I will develop in the next chapter, is the methodological tool for incorporating the analysis of productive and dynamic efficiency of law-making into positive and normative law and economics discourse. Finally, section 1.3 examines the limitations of three alternative normative standards employed in legal-economic analysis: Pareto optimality, Kaldor–Hicks efficiency, and Wealth maximization.

1.1. Logical Features of Legal Efficiency

The easiest way to approach the notion of efficiency is to consider its most practical implication. In this respect Frank Knight indicates the search for economic efficiency and “reduction of waste” as relevant contributions afforded by the economics discipline to social policy analysis. He defines efficiency as “avoidance” or “reduction” of “waste”. This simple definition tells us what efficiency is not (i.e., the wasteful use of resources),

---

20 See Frank H Knight, “Review of Melville J. Herskovits’ ‘Economic Anthropology’” (1941) 49 J. Pol. Econ., 252, [Review] (“[...] men in general, and within limits, wish to behave economically, to make their activities and their organization ‘efficient’ rather than wasteful. This fact does deserve the utmost emphasis; and an adequate definition of the science of economics [...] might well make it explicit that the main relevance of the discussion is found in its relation to social policy, assumed to be directed toward the end indicated of increasing economic efficiency, of reducing waste.”)

21 Knight, Review, supra note 20.
but it does not provide a meaningful notion of efficiency for positive or normative analysis. To capture the full meaning of efficiency, two logical features must be recognized. Let me briefly introduce them through the following summarizing propositions. (i) Efficiency is a relative (or relational) concept inevitably based on some normative antecedent specifications; it becomes operative upon a clear definition of the goals to be pursued and upon a precise identification of the people whose interests count. Neglect of the relative nature of efficiency leads to tautological propositions. (ii) Efficiency is a comparative criterion. It guides choices over a set of alternative ways of doing things, under the constraint of a limited amount of available resources; neglecting this aspect leads to the Nirvana fallacy, one of the major methodological pitfalls from the viewpoint of an unwary legal theorist.

In the following sections, I discuss each of these characteristics along with their methodological implications.

1.1.1. Efficiency as a Relative Concept

Institutional law and economics has long recognized that the use of the efficiency criterion inevitably presupposes an antecedent normative specification. Efficiency is a “relative condition”, and “is not a function of any intrinsic substance or relation per se”. The relative nature of efficiency has been at the same time the most attractive

---


24 Ibid. at 109. In more philosophical terms, Peter Greach well explains the difficulties associated with the use of relational concepts: “Relational concepts in general raise difficulties of which abstractionists are usually unaware […]. The trouble is that a relation neither exists nor can be observed apart from its converse relation; what is more, the concept of a relation and of its converse is one and the same indivisible mental capacity, and we cannot exercise this capacity without actually thinking of both relations together; relativa sunt simul natura et intellectu. It is difficult, in view of this, to give any account of our understanding the difference between “the knife is to the left of the book” and “the knife is to the right of the book”. But I do not see how an abstractionist of it is even possible; for all that is obtainable by abstraction, one would think, is ability to recognize the “recurrent characteristic” of right-
feature of economic analysis and the source of major disagreements among different schools of economic thought. In particular, neoclassical economists tend to see the concept of efficiency as the methodological expedient through which economic analysis is freed from value judgments. This positivistic attitude is grounded in the search for a scientific foundation of economic analysis, which is viewed as an empirical science separated by any normative concern. However, despite the claim of neutrality, most optimality reasoning conducted by economists rests tacitly on some implicit (sometimes hidden) normative premise. With specific reference to law and economics scholarship, a large part of conventional analysis is based on implicit normative premises that often are not assumed for precise methodological reasons, but stem from an unawareness of the relative nature of legal–economic efficiency. In particular, I identify the following three

left ordering, not ability to tell which thing is to the left and which to the right. [...] it is only if there is at least an ability to tell which thing is to the right and which to the left, that we can say there exists a concept of right and left”, in Mental Acts. Their Content and their Objects (New York: Humanity Press, 1957).


Friedman, “Methodology”, supra note 25


For some of the major theoretical contributions to the debate on efficiency as a legal concern, see “Symposium on Efficiency as a Legal Concern”, (1980) 8 Hofstra L. Rev. 485, including articles by Richard A Posner, Jules L Coleman, Ronald M Dworkin, Lewis A Kornhauser, Richard S Markovits, Lucian A Bebchuk et al.

theoretical issues that are generally not considered in standard efficiency analysis of legal rules:

(i) “Efficiency is rights-structure specific; it is a partial function of law”;  
(ii) the determination of rights is an inherently normative activity; separation between efficiency and distribution is highly problematic in legal discourse; 
(iii) efficiency is usually conceived as a technical criterion based on the means given the ends; however, the means–ends relationship itself is part of normative discourse and content changes in different institutional settings.

The failure to consider these three points carries the risk of falling into a logical circularity problem.  In the following sub-sections, I elaborate on these three important points and emphasize the unavoidable normative dimensions of any efficiency analysis.

1.1.1.1. Efficiency as a Function of Rights

Warren J. Samuels has devoted much of his academic scholarship to investigating the complexity of what he labelled “the legal–economic nexus”.  His account of the relationship between economics and the structure of the legal environment helps illuminate the difficulties of adopting efficiency as a tool for legal analysis. First, Samuels accords primary importance to the process through which the outcomes of legal systems come into existence. In his view, economic performance and legal rights are

---


29 The problem of logical circularity of efficiency analysis in the legal context has long been identified by institutional economics; see bibliography indicated at note 6. However, legal scholars outside the institutional law and economics movement as well have recognized the risk of circularity associated with the use of efficiency in the legal discourse; see, for example, Leff, “Economic Analysis”, supra note 27.
jointly produced, and both are the result of choice processes that take place at various institutional levels. The importance of economic efficiency is not rejected; rather, an investigation of the decision-making processes at work in the legal-economic nexus allows for a better account of economic efficiency. Second, economic efficiency is conceptualized as a function of rights. The allocation of legal entitlements determines the fundamental variables of an economic system (wealth distribution, prices, costs, and so on); thus, efficiency is right-based as opposed to legal rights being efficiency-based. From this premise, it follows that there is no unique efficient solution and, depending on the structure of rights, it is possible to identify different efficient solutions. These assumptions provide the basis for a rights-based conception of efficiency that recognizes that law and economics are neither independent nor merely interacting; rather, they simultaneously arise in a variety of complex decision-making processes.

1.1.1.2. Efficiency and Distribution

Conventional economic analysis focuses on marginal improvements in allocative efficiency. Marginal analysis presupposes that the initial distribution of wealth is determined exogenously. However, when imported into the legal context, this assumption becomes a key source of difficulties for standard efficiency analysis. While it might be

31See Samuels, “Nexus”, supra note 30 at 1577: “This essay affirms the existence of a fundamental legal-economic nexus in which certain behaviour is simultaneously legal and economic, in which the two nominally different processes are actually aspects of one fundamental process. From this perspective, the idea of the self-subsistent existence of economy and polity seems naive and incomplete. The central element of this legal-economic nexus is governance, in the sense of a process in which important decisions are made, whether by legislatures, courts, or administrative agencies; by giant manufacturing corporations, cartels, trade associations, pension funds, major banks, and so on; or by alliances of governmental institutions and private organizations. […] the emphasis must be on the process by which actors who are simultaneously both economic and political actors themselves work out solutions to problems that are simultaneously both political and economic”.

32See Mercuro et al. “Institutional Law and Economics” supra note 22 at 440: “[…] economic activity - prices, costs, outputs, risk, income, wealth, and so on - is not some sort of natural phenomenon, but rather is determined by the structure of rights that exists in society, with the levels of and changes in each of these variables being in part a function of the legal structure and legal change over time […]. Each particular rights structure will give rise to a particular set of prices, costs, outputs, and so on and thus to a particular efficient allocation of resources. Hence, there is no unique efficient result. […]. Different specifications of rights will lead to different (and economically non-comparable) minimizing or maximizing valuations. The result is that an outcome that is claimed to be efficient is efficient only with regard to the assumed initial structure of rights […] the latter of which is often the very matter at issue”). See also Robert Cooter, “Liberty, Efficiency, and Law” (1988) at 142 [“Liberty”]: (“An initial distribution of resources is the Archimedean point by which efficiency analysis is levered. If this point detaches itself and moves, the comparison of alternative policies in terms of their relative efficiency becomes incoherent”).
well justified in the context of neoclassical models, the assumption significantly undermines the validity of efficiency analysis in the legal context because the formation of the initial wealth distribution constitutes the object of the choices that are at work in the legal–economic nexus. Thus, assuming the initial wealth distribution as exogenous leaves out of the analysis the institutional processes generating the baseline of marginal efficiency. By contrast, the shift of focus from marginal efficiency improvements to mechanisms generating legal entitlements allows us to internalize into the economic analysis the institutional processes through which the antecedent normative specifications of efficiency are formed. I briefly expand on this point.

In the field of welfare economics, the separation between efficiency and distribution has been the methodological tool for avoiding the problem of interpersonal comparison of utility or welfare.\(^{33}\) Pareto efficiency, as applied by welfare economists, does not take into account any distributional concern and is exclusively based on allocative considerations. In the same perspective, Kaldor-Hicks efficiency is premised upon the idea of separating efficiency and equity so as to avoid interpersonal comparisons of utility.\(^{34}\) Based on the latter view, the measure of economic efficiency is the change in aggregate gains, and distributional concerns are evaluated on a separate basis. The underlying idea is to free economic analysis from the value judgments that are typically related to distributional questions—involving subjective moral evaluations—thereby conferring a “scientific” status to the economics discipline.

Mainstream law and economics scholarship has largely adopted the separation between efficiency and distribution.\(^{35}\) Although distributional concerns (i.e. the presence of

---

\(^{33}\) Distributive issues entail comparison of the welfare gains and losses that different individuals receive by the resource reallocation. However—as I will explain subsequently—since it is difficult to find general agreement on a priority scale of social values to assess distribution policies, economists have generally assumed a methodological separation of efficiency from distribution. On this point, see the classical contributions of Robbins and Freidman cited in note 9.


\(^{35}\) For a useful overview on the debate on the positive/normative nature of economic analysis of law, see, Hendrik Kerkmeester, “0400 Methodology: general” in de Geest, “The History and Methodology of L&E” supra note 22 at 390-391 and the bibliography there cited.
“winners and losers” as a consequence of legal change) have not been neglected entirely, the prevailing approach has been firmly based on the assumption that allocative efficiency and distributional equity should be achieved at different functional stages. One task involves understanding the initial distribution of normative resources; another task is to show how to maximize the value of said resources. While the former is a task for philosophers, historians, or legal theorists, the latter is a proper object of efficiency analysis of law that ought to be concerned only with marginal improvement with respect to a presumptive wealth distribution. Consistent with this perspective, the chief concern of normative law and economics has been traditionally identified with wealth maximization, which is firmly grounded in the efficiency-distribution dichotomy.

In sum, the prevailing law and economics approach has imported the welfare economics conceptual apparatus without carefully adapting it to the specific characteristics of legal discourse. Unlike many of the economic goods described by the neoclassic models, legal rights are normative resources, i.e. resources that cannot be allocated or re-allocated without entailing normative evaluations at the systemic level. The theoretical weakness of importing the efficiency–distribution dichotomy into the legal context is easily recognizable once a right-based conception of economic efficiency is adopted. In the light of this view, legal rights determine the initial distribution of wealth that law assigns to people; hence, the problem of the efficiency of legal entitlements cannot rest exclusively on the assumption that the bundle of rights that constitutes parties’ initial endowments is exogenous to the model.

In conclusion, within the legal context, efficiency and distribution are inextricably interlinked. Since economic efficiency is the result of choices made at the institutional level, its determinants are located in the same domain as the initial distribution of wealth (assumed as given by mainstream economists). This analytical point marks the difference between the neoclassical and institutional approaches. While the latter encompasses the processes determining the initial set of allocative choices in its analytical scope, the former starts its analysis where the legal–economics process ends.

1.1.1.3. Means-Ends Relationship

Thus far, I have emphasized the tension between efficiency as a value-free criterion and efficiency as a relative concept. An analogous tension arises in relation to another cornerstone of neoclassical methodology: the theory of technical efficiency as based on the means-ends dichotomy. This conception is a major theoretical underpinning of positive law and economics: the focus on the means (i.e. resource allocation) enables

---

38 See Nicholas Mercuro and Steven G Medema, Economics and the Law: From Posner to Post-Modernism, (Princeton: Princeton University Press, 1997) [Mercuro and Medema, “Economics”] at 229: “Rights determine the distribution of income and wealth, which in turn determines the efficient solution that is reached—but at the same time, the specification of rights, and the resulting efficient outcome, structure the future distribution of income and wealth in society”. See also supra note 30, and the bibliography there cited.

39 See Robert Cooter, “Liberty” supra note 32 at 142 who recognizes that efficiency is a condition relative to a given distribution but at the same time identifies the proper object of efficiency analysis of law as the issue of wealth maximization (“Since the efficiency analysis of law cannot turn upon itself and prove what it assumes, economists must wonder how they are to understand an initial distribution of normative resources. One approach is to treat an initial distribution as an historical artefact. This approach is congenial to sceptics who accept that norms can be described, but doubt whether they can be explained or justified. Another approach taken in jurisprudence is to offer a philosophical justification and critique of fundamental legal values. This approach is congenial to philosophically inclined legal scholars. Regardless of the approach, once an initial distribution of legal rights and duties is described, the efficiency analysis can show how to maximize its value. This is important in principle and useful in practice. There is, then, a division of labour in legal studies: Jurisprudence or history provides an account of basic normative resources, while economics shows how to maximize their value”).

40 This theory is usually invoked to support the description of positive law and economics as focused on efficiency as a “technocratic principle of unimprovability” immune from value judgments, C. Veljanoski, Economic Principles of Law, (Cambridge: Cambridge University Press, 2007) at 34.

41 The separation between means and ends finds a clear exposition in Lionel Robbins’ influential essay on the nature of economic science; see, Robbins, “An Essay”, supra note 25 at 15 (“The economist studies the disposal of scarce means. He is interested in the way different degrees of scarcity of different goods give rise to different ratios of valuation between them, and he is interested in the way in which changes in conditions of scarcity, whether coming from changes in ends or changes in means—from the demand side or the supply side—affect these ratios. Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.”) Robbins describes the discipline of economics as the study of allocation of scarce resources to competing ends, such ends being beyond the scope of economists’ investigation.
economists to engage in neutral economic analysis without requiring any normative commitment about whether the ends are good. Similar to the efficiency–distribution dichotomy, the means–ends dichotomy is assumed to ensure the “scientific” status of economics investigations. However, the means–end dichotomy is highly problematic when imported into the legal context. Let me briefly discuss its two major criticisms: (i) problem of infinite regress, and (ii) dynamic interaction between means and ends.

(i) Every possible end can become a means to a more ultimate end. For instance, “my end of catching the next train to town is merely a means of transportation to work, which, in turn, is a means towards personal remuneration, and so on”. This raises the logical problem of an infinite regress unless the ends are stipulated clearly before weighting the efficiency of the means against them.

(ii) The means and ends are not always logically distinct, as standard economic theory tacitly assumes. First, “to the extent that means can be made relatively less scarce, or stretched further, more ends or goals of individuals or communities can be realized”. Second, means affect the determination of the ends, which are often defined retrospectively. Thus, assuming ends as exogenous to the economic model entails rigidity in the analysis and leads to the tacit assumption that economic actors have in mind clear and stable objectives to achieve. This in turn leads to the disregard of an informational problem: “ends are not always fully formed and clearly articulated, and often, their ramifications are too complex to give rise to a fully informed choice of means”. In other words, “means can alter ends in the process of achieving them”.

---

42 In that manner, “economists can safely restrict their attention to means, that is, to efficiency, knowing that some other actor or institution will see to the content of the ends”, Avery Wiener Katz, “Positivism and the Separation of Law and Economics” (1996) 94.7 Mich. L. Rev. 2229 at 2264.  
45 See, M Hollis and EJ Nell, Rational Economic Man, (Cambridge: Cambridge University Press, 1975) at 51 (“[…] means indisputably affect ends. As many studies show, large corporations rarely decide to expand, diversify, or merge in order to achieve clearly defined ends. The ends are understood and defined only afterwards, and by then, they often change”).  
46 Ibid.  
47 Ibid.
The foregoing considerations emphasize that the means–ends dichotomy can be employed effectively only to the extent that (i) the ends to be pursued are stipulated, and (ii) the underlying personal preferences are stable. Lacking these two conditions, the assumption that technical efficiency can be evaluated by isolating the means from the ends is likely to be misguided.\footnote{In a famous article discussing the decision-making process of public administrations (“The science of ‘muddling through’” (1959) PAR 79), Charles Lindblom draws a distinction between two alternative approaches: “the root method” and “the branch method”. The root method is based on the choice of the means that best satisfy the goals previously clarified and ranked on the ground of a systematic inquiry into the prevailing values in society. By contrast, the branch method “combines the choice among values and the choice among instruments for reaching values into a single choice”. As Lindblom points out, the first method cannot be employed realistically by public administrations for complex problems; it is much more likely that practitioners will follow the second approach, wherein means and ends are formed by a repeated sequence of mutual interactions. These considerations prove useful in assessing the technical efficiency of the law-making processes and in appreciating the limits and strengths of an analysis based on a pure technical relationship between means and ends.}

1.1.2. Efficiency as a Comparative Criterion

Once ends and means have been specified carefully, the efficiency criterion proves useful for comparing \textit{alternative} uses of resources. In this respect, one major criticism of the outcome-oriented law and economics approach is that it is based on a \textit{single}-institutional analysis rather than on a \textit{comparative} institutional analysis.\footnote{See Komesar, \textit{Imperfect Alternatives}, supra note 7 at 14–28.} The assumption of an exogenous law-making process assumes away the procedural component of efficiency; therefore, the analytical strength of efficiency analysis is halved because comparative assessment focuses only on substantive legal rules while losing sight of the procedural dimension of law. As a consequence, law and economics analysis is often based on the wrong or simplistic assumption that an inefficient legal rule should necessarily be replaced, without careful assessment of the comparative efficiency advantages of the \textit{feasible} alternatives.

Scholars have long emphasized the risks of single institutional analysis. Harold Demsetz coined the expression “nirvana fallacy” to describe a common tendency of many public policy economists.\footnote{Harold Demsetz, “Information and Efficiency: Another Viewpoint” (1969) 12.1 J. Law E. 1.} Theorists commit the nirvana fallacy when they fail to engage in a \textit{comparative} analysis, namely, fail to consider whether an alternative institutional arrangement is better or worse than the existing institution for coping with a given market...
failure. Desmetz described scholars’ tendency to identify a discrepancy between first-best solutions and existing imperfect institutions and then to propose an alternative allocation of the decision-making power, without properly investigating the comparative advantages of the suggested alternative solution relative to the existing one.

In the same vein, but more specifically with reference to law and economics, Komesar demonstrated that the mainstream approach—based on implicit single-institutional analysis—does not consider the relative abilities of alternative law-making institutions to identify the most efficient legal rules. Based on this premise, Komesar called for a comparative institutional analysis that identifies the costs and benefits of three classes of institutional processes: politics, courts, and the market.

In conclusion, paying exclusive attention to outcomes unduly restricts the scope of comparative efficiency analysis. To avoid the pitfalls of nirvana fallacy and single-institutional analysis, the proposed structural methodology is inherently comparative. It aims to integrate substantive efficiency analysis with the results of comparative assessments of alternative law-making institutions.

1.1.3. Theoretically Sound Definition of Efficiency: Avoiding Logical Circularity

The prior discussion provides important insights into the use of economic efficiency for legal analysis. First, any efficiency solution is relative to an initial structure of rights. Second, efficiency depends on a given wealth distribution, which is, in turn, dependent on the rights structure. Third, the logic of efficiency analysis is based on the distinction between means and ends. To avoid circularity, the ends must be specified before assessing the means; to avoid rigidity of the analysis, it should always be kept in mind that the available means retrospectively affect the identification of the ends. Fourth, in the context of legal discourse, efficiency analysis requires an accurate comparative analysis of the relative merits of available alternative institutional solutions.

51 See, Komesar, Imperfect Alternatives, supra note 7
These preceding four analytical steps demonstrate that legal–economic discourse is *positional*, rather than being merely technical. It rests on antecedent normative premises that are internal to legal–economic discourse. The positional nature of legal–economic discourse has important methodological consequences: when the assumptions underlying efficiency analysis are *not* articulated clearly, a serious problem of logical *circularity* arises. Logical circularity deprives efficiency analysis of any meaningful content. As Samuels puts it, in the case of logical circularity “the analyst assumes something about the object to be determined that govern the determination”.\(^5^2\)

To make the normative premises as explicit as possible, any construction of the notion of efficiency should define in advance: (i) the output, (ii) set of alternative choices, and (iii) set of relevant interests.

(i) Output

Efficiency discourse often assumes that (individual and institutional) behaviours are *output*-oriented, that is, behaviours are conceptualized as performances directed towards the achievement of *given goals*. The definition of output is the prime way through which tacit normative premises intrude in a positive discourse and drive the conclusion of what is an efficient solution. The definition of output is a value-laden one, and ultimately determines the conclusion of what constitutes an efficient outcome.\(^5^3\) Samuels makes this point clearly: “The definition of output [...] is one of the most subtle ways through which an analyst or policymaker assumes something about the object to be determined, which governs the determination and which becomes the subject of profound controversy”.\(^5^4\) It follows that a theoretically sound definition of efficiency requires clarification of the normative specifications of the definition of output as the object of maximization.\(^5^5\)

---

\(^{5^2}\) Samuels, “Normative” *supra* note 22 at 100.

\(^{5^3}\) See, Mercuro at al., “Economics” *supra* note 38 at 228.

\(^{5^4}\) Samuels, “Normative” *supra* note 22 at 103.

\(^{5^5}\) See, Allan A Schmid and James D Shaffer, “Marketing in Social Perspective” in Vernon L. Sorenson ed., *Agricultural Market Analysis*, (East Lansing: Michigan State University Bus. Stud., 1964) at 29 (“Economic theory provides a method of calculating positions of maximum efficiency or optimum advantage [...] These calculations are valid, however, only within any given set of exchange system rules which defines the qualitative makeup "analysis [...] developed for the current debate over of the inputs and outputs to be included").
(ii) Choices

Efficiency is defined in relation to a specific set of alternative options. There is no efficiency without a set of feasible alternatives. It follows that a theoretically sound definition of efficiency requires a clear definition of the set of available alternatives. In particular, with respect to the analysis of institutions, the imperfection of an institution does not automatically entail its inefficiency. An institution is inefficient only if there is a less imperfect viable alternative. Hence, it is important to identify the set of alternative institutional choices.

(iii) Interests

The essence of the law is to protect interests through the assignment of enforceable legal entitlements; thus, any legal or regulatory activity is necessarily premised on the previous selection of set of interests that are worthy of legal protection. Given that the choice of interests is logically prior to the definition of legal entitlements, each efficiency definition builds in the normative premises as to whose interests should be served by efficient legal institutions. Moreover, as previously pointed out, the choice of interests worthy of legal protection compels the allocation of legal entitlements and the distributive consequences of legal rules. In conclusion, a theoretically sound definition of efficiency in legal discourse requires that the set of interests to be protected be specified.

1.2. Three Notions of Economic Efficiency

This section briefly discusses the three notions of efficiency traditionally used in economic discourse: allocative, productive, and dynamic efficiency. It clarifies that the prevailing law and economics approach that is mostly focused on the allocative efficiency of legal rules should incorporate productive and dynamic efficiency of the law-making process into the analysis.

---

1.2.1. Allocative Efficiency

The concept of allocative efficiency plays a central role in mainstream law and economics analysis. Typically, allocative efficiency is defined in the context of general equilibrium analysis, and is identified by the simultaneous concurrence of three conditions: (i) efficiency in production, (ii) efficiency in exchange, and (iii) efficiency in production composition.

Efficiency in production occurs when the production of one good cannot be increased without having to decrease the production of some other economic good. That is, when perfectly efficient production has not been attained, it is possible to increase the production of any economic good just by making inter-sectoral reallocations of input factors, and without limiting the production of any other economic good. To clarify this concept, it is useful to introduce the notion of the production possibility frontier. Let us consider the classic example of an economy with only two goods (e.g., butter and guns) in which the amount of input factors is given. The production possibility frontier describes the maximum levels of the two goods that can be produced given the amount produced of one of the goods. Each point located along the frontier represents a production bundle such that it is not possible to increase the production of guns without reducing the production of butter, and vice-versa. By contrast, holding constant the total amount of input factors in the economy, each point located below the frontier represents production bundles such that it is possible to increase the production of guns without reducing the production of butter, and vice-versa.
The negative slope of the curve reflects resource *scarcity*; along the frontier, an increase in the production of one good can be achieved only by reducing the production of the other good. The concave nature of the curve represents the *increasing opportunity cost* of producing more of one good. As the economy gets closer to the horizontal or vertical intercept, the production of additional units of one good imposes to renounce the production of increasing amounts of the other good.\(^{57}\)

---

\(^{57}\) The efficiency in production can be described using a two-input model in which firms choose among alternative production-bundles, given the input prices. In this context, the efficiency of production can be described in terms of cost-minimization. Graphically, the point of tangency between isoquants (the combination of inputs producing the same level of output) and isocosts (the combination of inputs associated to the same level of costs) represents the cost-minimizing production technique. Algebraically, the efficiency in production is attained at the point where the marginal rate of transformation between input factors (i.e. the slope of isoquants) is equal in all production sectors to the relative prices of inputs (i.e. slope of the isocosts). Until the marginal rate of transformation of different sectors is equated, it is possible to increase the overall production of the economy by reducing the production of some goods which require a higher marginal opportunity cost per additional unit of input (in terms of the foregone units of the other of input), and increasing the production of other goods whose production requires a lower marginal opportunity costs per additional unit of inputs.

In chapter 2, I will discuss the difficulties of adopting this neoclassical representation of the concept of “efficient production” to the lawmaking process.
Efficiency in *exchange* is reached when no further mutually advantageous exchanges are possible. Until the economy reaches the point of efficient consumption, it is possible for the economic actors to increase their level of utility by simply exchanging goods according to their relative preferences. Once efficient allocation has been reached, assuming a fixed available amount of goods and individuals’ preferences, any exchange that makes one of the parties better off necessarily makes some other party worse off. To clarify this point, it is useful to briefly introduce the *utility possibility frontier* concept. Assume an economy with two individuals, say Friday and Crusoe, in which the amount of goods available is given. The utility possibility frontier describes the maximum level of utility that is attainable by the two individuals, given the other individual’s utility level. At each point located along the frontier, it is not possible to increase Friday’s utility without making Crusoe worse off, and vice-versa. By contrast, if the economy is located below the frontier, Friday’s utility can increase without decreasing Crusoe’s utility, and vice-versa.\textsuperscript{58} It must be emphasized that when efficiency in exchange is attained, the individual economic surplus (i.e., the difference between their willingness to pay and the actual price they pay) is maximized; the two individuals reach equilibrium because neither of them can improve their surpluses through further exchanges. Mainstream law and economics focuses on allocative efficiency understood as the maximization of individual surplus.

\textsuperscript{58}Efficiency in consumption can be described using a two goods model, i.e. a model in which each individual chooses consumption bundles of two goods based on preferences (as represented by indifference curves) and purchasing power (as represented by the budget line). In this context, consumption efficiency is attained at the point where the marginal rate of substitution for goods and services is equal for all consumers at current market prices. Until the marginal rate of substitution is not equalled, it is possible to increase the overall utility of the economy by reducing the consumption of some goods that require a higher marginal opportunity cost (i.e. in terms of the foregone units of the other goods) per additional unit of consumption and by increasing the consumption of other goods with a lower marginal opportunity cost.
The third condition for allocative efficiency is optimal coordination between efficiency in consumption and production, given production technology and consumer preferences. Coordination between consumption and production can be described in terms of efficiency in production composition. Let us consider the economy where only guns and butter are produced from an aggregate perspective. Let us also assume that given the available technology and quantity of input factors, the maximum quantity of each good is produced; namely, the economy produces the maximum quantity of butter and guns possible, to use the previous example. Under these assumptions, when efficiency in production composition has not been attained, it is possible to increase the total utility from guns and butter by simply reallocating the production from one good to another. By contrast, when production is efficiently allocated between two goods, it is not possible to reallocate the production from butter to guns or vice versa, without concurrently decreasing the total utility.\(^59\)

\(^{59}\) In more technical terms, the marginal rate of substitution and the marginal rate of transformation equate each other and, simultaneously, are equal to the relative prices of productive inputs and consumption goods.
Having clarified the notion of allocative efficiency, the discussion should turn towards how law and economic scholars use allocative efficiency to analyze legal rules.

1.2.1.1. Allocative Efficiency and Economic Surplus

Generally, when law and economics scholars discuss the allocative efficiency of legal rules, they consider the maximization of individuals’ economic surplus. The efficient legal rule maximizes the sum of the individuals’ surpluses subject to the rule. That is, when a legal rule is efficient—setting aside for a moment the issue of transaction costs—it is not possible to increase the aggregate surplus magnitude by shifting to an alternative legal regime. Given the existing technology, available resources, and individuals’ preferences, no additional gain can be obtained through a change in the legal regime because the sum of individuals’ surpluses has been maximized.

Positive law and economics identifies alternative mechanisms for allocating legal entitlements so as to maximize social surplus. Sometimes the law simply facilitates individual bargaining (i.e., “lubricating approach”); alternatively, when private bargaining is too costly, the law mandates cooperative behaviours among parties (i.e., Hobbesian approach) or allocates legal rights to those individuals who value them most highly (hereinafter “the efficiency right holders”). An exhaustive discussion of the alternative surplus-maximizing legal techniques is beyond the scope of this brief summary. It is sufficient here to emphasize the focus of mainstream law and economics on the efficiency of alternative legal entitlement allocation. The following two subsections emphasize the implicit (often inaccurate) institutional assumptions underlying this analysis style.

1.2.1.2. Allocative Efficiency and Single Institutional Analysis

The privileged domain of inquiry by legal economics scholars has been the problem of economic externalities, namely, interference of an economic actor’s activity with the

---

60 See the foundational contribution of Ronald Coase, “The Problem of Social Cost” (1960) 3 J. Law and Econ. 1 [Coase, “The Problem”].
62 See, Coase, “The Problem” supra note 60.
utility or cost function of other economic actors. The traditional approach proceeds by asking two questions: (i) who is entitled to the right of interfering or, symmetrically, being free from interference (hereinafter “the entitlement” issue), and (ii) how the entitlement is to be protected (hereinafter the “the protection” issue). As to the entitlement issue, allocative efficiency dictates that the right be assigned to the party who values it the most. Let us consider the case of negative externality from pollution. Here, the entitlement issue is whether the polluter or the victim values the right more highly. If the polluter values the right more, then he or she should be assigned the right to interfere in the victim’s utility function. Conversely, when it is the victim who values the right more, he or she should be entitled not to suffer interference in his/her utility or cost function.

The protection issue is traditionally related to the distinction between property and liability rules. Property rules assign to the owner the right of being free from interference. An owner protected by a property rule can obtain an injunction enjoining others from encroaching on his or her rights. In the pollution case, a property rule protecting the right of the polluter gives him or her the right to exercise the polluting activity without interference; here, the externality cost is suffered entirely by the victim. Conversely, a property rule protecting the right of the victim gives him or her the right of being free from interference and enjoins the polluter from generating the externality; here, the polluter fully internalizes the externality costs. Unlike property rules, liability rules assign to the entitled person the right to obtain compensation for damages due to external interference. In the pollution case, a liability rule protecting the polluter gives him of her the right to carry out an activity and obtain damages should the victim interfere. Conversely, a liability rule protecting the victim gives him or her the right of obtaining damages equal to the costs suffered as a consequence of the polluting activity.

---
As a practical matter, the choice between property and liability rules is influenced by many factors. Law and economics practitioners particularly emphasize the importance of transaction costs. In cases where transaction costs are low enough to allow private bargaining, efficient resource use results from negotiations among interested parties. It is maintained that if transaction costs are low, the clear assignment of property rights facilitates private bargaining. In particular, as identified by Coase, under the assumption of a frictionless world, private bargaining reaches the efficient allocation of rights regardless of the initial assignment of property rights. When transaction costs are high enough to prevent parties from bargaining, the efficient use of resources will depend on the assignment of property rights; because private bargaining cannot attain allocative efficiency, the law allocates the right to the efficient holder.

The abovementioned principles constitute the traditional teaching of law and economics. As we can see from this brief summary, the analytical focus is on efficiency improvements at the margin. That is, the allocation of rights is assessed without factoring into the efficiency calculus the impact of alternative institutional processes through which

---

64 E.g.: who can abate the externality at least cost; whether the magnitude of transaction costs allows for a bargaining solution; what are the effects on the ex ante incentives to undertake economic activities.


From a law and development perspective, see Douglass C North “Institutions, Transaction Costs and Economic Growth” (1987) 25.3 Economic Inquiry, 419.


Finally, for a general comprehensive overview, see Douglas W Allen, “What are Transaction Costs” (1991) 14.0 RLE 1, and the bibliography there cited.


67 The example discussed in Coase’s article rests on a set of simplifying assumptions regarding the functioning of the bargaining process: (1) parties are rational, (2) fully informed, (3) negotiate under the conditions of zero transaction costs, (4) in a perfectly competitive environment where prices of goods are given, and where (5) there is no income effect. Hereinafter, I will refer to the set of assumptions characterizing Coase’s article as to the “Coaseian environment”.
rights are allocated. The transaction costs are considered to the extent that they affect the magnitude of individual surplus, but the source of the transaction costs is exogenous to the model. When transaction costs are low, the economy can move to a point along the possibilities frontier via private bargaining; if transaction costs are higher than the cooperative surplus attainable via private bargaining, individuals will not engage in bargaining activity. In the light of this view, the outcome of law-making (i.e., the allocation of rights) is insulated from the law-making process itself (i.e., incentives, preferences, and constraints of those who decide rights allocation). Seemingly, transaction costs are high or low independent of the costs generated by alternative institutional processes for rights allocation. In more recent years, many scholars have identified additional conditions—besides the level of transaction costs—that affect the efficiency of property rules and liability rules. As a result, the prevailing analysis style has progressively become more sophisticated. However, notwithstanding these analytical advances, the methodological focus remains on marginal efficiency.

Insulation of the legal rule from the law-making process is methodologically misguided. Traditional prescriptions for the use of property rules and liability rules are based on an implicit institutional analysis. One could rephrase them as follows:

(i) when transaction costs are sufficiently low, the market is the institution better suited for allocating the right to the efficient holder;
(ii) when the costs of operating markets exceed the cooperative surplus generated by private bargaining, courts are deemed to achieve better allocative efficiency by assigning the right to damages or property rights protection to the efficient holder.

There is a nirvana fallacy here: the reasoning tacitly assumes that when the market does not work efficiently (in the case of high transaction costs), courts can do a better job. The efficiency of the courts is inferred from market imperfections. This is not a correct way of efficiency analysis for the reasons clarified previously. First, the set of viable institutional alternatives is not identified clearly (e.g., why not compare the market and

---

68 See, in particular, Kaplow et al. “Property Rules”, supra note 63.
69 See, Komesar, Imperfect Alternatives, supra note 7 at 21.
the courts with the political or administrative processes?). Second, there is no examination of the preferences and incentives of the actors involved in the adjudication process (and its alternatives). Third, the bounded rationality and limited information of the individuals involved in alternative institutional law-making processes are not considered. In short, as Komesar pointed out, the traditional prescription of law and economics is based on a single institutional analysis.70 Yet, as we have seen, a reliable institutional analysis proceeds by comparing the relative performances of a defined set of viable alternative institutions. Drawing inferences about the efficiency of one institution directly from the failure of its alternatives lead to unreliable conclusions. Only comparative analysis allows for the identification of the efficient (i.e., less imperfect) institution.

In conclusion, insulation of the output from the process is illusory and misguided. It is an illusion because there is no real coherent analytical separation; rather, there is an implicit assumption on what is the efficient process. It is misguided because doing so leads to unreliable conclusions that are not verified by methodologically sound comparative institutional analysis.

1.2.2. Productive Efficiency

Productive efficiency can be intended in two different ways. Under the first approach, it refers to the efficient allocation of input factors. In this sense, it is a synonym of “technical efficiency”: producing the largest possible output from a given set of inputs. If we look at the cost function, productive efficiency is achieved when the output is produced at the minimum average total cost. In the language of modern microeconomic theory, productive efficiency corresponds to the point of tangency between isocosts and isoquants, in which the producer chooses the input bundle that minimizes the production costs while maximizing the output level.71

70 Ibid.
71 The next chapter provides a more detailed discussion of the neoclassical theory of production, as well as the strengths and limitations of its use, to explain efficiency in the production of legal rules.
From another perspective, Leibenstein has introduced the concept of “X-Efficiency” to elucidate a second dimension of productive efficiency. Leibenstein emphasizes the limitations inherent in focusing economic analysis exclusively on welfare improvements achievable through the efficient allocation of productive inputs. In his seminar paper, Leibenstein contends that welfare improvements could be achieved by intervening in the organizational factors that determine the efficient utilization of productive inputs—he called this X-efficiency, which largely depends upon the incentives provided to managers and workers by the competitive pressures of an environment. In particular, economic theory has focused on (i) excessive employment of production factors (i.e. non-optimal labour/capital ratio), (ii) higher payment for production factors, and (iii) lower productivity levels.

Leibenstein’s reasoning is well suited for elucidating the central concern of this study. To obtain a better understanding of how the law can improve social welfare, law and economics should investigate the internal inefficiencies of law-making and identify the mechanisms through which said inefficiencies affect law-making costs.

---

72 Harvey Leibenstein “Allocative Efficiency vs. X-Efficiency” (1966) 56 Am. Econ. Rev. 2 394.
73 Ibid. at 394–395: “The empirical evidence, while far from exhaustive, certainly suggests that the welfare gains that can be achieved only by increasing allocative efficiency are usually exceedingly small at least in capitalist economies […] the welfare loss of allocative inefficiency is of trivial significance”.
74 To be clear, “X-efficiency” denotes a concept different from “efficiency in production”, previously identified. X-efficiency is related to the efficient organization of the production process, with particular regard for the incentives and motivational factors of the actors involved in the production process; efficiency in production concerns the efficient allocation of the inputs to be employed in the production process (e.g. labour and capital).
75 Leibenstein identified three determinants of X-efficiency; see ibid. at 412–413: “(a) contracts for labour are incomplete; (b) the production function is not completely specified or known, and (c) not all inputs are marketed or, if marketed, are not available on equal terms to all buyers”.
76 It is important to clarify that although Leibenstein’s reasoning on the limits of allocative efficiency can be usefully extended to law and economics theory, the X-efficiency criterion does not provide a sufficient account of the efficiency of the law-making process. The concept of productive efficiency needs to be adapted to the legal context. X-efficiency does adequately account neither the demand side of the law-making process nor the output quality. Moreover, it focuses exclusively on the minimization of the ratio between inputs and output. As I will explain in the next chapter, process efficiency cannot be reduced to a cost-minimization problem and should be integrated into a richer account of other economic values that come into play in the analysis of the law-making process.
1.2.3. Dynamic Efficiency

Both allocative and productive efficiency are focused on the static aspect of efficient resource allocation and usage. They assume constant given technology and resources, and reduce the economic problem to the avoidance (or minimization) of resource wastage. By contrast, the idea of dynamic efficiency rests on the assumption that the given technology and resources are not constant, i.e. the ability of economic actors to innovate the economic environment enables them to overcome technological constraints in economic production. In essence, dynamic efficiency looks at efficiency in an inter-temporal dimension.77

The Pareto possibility frontier idea helps understand the notion of dynamic efficiency. From the static perspective, the economy is efficient when located along a given production possibilities frontier; by contrast, from a dynamic standpoint, the most important factor is the ability of the economic actors to shift the production possibilities curve progressively towards the right by moving the technical constraints forward. Put differently, allocative efficiency is an attribute of a state of economic affairs; with this view, the economic problem is reduced to a technical issue of constrained output maximization, given the available level of inputs and the available existing technology. By contrast, dynamic efficiency comprises the ability of the economic actors to recognize new opportunities that appear in the economic environment and adapt their activity for benefitting from such opportunities. Rather than a state of economic affairs, dynamic efficiency represents an attribute of the economic process, which displays long-term effects in terms of economic growth, knowledge diffusion, and innovation. It is important to consider that the static allocative efficiency does not necessarily entail long-term dynamic efficiency, and, conversely, that the latter may not always be compatible with static efficiency in the short term.

77To be clear, what we call “dynamic efficiency” has been the central concern for the economists since the early stage of the history of the economic thought. For instance, from the perspective developed by Adam Smith, the most valuable feature of the free-market based economy is that it grows rapidly, not that it is allocative-efficient. The term “efficiency” never appears in “The Wealth of Nations”. 
The output-oriented approach often remains incapable of identifying the dynamic implications of legal rules. In particular, the identification and explanation of a possible trade-off between allocative and dynamic efficiency represents one of the major challenges in the economic analysis of legal rules. Although an exhaustive account of the dynamic efficiency of law-making is beyond the scope of this study, it is assumed here that the ability of the law-making process to reduce the costs of transition to the new efficient legal standard facilitates the dynamic efficiency of law-making.

1.3. Limitations of Output-oriented Approach

This section closely examines three normative standards employed in legal-economic analysis: Pareto optimality (hereinafter “PO”), Kaldor–Hicks (hereinafter “KH”) efficiency, wealth maximization (hereinafter “WM”), and social welfare (hereinafter, “SW”). The discussion summarizes first the logical structure and normative content of each standard, and then emphasizes the major criticisms that are attributable to “output bias” including (i) logical circularity, (ii) failure to identify increases in social welfare, (iii) failure to account for distributional concerns, and (d) failure to account for the dynamics of legal change.

1.3.1. Pareto Efficiency

The Italian economist Vilfredo Pareto developed the concept of Pareto efficiency for constructing a value-free social decision-making theory and for avoiding the difficulties
of the so-called “interpersonal comparison of utility” (hereinafter “ICU”).\textsuperscript{82} Pareto’s idea of replacing the ICU with the \textit{unanimity principle} is simple and powerful. When people consent to decisions, there is no need for comparing different individuals’ utility functions. It is assumed that the consensus follows personal preferences. The following rule for comparing alternative social states operationalises the unanimity principle: \textit{if each member of society regards x to be at least as good as y, and at least one individual strictly prefers x to y, then the society should prefer x to y.} It is important to distinguish between PO and Pareto \textit{superiority}. Pareto superiority compares \textit{two} states of affairs: e.g. $x$ is Pareto-superior to $y$ if the move to $x$ generates no losers and at least one winner. PO is the maximal element of ordering based on Pareto superiority: e.g. one social state of affair $x$ is Pareto-optimal when no Pareto-superior moves are possible considering \textit{all} possible alternative allocations. In essence, while Pareto superiority allows for further Pareto-superior moves, PO is the state of affairs where no further Pareto-superior moves are possible.

The logic underlying Pareto efficiency is based on three elements. First, the Pareto criterion abandons cardinal individual utility in favour of \textit{ordinal} individuals’ preferences. It compares individuals’ ordering of alternative social states, irrespective of the assignment of cardinal values to individual preferences. Second, rather than seeking an objective measure of individual preferences, the Pareto criterion shifts to an \textit{intrapersonal utility comparison}, namely, the variation of one individual’s welfare is assessed against said individual’s utility function, rather than other individuals’ welfare.\textsuperscript{83} Third, it is as if individuals were assumed to have “lexicographic” preferences with respect to their own personal utility. This means that individuals are assumed not to

\textsuperscript{82} Individual preferences are difficult to discern, and any distributional judgment is value-laden; on the other hand, there is no agreement on the priority order among different ethical values. For these reasons, it is hardly possible to ground the choice among alternative policies on a firm scientific base when distributional consequences are at issue.

\textsuperscript{83} On this point see Bruce Chapman, “Economic Analysis”, \textit{supra} note 34, which analyzes the contract as the typical form of Pareto-superior moves (“[…] the source of the social gain in a contract is not to be found in any such interpersonal comparison, at least if we restrict ourselves to the economist’s idea of Pareto superiority. Rather, the social gain in a contract is to be found in intrapersonal comparisons only, namely, in the idea that no one person (given contractual compensation) is worse off in terms of his or her own welfare or satisfaction after the exchange than that person was before it, and at least one person (willing to pay the compensation) is better off”).
compare their own personal utility with social utility or with other individuals' utilities; rather, if offered several social states, they invariably choose the social state that offers the highest personal utility, no matter how much social (or others') utility is derived from their choice.

The practical implication of the Pareto test based on the three preceding principles is that any loss suffered by any individual is not justified by any benefit gained by any other individual.

The reliance on intrapersonal comparison of utility and the assumption of “lexicographic” individual preferences are the sources of the Pareto principle’s major shortcomings. In particular, I identify and describe the following limitations in the discussion below: (1) logical circularity and incompleteness, (2) inability to compare alternative optimal legal changes and to identify the efficient right-holder, (3) neglect of distributional issues and related status quo bias, and (4) inability to explain legal change.

1.3.1.1. Logical Criticisms

Logical Circularity. Pareto efficiency is relative to a given initial distribution of resources and structure of rights. Since individuals’ wealth levels significantly influence their marginal utility curves, a resource allocation scheme can be defined as Pareto-efficient only when the initial endowment of individuals is known. In the absence of this information, not one but many Pareto-efficient allocative distributions are possible. In this respect, institutional scholarship has long emphasized that any Pareto optimal allocation is determined by the structure of legal entitlements. This raises one major criticism of Pareto efficiency as a normative criterion for assigning legal entitlements. Namely, Pareto Efficiency does not provide guidance in the choice among alternative

---

84 See Samuels, “The Legal” supra note 30 at 1563: (“Efficiency in the sense of Pareto optimality means the exhaustion of gains from trade of goods. But efficient-optimal allocations are specific to the definition and assignment of rights forming the basis on which trades take place. There is no unique optimal result, only rights-structure-specific optimal results.”).
optimal allocations. The structure of legal entitlements determines Pareto-optimality, and the reverse is not true.

Logical Incompleteness. When it comes to the legal realm, the Pareto test fails because its claim of neutrality is based on the idea of the possibility of no losers. Instead, the inevitability of losers and the related interpersonal comparison is to be recognized as an essential component of legal change. In fact, generally, a change in the assignment of legal rights results in someone gaining and someone else losing. In short, the Pareto principle is rarely applicable to the reality of legal change simply because the unanimity

---

85 See Samuels, “The Legal” supra note 30 at 1563: (“[…] efficiency in the Pareto sense cannot dispositively be applied to the definition and assignment of rights themselves because efficiency requires an antecedent determination of the rights […]. To say, then, that the law is wealth-maximizing is not saying anything more than that the law, in determining whose interests will count as rights, determines which non-unique optimal solution will be achieved. In this respect, then, the law is willy-nilly wealth-maximizing, whichever or whosoever interests it protects as rights. This non-unique choice is accomplished in the arena of the law and the economy, which is not self-subsistent”).

86 From a non-institutional perspective, J Mishan demonstrated that since the law influences the distribution of wealth, the optimal allocation of resources ultimately depends on the structure of the legal order; see, EJ Mishan, “Optimality” supra note 27. In particular, Mishan (ibid. at 255) makes the following three points:

“(1) […] the characteristics of an optimal solution are not uniquely specified but depend, in general, on the existing law,
(2) […] the costs incurred in realizing an optimal outcome, and the question therefore of its “feasibility”, also depend upon the existing law, and finally;
(3) […] an optimal solution emerging from conflicts of interest is optimal only with respect to an implicit constraint requiring the area in question to be used in common by the groups or persons having conflicting interests. Once separate areas, or separate facilities, are introduced solutions appear that are Pareto superior to the familiar constrained optimal outcomes”.

87 In general, that legal change usually involves the presence of losers is generally recognized in law and economics literature; see, for example, Francesco Denozza, “Fairness and Welfare. Are they really Competing Values?” in Legal Orderings And Economic Institutions (New York: Routledge, 2007) at 154 (“A given rule usually increases the well-being of some, at the same time decreasing the well-being of others”), and, more recently Trebilcock, Dealing with Losers, supra note 37.

88 This is true independently of the characteristics of the law-making process. For example, the judicial process by definition arises out of the conflicts between contrasting interests; it generally concludes with someone who wins and someone who loses. The winner-looser dynamic is pervasive also in legislative processes; an entire body of scholarship has been developed based on the idea that organized interest groups attempt to capture the legislative process to gain favourable treatment at the expense of larger unorganized public interests. Finally, in private markets, economic exchange is usually based on mutual consensus between parties; for this reason, economists typically tend to identify market transactions as the main source of Pareto-efficient changes. However, it is well known that the problems of bounded rationality, limited information, and asymmetric bargaining power in many cases undermine the Pareto-superiority of private exchange.
criterion is incompatible with the presence of losers, which is an essential part of most legal changes. 89

1.3.1.2. Social Welfare

As noted previously, positive law and economics is based on the idea that social surplus is maximized when resources are allocated to the person valuing them the most. However, the shift from the positive to the normative dimension calls for a richer account of the objectives of the legal order. It is hardly sustainable in theory as well as in practice to ensure that legal entitlements be based exclusively on surplus-maximization considerations. The pursuit of “social welfare” and “individuals’ well being” should be recognized as a more appropriate normative task than the mere maximization of individual economic surplus. This raises a serious challenge to the practicability of Pareto efficiency within normative legal–economic discourse. In fact, only interpersonal utility comparison (and abandoning of Pareto non-cardinality) allows for identification of the person who values the contested resource the most in welfare terms. Without measuring the relative welfare gains and losses of different individuals, we cannot say who values the right most highly and whether social welfare has been maximized. On the contrary, intra-personal comparison of utility identifies Pareto-superior moves, but does not allow for the identification of who values the right most in terms of welfare. As a consequence,

89 Normativity cannot be neutral. See Calabresi, “The Pointlessness”, supra note 27 at 1217 (“[…] the moment we deal with a real situation, we know something more about who wins and who loses. Once we are no longer ignorant, any number of differences may cause us to believe that losses or gains to some matter more than losses or gains to others. In other words, we are immediately in a real world where we must deal with competing theories of just distribution and in which a theory maintaining that losses and gains to everyone, rich and poor, good and bad, regardless of starting points, are of the same social significance, has little to commend it. […] Indeed, one conclusion of this Article will be that the failure of the Pareto criterion to be of any real guidance makes inevitable a thoroughgoing and open discussion of distribution and of interpersonal comparisons. […] There may also be situations in which a just distributional theory leads to the conclusion that we are indifferent between the actual winners and the actual losers, and hence that any Kaldor–Hicks superior change is desirable.36 But without a real-world look at a particular situation, followed by the application of an explicit distributional theory to that situation, the attempt to make Kaldor–Hicks efficiency into a general normative guide must fail. It can be no more than a disingenuous attempt to convert a distributional theory of uncertain validity into one that has genuine distributional neutrality. Let me be clear. That normative world is the one in which we live, and we can say a good deal about what is better and what is worse in it. What we cannot do, though, is to hide what is involved in such a world behind an interpersonally neutral criterion and act as though we are not making moral judgments, the effects of which we then impose on others.”) To put it in a slightly different perspective: in order to come to grips with reality and provide practical guidance in evaluating legal changes, the normative theory has to recognize the necessity of the discussion on distributional consequences.
an *intrapersonal* comparison might justify a Pareto-superior move (which makes everyone better off) that allocates the contested resource to a person that values it *less* according to an *interpersonal* evaluation.\(^9^0\)

In conclusion, while we can say that a Pareto improvement *increases* social welfare (at least, one person is left with increased surplus), we cannot know with certainty whether it *maximizes* social welfare (i.e. whether the resource has been assigned to the person that values it the most).

**1.3.1.3. Distribution**

*Pareto Optimality Does Not Capture The Presence Of Losers.* Pareto optimality does not distinguish between Pareto-superior and Pareto-inferior moves. Since this is a crucial shortcoming of the Pareto notion of efficiency, it is worth expanding on this point and providing an analytical demonstration. Let us consider the classic farmer/rancher example.\(^9^1\) A rancher and a farmer run their own activities on adjacent plots of land. The rancher runs cows, while the farmer grows corn. From time to time, the cows wander onto farmer’s land causing damage to the farmers’ crops; the interference among these conflicting activities (running cows and growing corn) represents a classical case of negative *externality*. Suppose that the rancher wants to raise one more cow, and that the marginal profit from running the second cow exceeds the marginal cost in terms of additional damage to the corn. Say for example that the profit to the rancher from an

\(^{90}\) On this point, see Bruce Chapman, “Economics Analysis” *supra* note 34, at 4–5: (“[…] intrapersonal comparisons are quite consistent with the possibility that, if an interpersonal comparison were made, one of the goods or services exchanged was actually going from someone who valued it more highly in welfare terms to someone who valued it less highly. To see the point vividly, imagine someone who is both cold and hungry, and who, in exchange for some much needed food, gives up her warm jacket to another person who is neither cold nor hungry. One can easily appreciate that this might be a Pareto superior exchange in which each party is better off, but that, nevertheless, the jacket was more highly valued in an interpersonal sense by the person who was cold. This is a significant point for understanding the value of Kaldor-Hicks efficiency and we will have reason to reconsider it below”).

\(^{91}\) The farmer–rancher example is discussed in Coase, “The Problem” *supra* note 60. This example is traditionally used in the literature to illustrate how private bargaining in a Coaseian environment reaches allocative efficiency regardless of the structure of legal rights. A textbook discussion of the farmer–rancher example is provided, for example, by Cooter et al. *Law and Economics*, *supra* note 61, at 85–91; Jules Coleman, “Efficiency, Exchange and Auction: Philosphic Aspects of the Economic Approach to Law” (1980) 68 Cal. L. Rev. 226 at 226–231 [“Efficiency"] provides a critical discussion of the farmer–rancher example that relates the allocative efficiency of Coasean bargaining to the Pareto-optimality of the final allocative results.
additional cow is $50, while the damage imposed on the farmer is $25. In this case, the rancher is better off by raising a second cow. Suppose also that the cost of building a fence preventing the cow from intruding farmer’s land is less than the damage to crops from the additional cow. Say, for example that the cost of fencing off the cow is $20. Since it is assumed that the marginal profit from an additional cow ($50) exceeds its marginal cost ($25), it is socially efficient (in the sense of allocative efficiency) to raise a second cow so that social surplus increases by $25. Now, assuming zero (or low) transaction costs, alternative legal regimes influence the distribution of this $25 increase in social surplus but do not prevent private parties from reaching an equilibrium point in which the second cow is raised.

Consider the initial scenario, S1, where only one cow is raised. Alternative legal regimes lead to the same efficient terminal allocation (i.e., a second cow is raised); however, depending on the structure of legal rights, the final point of optimality is reached through a series of either Pareto-superior or Pareto-inferior moves. Let us consider four alternative legal regimes.92

S2: a property rule assigns the right to the farmer to prevent the rancher from raising a second cow.

In this case, the rancher offers the farmer a sum $x$ to acquire the right of raising a second cow, where $25 < x < 50$. Since $x$ exceeds the damage caused to the crops from the additional cow, the farmer accepts rancher’s offer and allows him or her to raise the second cow. In conclusion, the rancher’s surplus is equal to $(50 - x)$ and farmers’ surplus is equal to $(x - 25)$; thus, the social surplus associated with S2 is equal to $(50 - x) + (x - 25) = 25$. The passage from S1 to S2 is a Pareto-superior move: both the rancher and the farmer will prefer S2 to S1.

S3: a liability rule assigns the farmer the right to receive damages equal to the value of the crop damaged by a cow’s wandering on his or her land.

92 I draw from Coleman, “Efficiency” supra note 91 at 226-231.
In this case, the rancher “buys” the right of raising the second cow and letting it wander on the farmer’s land by paying $25 in damages to the farmer. Since fencing off the cow is less expensive than paying damages, the rancher would prefer to pay a sum $y$ to the farmer to build a fence, where $20 \leq x < 25$. However, the farmer will prefer to receive $25$ damages than building the fence for a sum $x < 25$. In conclusion, in S3 the farmer allows the rancher to raise the second cow and is concurrently compensated for damages to his or her crops. The rancher’s surplus is equal to $(50 - 25) = 25$ and the farmer’s surplus is equal to 0; thus, the social surplus associated with S4 is equal to $(50 - 25) + 0 = 25$. The passage from S1 to S3 is a Pareto-superior move: on the one hand, the farmer is indifferent between S1 (no damages to his or her crops) and S3 (damages compensated); on the other hand, the rancher prefer S3 (generating marginal profit of $25 dollars) to S1 (no profit from additional second cow).

**S4: a property rule assigns the rancher the right to raise the second cow and let it wander on the farmer’s land.**

In this case, the rancher raises a second cow without paying for the damages caused to the farmer’s crops. The rancher’s surplus is equal to $50$, and the farmer’s surplus is equal to $-25$; thus, the social surplus associated with S4 is equal to $[50 + (-25)] = 25$. Unlike S2 and S3, the passage from S1 to S4 is a Pareto-inferior move. Although the rancher prefers S4 (generating marginal profit of $25$ dollars) to S1 (no profit from additional second cow), the farmer is worse off by S4 (he or she suffers damages to his or her crops).

**S5: a liability rule protects the right assigned to the rancher to raise the second cow and let it wander on the farmer’s land; in the case of others’ interference with the rancher’s right, he or she has the right to damages.**

Under this rule, if the farmer builds the fence (at the cost of $20$) to prevent the rancher’s cow from wandering on his or her land, he or she must pay $50$. Thus, the farmer prefers to bear $25$ in damages and not interfere with the rancher’s right, rather than fencing off the cow and paying $75$. In conclusion, the rancher’s surplus is equal to $50$, and the farmer’s surplus is equal to $-25$. The social surplus associated with S4 is equal to $50 -
$25 = $25. The passage from S1 to S4 is a Pareto-inferior move: while the rancher prefers S5 (generating marginal profit of $50 dollars) to S1 (no profit from additional second cow), the farmer is worse off with S4 (he or she suffer damages to his or her crops).

The table in Figure 4 summarizes the preceding discussion.

**Figure 4: The Rancher–Farmer Example.**

<table>
<thead>
<tr>
<th></th>
<th>Rancher Surplus</th>
<th>Farmer Surplus</th>
<th>Social Surplus</th>
<th>Rancher</th>
<th>Farmer</th>
<th>Pareto Superior/Inferior</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>50 - x</td>
<td>x - 25</td>
<td>25</td>
<td>Winner</td>
<td>Winner</td>
<td>Superior</td>
</tr>
<tr>
<td>S3</td>
<td>50 - 25</td>
<td>0</td>
<td>25</td>
<td>Winner</td>
<td>Indifferent</td>
<td>Superior</td>
</tr>
<tr>
<td>S4</td>
<td>50</td>
<td>-25</td>
<td>25</td>
<td>Winner</td>
<td>Loser</td>
<td>Inferior</td>
</tr>
<tr>
<td>S5</td>
<td>50</td>
<td>-25</td>
<td>25</td>
<td>Winner</td>
<td>Loser</td>
<td>Inferior</td>
</tr>
</tbody>
</table>

When the marginal benefit of an additional cow exceeds its marginal costs, it is efficient to expand rancher activity; namely, it is efficient to move from S1 to a Pareto-superior scenario in which a second cow is raised and a marginal profit of $50 is generated. From this standpoint, S2, S3, S4, and S5 are all characterized by (i) surplus-maximizing resource allocation (the social surplus is equal to $25), and (ii) Pareto-optimal allocation (it is not possible to rearrange the allocation without causing losses to the rancher). However, despite their equivalence from the PO standpoint, there are substantial differences between the alternative outputs. While in S2 and S3 the rancher realizes his profit after purchasing the right to raise an additional head of cattle (thereby compensating the farmer), in S4 and S5 the rancher realizes his profit by creating uncompensated damages to the farmer. In Paretian terms, in S2 and S3, the society reaches PO through Pareto-superior moves (at least one winner and no losers); on the

---

93 *Ibid.* at 230
contrary, in S4 and S5, the society achieves PO by undertaking Pareto-inferior moves (at least one loser).\textsuperscript{94} Schematically:

\begin{align*}
(S1 \rightarrow S2) & \text{ and } (S1 \rightarrow S3): \text{ Pareto optimality through Pareto-superior moves} \\
(S1 \rightarrow S4) & \text{ and } (S1 \rightarrow S5) \text{ Pareto optimality through Pareto-inferior moves}
\end{align*}

In conclusion, Pareto efficiency does not consider whether PO is reached by the legal order through a series of Pareto-inferior or Pareto-superior moves, thereby remaining blind to the presence of losers on the way to PO. The rancher–farmer example demonstrates that the “distributive” dimension of efficiency\textsuperscript{95} is strictly related to the process by which allocative efficiency is achieved. However, the narrow perspective of the output-oriented approach, which insulates analytically the output from the process, does not enable legal economic scholars to articulate a coherent explanation of the relationship between allocative efficiency and the law-making process.

\textit{Distributional Concerns and Status Quo Bias.} The foregoing discussion suggests that Pareto efficiency hinges on theoretical foundations that make it prone to privilege the status quo in resource allocation. In particular, two elements support this conclusion. First, the Pareto test is based on the unanimity requirement. As a practical matter, the objection of a single individual is enough to block resource reallocation even if everyone else in the society could be better off by doing so. Second, PO tacitly assumes as given the existing endowments of individuals. The implicit reception of the existing distribution

\textsuperscript{94}Ibid. at 227: “A Pareto optimal distribution can be reached either by Pareto superior steps or ‘moves’, by Pareto inferior moves, or by a combination of the two. Saying that a distribution is Pareto efficient, therefore, may call attention to the efficiency of the existing distribution regardless of the efficiency of the steps along the way to it. In contrast, however, labelling a distribution Pareto efficient might mean not only that the existing distribution is Pareto optimal, but also that it is the result only of Pareto superior (and efficient in that sense) moves”.

\textsuperscript{95}As explained previously, the prevailing neoclassical approach is based on the idea of separation between efficiency and distribution; from this perspective, it makes no sense the talk about a “distributive dimension of efficiency”. However, the aim of this study is to incorporate distributive consequences, which are related to the institutional choices, into the efficiency analysis of law.
of wealth constitutes in itself an implicit value judgment, which exposes the Pareto criterion to the criticism of favouring the status quo.\footnote{See Julian Le Grand, \textit{Equity and Choice. An Essay in Economics and Applied Philosophy}, (London: Harper Collins Academic, 1991) [Le Grand, \textit{“Equity and Choice”}] at 32 (\text{“}[…] although commonly described in the relevant literature as a definition of efficiency, Pareto-optimality is more correctly interpreted as a form of value or social welfare function—one that in certain forms, actually incorporates a notion of equity.\text{”}\text{})}

1.3.1.4. Legal Change

Pareto optimality does not tell us anything about the origin of transaction costs, process of legal change, and how it affects the efficiency features of the outcome. Yet, the central concern of law and economics analysis is not the description of a frictionless world, rather the understanding of \textit{“what to do given the existence of transaction costs”},\footnote{\textit{Ibid.} at 1215 (emphasis is mine)} the barriers (of any sort) that impede legal innovation towards efficiency, and methods for minimizing said barriers. In, short, PO does not provide guidance in choosing from among non-optimal Pareto alternatives through available existing legal processes.

1.3.2. Kaldor–Hicks Test

In the attempt to overcome the limitations of PO, Nicholas Kaldor\footnote{Nicholas Kaldor, \textit{“Welfare Propositions of Economic and Interpersonal Comparisons of Utility”} (1939) 49 Economic Journal 549.} and John R. Hicks\footnote{John Hicks, \textit{“Foundations of Welfare Economics”} (1939) 49 Economic Journal, 696; John R Hicks, \textit{“The Rehabilitation of Consumer’s Surplus”} (1941) 8 The Review of Economic Studies, 108.} introduced the \textit{“Kaldor–Hicks test of efficiency”}. Within the law and economic context, KH represents the prevailing idea of economic efficiency. KH abandons the idea of a unique social optimum, and introduces the principle of \textit{hypothetical compensation} that allows for an assessment of the efficiency of Pareto non-optimal points.

According to the KH test, a social state $x$ is superior to another social state $y$ if, and only if, those who prefer $x$ (winners) can compensate for those who prefer $y$ (losers) so that the latter remain indifferent between $x$ and $y$, and the former are still better off in $x$ than in $y$. Put differently, the change from $y$ to $x$ can be qualified as KH-efficient if the benefit of moving from $y$ to $x$ is such that the losers can be fully compensated, and after the compensation the winner still have a net benefit from the decision to move from $y$ to $x$. It is a crucial feature of KH that the compensation is \textit{not necessarily paid} to the losers. KH
only requires a theoretical possibility of compensation, and not real compensation. For this reason KH efficiency is also called “potential Pareto efficiency”. The assumption is that the possibility of hypothetical compensation is adequate for members of the society to arrive at unanimous consensus.

At first glance, the advantages of this efficient criterion look significant. The compensation test seems to reconcile the Pareto concept of efficiency with the unavoidable presence of losers in the reality of policy-making. It accommodates the desire of freeing economic analysis from value judgment by concurrently separating distributional concerns from the issue of efficiency. However, KH has been the target of many objections. In the following subsections, I briefly summarize the most important criticisms.

1.3.2.1. Logical Criticisms

Logical Inconsistency. David Ellerman has recently contended that a numeraire fallacy undermines the logical consistency of KH efficiency. In many respects, Ellerman’s reasoning is consistent with our previous considerations of the logical features of the efficiency concept.

Similar to PO, KH is based on the separation of efficiency from equity. However, efficiency is assessed with the same numeraire whereby distribution is measured. If the transfer of resources carried out to compensate the losers is measured with the same numeraire used for measuring the size of the pie, there will never appear to be any increase or decrease in the size of the pie, and any KH move will look like a mere redistribution of resources. The idea that “money compensations do not change total social

---

100 For a clear explanation of the principle of compensation as a theoretical device for separating efficiency and equity see Robin W Broadway and Neil Bruce, Welfare Economics (Oxford: Blackwell, 1984) at 97.

101 It is interesting to observe that many shortcomings of KH come from its Paretian lineage. In fact, conceptually, KH is rooted in the same conceptual apparatus as Paretian efficiency; it represents only an attempt to adapt its practical usage to the reality of legal change and policy transitions.

wealth as measured in money is only the same sort of tautologous restatement of the consequences of the choice of numeraire”.\textsuperscript{103}

For the purpose of our analysis, Ellerman’s critique is important in two respects. For one thing, it dispels the methodological difficulty of separating equity and efficiency; Ellerman reaches the same conclusion as do institutionalists by placing emphasis on the need for antecedent normative specifications in any efficiency analysis. Similarly, Ellerman points to the tautological character of efficiency analysis when the normative premises underlying the definition of output are not stated expressly.

1.3.2.2. Social Welfare
We have seen that PO is not apt for ensuring SW maximization. The shift proposed by Kaldor and Hicks from actual to “hypothetical” compensation exacerbates this criticism related to Paretian logic.

Actual compensation allows losers to evaluate their losses according to their welfare function. Thus, it ensures that social welfare is not diminished—losers would not agree to compensation that does not effectively offset their suffered loss.\textsuperscript{104} On the contrary, hypothetical compensation entails that losses are not evaluated by losers—on the basis of their personal welfare function—but are assessed by some collective decision-making procedures through hypothetical evaluation based on market prices. This does not ensure that the losses suffered by losers in resource reallocation are compensated adequately by the utility that losers receive from monetary compensation.\textsuperscript{105}

\textsuperscript{103}Ibid. at 99.
\textsuperscript{104}Though it does not ensure that social surplus maximization is achieved, as previously argued.
\textsuperscript{105}See, for example, Chapman who raises the doubt that an increase in wealth resulting from reallocation in favour of the winner does not correspond to an increase in welfare large enough to outweigh the loss of welfare borne by the uncompensated loser. Should this happen, KH might qualify an economic change that creates a net loss in terms of welfare as desirable. See, Chapman, “Economic Analysis” supra note 34 at 9 (“[…] in a Kaldor-Hicks superior re-allocation, […] we have lost the assurance that any possible loss in the transfer of the good to someone who values it less highly has been properly offset for that person by the welfare gain she receives in having money to spend on other goods she prefers. Thus, it is possible that all we have accomplished is the transfer of a good from someone who values it more highly to someone who values it less highly and, therefore, a loss of total welfare overall”). This point is made clear also in Le Grand, Equity and Choice, supra note 96 at 33–34.
The reasons why monetary hypothetical compensation does not ensure full compensation can be different. It might happen that despite gains exceeding losses in monetary terms, the monetary compensation received by the losers does not offset their loss of welfare. This might depend on the fact that the marginal utility of money is different across persons. First, as already emphasized, the income effect should be considered: “the poor person’s utility loss might exceed the rich person’s utility gain”. Second, external factors might determine a differential of utility between gainers and losers. For example, the availability of goods that can effectively replace the lost resource might not be sufficient to compensate the losers in terms of utility. Third, individuals have different “productivity” in converting the monetary compensation received in actual utility. This might depends on many subjective factors. Of course, these elements can operate in the opposite direction as well. Namely, it might be that despite in monetary terms the gainer’s gain is larger than the loser’s loss, the monetary value received by the gainer does not enable them to generate a utility that more than offsets the loser’s losses.

To summarize, KH superiority is not apt for ensuring any improvement in social welfare. The reason is simple. The shift from actual to hypothetical monetary compensation eliminates the guarantee that the loser’s welfare loss is effectively compensated for, and that the gainer’s welfare gain is effectively larger than the welfare loss to the loser.

---


107 This case is considered by Matthew D Adler, “Beyond Efficiency and Procedure: A Welfarist Theory of Regulation” (2000) 28 Fla. St. UL Rev. 241 at 246 (“Beyond Efficiency”) (“money has (or can have) differential welfare productivity across persons.”)

108 Despite the authors’ use of different terminology and different arguments, on this point there is substantial agreement in literature. See, e.g., Chapman, “Economics Analysis” supra note 34 at 8–10; Mathis, “Efficiency Instead Of Justice?” supra note 106 at 47; Adler “Beyond Efficiency” supra note 107 at 246.
1.3.2.3. Distribution

In accordance with its Paretian lineage, KH ignores any distributional concern.\textsuperscript{109} However, its reliance on market prices exacerbates this criticism of Paretian logic. Two aspects come to the fore here. First, hypothetical compensation does not factor in the different impact on individuals’ utility of money depending on income level.\textsuperscript{110} One could say that KH implicitly assumes a constant and equal marginal utility for all individuals. Rather than achieving neutrality, this conceals a peculiar social welfare function that systematically values rich people more than poor people.\textsuperscript{111} Second, gains

\textsuperscript{109}All that matters is the aggregate increase in real income for the whole of society: individuals’ well-being as such does not count for the purpose of KH. In this respect, KH is vulnerable to the same criticisms as the Pareto efficiency criterion.

\textsuperscript{110}The principle of decreasing the marginal utility of money has long been recognized in the history of economic thought. See, e.g., Jeremy Bentham, “Pannomial Fragments” in Stephen G. Engelmann ed., Jeremy Bentham Selected Writings, (New Haven: Yale University Press, 2011) at 275 (“the quantity of happiness will not go on increasing in anything near the same proportion as the quantity of wealth—ten thousand time the quantity of wealth will not bring with it ten thousand time the quantity of happiness. It will even be a matter of doubt whether ten thousand times the wealth will in general bring with it twice the happiness […] The effect of wealth in the production of happiness goes on diminishing, as the quantity by which the wealth of one man exceeds that of another goes on increasing: in other words, the happiness produced by a particle of wealth (each particle being of the same magnitude) will be less and less at every particle.”). See also, AC Pigou, The Economics of Welfare (London: Macmillan, 1924), and A Study in Public Finance (London: Macmillan, 1928); Richard A Musgrave, The Theory of Public Finance: a Study in Public Economy (New York: McGraw-Hill, 1959) at 102-105.

\textsuperscript{111}From an opposite perspective, neoclassical economists have argued in favour of the assumption of constant marginal utility of money, see Paul Samuelson, “Complementarity: An Essay on the 40th Anniversary of the Hicks–Allen Revolution in Demand Theory” (1979) 12 J. Econ. Lit. 1255 at 1264 (“The money-metric marginal utility of income is constant at unity. For how could it be otherwise? If you are measuring utility by money, it must remain constant with respect to money: a yardstick cannot change in terms of itself.”).

\textsuperscript{111}This point is well explained by Le Grand, Equity and Choice, supra note 96 at 33-34 (“Suppose there is a change of some kind that benefits a rich person by an amount estimated $10, but also imposes a cost on a poor person of $9. Clearly, the rich person could fully compensate the poor persona and remain a gainer. So, according to the compensation principle, whether or not the compensation actually takes place, there has been an increase in social welfare. But now consider the utility change involved. On the not unreasonable assumption of diminishing marginal utility of income, if there is no compensation it is likely that the reduction in utility for the poor person due to the change is greater than the increase in utility for the rich person. Therefore, the only way that there could be a net increase in social welfare (in the absence of any actual compensation) is if the (smaller) change in utility for the rich person is “weighted” in such a way that it increases welfare more than the (greater) utility change for the poor person reduces welfare. And this could only occur if the social welfare function weights rich people more (value them more) than poor people.”)

With specific respect to the use of Kaldor–Hicks efficiency in law and economics, Markovits raises a similar point by emphasizing the distortive effect of standard monetized Kaldor–Hicks operationalisation, which does not consider that the monetary evaluation of gains and losses is positive wealth elastic (“In particular, the Kaldor–Hicks test is generally biased in favour of the status quo because the diminishing marginal utility of money normally results in the winners’ positive and the losers’ negative equivalent-dollar valuations of a policy’s being positively wealth-elastic. Thus, to the extent that a choice’s winners’ equivalent-dollar gains are positively wealth-elastic, the correct answer the winners will give to the
and losses are typically measured in terms of willingness to pay. Since willingness to pay depends on the ability to pay, it is easy to observe that less wealthy people, who by definition have more severe budget constraints than the rich, are normally less willing to pay than the latter. Thus, it is predictable that KH leads to favour the allocation of resources to the rich. However, in that manner, the rich become even richer, thereby further acquiring higher willingness to pay relative to the poor. The acquired higher willingness to pay will again favour the rich in future resources allocations, thereby allowing for higher willingness to pay and so on. As a result, “the rich would get even richer, and their ability and willingness to pay would rise continually”. In this sense, one could say that KH is affected by a bias against the poor.

1.3.3.4. Legal Change

KH implicitly assumes an institutional decision-making process that is apt (i) for measuring losers’ loss of welfare and (ii) evaluating the plausibility of compensating individuals’ losses hypothetically. If this assumption is dropped, the principle of hypothetical compensation turns into nonsense. In other terms, KH poses a fundamental question of institutional choice. In particular, it raises the problem of identifying the institutional process that can better measure the gains and losses in resource reallocation and evaluate the plausibility of compensatory measures. That is, the insurance that KH-superior moves effectively improve social surplus hinges on the choice of institutional mechanism for evaluating the losses and the resulting compensation. Only the comparative institutional analysis of alternative decision-making processes is apt for identifying the conditions necessary to enhance social welfare. The principle of hypothetical compensation does not capture the institutional problem associated with gain–loss measurement. In cases where the informational and measurement problems do
not allow for an assessment with certainty of the magnitude of gains and losses and the hypothetical compensation characteristics, KH does not ensure a potential Pareto improvement in social welfare. Similarly, when the costs of the compensation mechanism are likely to be greater than the benefits of resource reallocation, KH might lead to wrong results.

1.3.3. Wealth Maximization

Richard Posner’s scholarship\textsuperscript{114} provides a theory of WM that represents what law and economics practitioners typically mean when referring to economic efficiency.\textsuperscript{115} Posner’s WM theory rests on a positive and a normative claim. The positive claim simply states that the common law is efficient. The normative claim contends that the law should be grounded in the idea of WM. Posner turns WM into an “auction rule”: a legal entitlement should be conferred on the party who would have purchased it, if the level of transaction costs had not made it inefficient to do so. This criterion is typically referred to as the “highest bidder” rule because it aims to replicate a market mechanism in which the right is allocated to the person who is willing to pay the highest price. Willingness to pay is the measure employed to identify who values the right the most with a view of pursuing allocative efficiency.

1.3.3.1. Logical Criticisms

The major criticisms of WM come from its reliance on the price system that generates the fundamental problem of circularity. As repeatedly emphasized in the prior discussion, efficiency is a right-specific concept. The existence of an initial set of individual legal entitlements is a necessary prerequisite for operationalising any efficiency criterion. This applies to WM as well, which is based on market prices: an antecedent specification of


rights is necessary to determine those prices on which WM relies upon. Therefore, WM cannot be employed to determine rights on which it ultimately rests.\textsuperscript{116}

1.3.3.2. Social Welfare

Wealth is an “instrumental” value. It serves as a means to promote other values. As Coleman puts it, “Wealth is not something of intrinsic value. If the pursuit of wealth is a good, it must be because pursuing wealth promotes other things of value”.\textsuperscript{117} The instrumental nature of wealth generates the need to further specify wealth as the object of maximization. In fact, WM leads to different outcomes depending on the distributional elements incorporated in the definition of wealth\textsuperscript{118} and the assumptions about initial rights distribution.\textsuperscript{119}

1.3.3.3. Distribution

Like KH, the failure to recognize the distributional effects of WM leads WM to privilege and reinforce the existing wealth distribution.\textsuperscript{120} The measurement of increase in the aggregate wealth level (“the size of the pie”) rests upon market prices. Yet, as already noticed, prices reflect a given distribution of wealth, so measuring preferences of people

\textsuperscript{116} Coleman explains the incapability of wealth maximization to generate an initial assignment of rights; see “Efficiency, Utility” \textit{supra} note 115 at 524 (“The problem of relying on prices arises once again when we reach the question whether one could employ wealth maximization to assign basic entitlements. Wealth maximization requires a fixed set of relative prices. The prices of goods depend, among other things, on the relative demand for them. The demand for goods depends in turn on the distribution of wealth. And the distribution of wealth is of course a function of what individuals are entitled to. Therefore, the system of wealth maximization must presuppose a set of initial entitlements in order to get started; and these initial entitlements cannot, by hypothesis, be accounted for on wealth-maximizing grounds. The system of wealth maximization therefore cannot provide a basis for an initial assignment of entitlements.”).

\textsuperscript{117} Ibid. at 527

\textsuperscript{118} See Warren J Samuels, “Maximization of Wealth as Justice: an Essay on Posnerian Law and Economics as Policy Analysis” (1981) 60 Tex. L. Rev. 147 at 150 (“[…] Wealth can be specified in different ways, depending in part upon the distribution of income and wealth. Changes in distribution will affect the substance of wealth, the object of maximizing behaviour. Whether behaviour is rational remains dependent on the specification of its goal, and this is no less true if the goal is wealth, for wealth is necessarily subject to further specification”).

\textsuperscript{119} In this respect, see Lucian A Bebchuck, “The Pursuit of a Bigger Pie. Can everyone expect a Bigger Pie?” (1980) 8, Hostra L. R. 671 at 684–87 [“The pursuit”] which demonstrates the “indeterminacy” of WM. Bebchuck’s idea of indeterminacy corresponds to the general principle that efficiency is a “relative concept”; namely, WM yields different results depending on the initial distribution of legal entitlements. Put succinctly, to not be indeterminate, WM requires a previous specification of the initial assignments of legal entitlements.

\textsuperscript{120} See Bebchuck, “The Pursuit”, \textit{supra} note 119 at 672. Bebchuck demonstrates that WM is biased against the poor in favour of the wealthy.
by the mean of market prices implies the assumption of status quo as exogenous to the wealth-maximizing decision.\textsuperscript{121} The preceding observation is a related to a further criticism. One of the criteria for a sound definition of efficiency is a prior definition of the interests that count. In this respect, the WM logic does not provide any selective criteria for determining whose interests should count in the WM calculus. Rather, WM is based on a tacit determination of the relevant interests that is reached outside of the WM logic. Only after the relevant interests have been selected (i.e., the decision on which interests should amount to a legal consideration as “legal rights”), the efficiency of alternative allocations of legal rights enter into the maximizing calculus.\textsuperscript{122} From this perspective, Posner’s WM theory would be better understood in the light of an explicit discussion of his tacit normative assumptions about the relevant interests to be counted as the ultimate object of WM.

1.3.3.4. Legal Change

Posner’s WM theory does not provide any positive account of judicial behaviour that can explain why and how courts should maximize wealth.\textsuperscript{123} The positive claim that common law is an efficiency-seeking system is grounded on weak theoretical foundations.\textsuperscript{124} Similarly, the normative claim that common law ought to be shaped so as to maximize social wealth seems to be lacking sound theoretical foundations. Without a coherent theory of judicial behaviour, which identifies the incentives of the actors involved in the adjudication process, Posner’s normative claim bears the features of an ideological

\begin{footnotesize}
\begin{itemize}
    \item \textsuperscript{121} See, Bebchuck “The Pursuit”, \textit{supra} note 119 at 684–87
    \item \textsuperscript{122} On this point, see Samuels, “Posnerian Law”, \textit{supra} note 115 at 112 (“Maximization of wealth in practice is functional with regard to and tautological with the determination of rights (of whose interests are to count) which drives the particular result said to be wealth maximizing. […] The logic of wealth maximization does not give a clear and unequivocal solution in any case without additional selective normative premises regarding interests”).
    \item \textsuperscript{123} Here, I refer to the theory formulated by Posner in his classic contributions, Posner “Utilitarianism” \textit{supra} note 114, and “The Ethical” \textit{supra} note 114. More recently, Posner has devoted attention to the explanation of judicial behaviour and to the identification of the structure and features of the utility function of judges. See, Richard Posner, \textit{How Judges Think}, (Cambridge, Massachusetts: Harvard University Press, 2008) [Posner, “\textit{How Judges Think}”]. However the results of these works have not been integrated in a coherent normative theory of how wealth should be maximized through the judicial process.
\end{itemize}
\end{footnotesize}
assertion rather than of a sound prescriptive principle.\textsuperscript{125} The lack of inquiry into the structure of the (common law) process, which is claimed to lead to efficient outcomes, prevents Posner from looking at the relative merits of judge-made law with respect to alternative sources of law. In brief, as Komesar explained extensively, Posner’s analysis is single-institutional.\textsuperscript{126} There is no comparative analysis of the relative advantages of common law over alternative law-making institutions. Thus, Posner’s positive claim falls into the nirvana fallacy.

Therefore, when arguing in favour of the hypothesis of the efficiency of common law, Posner indulges in a methodological error: he infers the characteristic of the process from the quality of the outcomes. The observed allocative efficiency of legal output leads Posner to conclude that the judicial process yields efficient results.\textsuperscript{127}

1.3.4. Social Welfare Maximization

Kaplow and Shavell - in their monograph Fairness versus Welfare (hereinafter: “FVW”)\textsuperscript{128} - maintain that law should promote social welfare regardless of any

\begin{footnotesize}
\begin{enumerate}
\item[125] See Nicholas Mercuro and Timothy P Ryan, Law, Economics and Public Policy, (Greenwich CT: JAI Press, 1984) at 125 [“Law”] (“Much of the criticism levelled at Posner’s concept of wealth maximization can be traced back to the fact that judicial behaviour, as yet, is not well understood and, as such, is not yet amenable to the economic analysis of the common law, positive or normative. […] an explicit, plausible, causal explanation of the behaviour of judges has not yet been established. That is the linkage between (a) the behaviour of individual participants in the legal-economic arena and (b) the incentives and thus the behaviour and decisions of judges is not well developed.”)
\item[126] See Komesar, Imperfect Alternatives, supra note 7 at 17–28
\item[127] On this point, see Mercuro “Law” supra note 125 at 129. In addition, it should be also noticed that this inversion of the process-outcome relationship is typical of the Kaldor–Hicks criterion.
\end{enumerate}
\end{footnotesize}
consideration of fairness. They assume social welfare “to be an increasing function of individuals’ well-being and to depend on no other factors”\(^\text{129}\). The related notion of well-being encompasses “everything that is of concern to an individual”\(^\text{130}\). This expanded definition of well-being allows the inclusion of _distributive_ concerns within the social welfare theory.\(^\text{131}\)

1.3.4.1. The Chore Thesis of “Fairness Versus Welfare”

FVW can be summarized as follows.\(^\text{132}\)

(i) _Normative Premise_. Individuals’ well-being is a desirable goal of legal policy-making.

(ii) _Central Argument_. Pursuing fairness _independently_ of considerations of individuals’ well-being results in a reduction of social welfare.

(iii) _Central Claim_. “[…] legal rules should be selected _entirely_ with respect to their effects on the well-being of individuals in society; […] the notion of fairness like corrective justice should receive no _independent_ weight in the assessment of legal rules”.\(^\text{133}\)

---

Kaplow et al. “Fairness” _supra_ note 128 at 24

Ibid. at 16

Distributional concerns play a role within the operation of the social welfare function in three related ways. First, the social welfare function is based on the principle of the decreasing marginal utility of income. Second, holding constant the total well-being of individuals, social welfare might be greater when individuals’ well-being levels are less unequal than when they are more unequal (Ibid. at 30). Namely, the distribution of well-being affects the social welfare function independently of the decreasing marginal utility of income. Third, an individual’s well-being is directly affected by his or her preferences about income distribution (Ibid. at 31) For instance, individuals might have “a taste for a notion of fairness, just as they may have a taste for art, nature, or fine wine. […] In such cases, satisfying the principle of fairness enhances the individual’s well being, just as would satisfying his preference for wine” (Ibid. at 21).

The central claim is normative in nature: it is a value prescription of the objective of legal policies. The central argument is logical: it concerns the relationship among two alternative normative standards with the value prescription contained in the central claim. On this distinction see Mattiacci, _supra_ note 128 “Godel” at 501

Kaplow and Shavell, _Fairness Versus Welfare, supra_ note 128 at 4 (emphasis is mine). FVW provides a second argument concerning the rationale of fairness and a positive explanation of the reason why scholars, policy analysts and common people are usually attracted by non-welfarist and non-consequentialist views, and find them appealing as a normative criterion for evaluating legal rules. I will not discuss this second argument, which is beyond the scope of my analysis; for a critical discussion see,
FVW provides a negative definition of fairness including all the normative conceptions which are not exclusively concerned with the impact of legal policies on individuals’ well-being. Based on this definition, FVW makes two fundamental claims: (1) the pursuit of fairness can lead to the choice of legal rules resulting in the reduction of social welfare; (2) when parties are symmetrically situated and fairness-based prescriptions differ from those of welfare economics, pursuing any notion of fairness always makes everyone worse off. This last stronger claim is crucial in FVW. To support it the authors introduce a line of reasoning that brings a procedural dimension to the normative debate in law and economics. This is a decisive point in the debate on efficiency as a legal concern and is worth attention.

1.3.4.2. Criticism

The central claim of FVW is that the impact of SW should be the exclusive concern when evaluating legal policies. The central argument supporting the claim is that fairness, as an independent normative criterion, prescribes solutions that make everyone worse off (hereinafter, the “central argument”). One major criticism of FVW is that the logical structure of its central argument constitutes a tautology: to the extent fairness is incorporated in the preferences of people (i.e., it is a “constituent” of SW), it is compatible with the pursuit of social welfare; when fairness is weighted up independently of people’s well-being, it is in conflict with social welfare. This reasoning is devoid of real prescriptive content, and, as such, does not support the central claim of FVW (i.e., that welfare should be the exclusive standard for assessing legal rules and policies). As Coleman has demonstrated, to support the central claim some preliminary logical steps are needed. First, Kaplow and Shavell should identify the requirement of any criterion

---

\(^{134}\) The symmetric setting is one in which each actor is subject to the same probability of finding himself in any of the possible future anticipated situations to be regulated. For further discussion of the notion of symmetric setting, see infra note 160.

\(^{135}\) FVW has been the object of many lines of attack Here, I focus only on those criticisms concerning the relation between their social welfare and distributional concerns.

\(^{136}\) FVW is based on the distinction between distribution as a constituent of social welfare (I borrow this expression from Coleman, “The Grounds”, supra note 128 at 1523) and distribution as an independent principle for legal policy evaluation.

\(^{137}\) Ibd. at 1522
suited for assessing the legal rules, and explain why fairness and social welfare are adequate for a normative evaluation of legal policy. On the contrary, FVW assumes from the beginning that the range of possible normative views comes down to either welfare–based or fairness-based criteria, and do not distinguish among the many different theories that can be included in the two general categories. Second, once explained why social welfare and fairness are appropriate standards for evaluating legal rules, a criterion for choosing between the two normative standards is required. The central argument is that fairness ends up reducing social welfare, but there is no explanation why social welfare should be the appropriate exclusive criterion for the evaluation of legal rules. The reader is left without an account of the value of social welfare as a sound theoretical basis for evaluating legal policies. Its adequateness to ground a normative law and economics approach is assumed, but never explained or demonstrated. Third, the only concern guiding the reasoning is to show that non-welfarist (and non-consequentialist) views lead to a reduction of the social welfare. But since no criterion of choice between fairness or social-welfare is established, what FVW demonstrates is just that each evaluation exclusively based on either fairness or welfare is made at the costs of the unchosen standard of choice. The result is a tautology devoid of content, which can be logically reversed and still remain valid. On the ground of the argument provided in FVW, both the following propositions are equally sustainable: welfare-based policies might reduce fairness, or, alternatively, fairness-based policies might reduce social-welfare. The tautological nature of the central argument of FVW is also reflected in the adopted definition of fairness as opposed to social welfare. FVW provides a negative definition of fairness

138 See Dari Mattiacci, “Godel” supra note 128, at 504-506 that demonstrates how the central argument (that he defines “the logical claim”) derives its soundness from (a) the value judgments assumed by FVW as normative starting points, and (b) by the way the concept of fairness and welfare are defined.

139 See Koleman, “The Grounds”, supra note 128 at 1528 (“Kaplow and Shavell […] fail to appreciate that if what they offer up is a good argument, it is as telling against the welfarist as it is against the deontologist. For if conforming to justice can greatly diminish welfare, it is equally true that pursuing welfare can greatly diminish justice. The very same counterexamples to utilitarianism that presuppose the extent to which one must forgo welfare to conform to the demands of fairness can be read as indicating the extent to which pursuing welfare imposes costs on fairness. Isn’t that just the point of the punishing-the-innocent kind of example? That is, if all we have in mind is pursuing welfare, then we run the risk of imposing the greatest sorts of injustices—including punishing those we know to be innocent of wrongdoing. Similar remarks are in order for all such counterexamples to utilitarianism. In just the same way that they require the deontologist to confront the cost to welfare of conformity with justice, they force the welfarist to confront the cost to fairness of a single-minded pursuit of welfare. That is why so much ink has been spilt on these examples. They vividly raise the conflicts between two different conceptions of right action, and the costs of a single-minded devotion to either.”)
fairness as including all normative views that do not consider the impact on individuals’ well-being. Constructing fairness as the negative image of social welfare is the logical premise for a reversible statement. In conclusion, the only point FVW demonstrates is that welfare and fairness - conceptualized as opposite views – are in conflict when it comes to assessing alternative legal policies. But there is neither an elaborated definition of the two normative standards, nor a sound discussion of their value and the way they should be compared. What Kaplow and Shavell explain is that fairness and welfare – as they are defined in FVW - are mutually conflicting.

Conclusions

The following summarizing conclusions can be drawn from the foregoing discussion.

(i) The logical features of efficiency expose the economic analysis of legal rules to the risk of tautological propositions, mostly predetermined by tacit normative assumptions. To avoid the risk of logical circularity, any definition of efficiency should preliminarily clarify the following elements: (a) Output, understood as the object of the maximization process, with all its normative specifications; (a) Set of Choices, i.e. the set of alternative viable economizing solutions; (c) Relevant Interests, i.e. the set of individuals whose interests are served by efficient resource use.

(ii) While conventional output-oriented scholarship focuses on the allocative efficiency of legal entitlements, my proposed structural approach focuses on productive and dynamic efficiency. This shift of focus allows one to account for the efficiency implications of the process of producing legal rules and enables economic analysis to gain a better

---

140 *Ibid.* at 1539 (“On their view, it is not necessary to explain what fairness is or why pursuing fairness is valuable because their argument is that whatever fairness is and whatever is valuable about it, pursuing it is detrimental to welfare.”) This point is well clarified in the Dari Mattiacci

141 *Ibid.* at 1528 (“All Kaplow and Shavell have provided are vivid examples of how the conflict between welfare and fairness—a conflict that is inevitable, given the way they conceptualize the two—might play out in actual practice) and *Ibid.* at 1539 (“If truth be told, the only serious claim they make is that fairness is detrimental to welfare […]”).

See also Dari Mattiacci, “Godel”, *supra* note 128 at 506-507 (“[…] FVW does not prove that the maximization of individuals' well-being ought to be the only criterion of policy assessment. Rather, FVW proves that if the maximization of individuals' well-being is taken as the aim of social decision-making and if welfare economics and notions of fairness are defined as explained, then the consequences that follow from the implementation of notions of fairness are inconsistent with (contradict) the outcomes of the welfarist criterion.”).
understanding of the impact of legal rules on the magnitude of social surplus. In the following section, I will expand on this latter point by clarifying the objectives and method of process efficiency analysis.

(iii) The output-oriented approach is susceptible to the four following criticisms: (a) logical circularity or logical incompleteness; (b) failure to account for increases in social welfare; (c) failure to account for the presence of losers; (d) failure to account for the dynamics of legal change. The reader will notice the correspondence between these major criticisms of the substantive efficiency criteria and the previous discussion of the logical features of efficiency in part I.
Chapter 2
PROCESS EFFICIENCY

 [...] man is a frivolous and unseemly being, and perhaps, similar to a player, likes only the process of achieving the goal, but not the goal itself.

Fyodor Dostoevsky, Notes from Underground, 1864

After having identified the problematic nature of legal efficiency and the limitations of the traditional output-oriented approach, this chapter introduces the theoretical foundations of my proposed structural approach based on the distinction between “output efficiency” and “process efficiency”. In particular, it (i) provides a definition of process efficiency, (ii) develops a taxonomy of law-making costs, and (iii) demonstrates that the relationship between these two efficiency dimensions should be understood from the perspective of social welfare maximization.

The discussion proceeds as follows. Sections 1 and 2 provide alternative arguments supporting my central claim that process efficiency and substantive efficiency are related causally. Section 1 elucidates the relationship between process efficiency and output efficiency through the Edgeworth Box model of economic exchange between two individuals in a setting preceding the establishment of legal order. The discussion in this section borrows from Buchanan142 some insights on the role of rules of the game in the definition of the set of optimal allocations. From this perspective, the discussion demonstrates that outcome-oriented analysis must be integrated with a complementary inquiry into the choices occurring at the “constitutional” level.143 Section 2 relies on a simplified model of individual constitutional choice to explain and identify the two

143 In this context, I use the expression “constitutional” in the specific sense of “constitutional stage of choice”.
components of the process efficiency problem (i.e. minimization of law-making costs and outcome quality). It also emphasizes that the input–output separation—implicitly assumed by the conventional law and economics approach—cannot be properly imported into the realm of legal theory without careful adaptation. In the remaining two sections, I construct the analytical tools for process efficiency analysis. Section 3 defines the notion of process efficiency and constructs a taxonomy of law-making costs. Finally, section 4 identifies the logic underlying the process efficiency analysis, which distinguishes my approach from other strands of scholarship. In particular, it defines my proposed efficiency test, discusses the assumption of collective ignorance of the social welfare function, and clarifies the concepts of demand for and supply of law.

Introduction: Law-making and Social Welfare Maximization

I assume that the ultimate goal of law-making is the maximization of social welfare, which I define as the vertical sum of individuals’ economic surpluses. Although contentious, this definition reflects the fundamental tenet of mainstream law and economics that the law should allocate rights to the party who values them the most. In particular, it is traditionally thought that legal rules lead to the maximization of social surplus, intended as the joint surplus of the actors involved in a regulated situation. Under these assumptions, the lawmaker maximizes social welfare through the creation and

---

144 This definition differs from Kaplow and Shavell’s idea of social welfare as the sum of individuals’ well-being (Fairness Versus Welfare, supra note 128 at 24). Subsequently in this chapter (sub-section 2.4.2.), I discuss in greater detail the assumption of a social-welfare-maximizing law-making process. For an excellent discussion of the methodological difficulties associated with the definition of the concept of social-welfare, see Richard B. Brandt, “The Concept of Welfare” in Sherman Roy Krupp (ed.), The Structure of Economic Science (Prentice-Hall: London, 1966).

145 The summation is vertical because law is a public good characterized by non-rival consumption. The social benefit of a public good is the sum of the marginal benefit obtained by each individual belonging to the community to which the public good is provided. Thus, to obtain the aggregate welfare, we need to use vertical summation of the individual marginal benefits for each produced level of good. In the case of legal rules, the aggregate benefit is the sum of individual surpluses (i.e. substantive efficiency, as I will explain shortly below) generated by the application of legal rules. Different is the case of the aggregate benefit of private goods characterized by rival consumption. In the latter case, the summation is horizontal because only individuals who buy the good gain utility from consumption; hence, aggregation follows the number of goods bought by consumers in the market, rather than the number of individuals in a society.

146 On economic surplus as a measure of social welfare see, generally, Currie, John Martin, John A Murphy, and Andrew Schmitz “The Concept of Economic Surplus and Its Use in Economic Analysis” (1971) 81 The Economic Journal 741. For the larger debate that was held in the 1960s and 1970s on the broader, controversial issue of the possibility of a social welfare function, see infra note 229.
enforcement of efficient legal rules that maximize the joint surplus of individuals who are subject to the law. That is, the allocative efficiency of legal rules becomes a means of maximizing social welfare. This component of the economic efficiency of legal rules is the traditional object of inquiry in the mainstream law and economics approach. I will call it “substantive” efficiency, in that it relates to the decision of who should be the owner of legal entitlements and how the law should most efficiently protect the assigned legal rights.

Scholars have traditionally investigated the problem of substantive efficiency by focusing on the “output” of law-making, as if substantive decisions were taken exclusively at the moment of defining the content of legal entitlements. Proceeding this way, substantive efficiency analysis has been mostly confined to the allocative efficiency of legal rules. In this sense, I define the traditional law and economics approach as “output-oriented”. From a different perspective, I emphasize the causal relationship between substantive efficiency and the institutional characteristics of law-making. Institutional actors make choices about and within the law formation process itself, thereby affecting the efficiency of the final outcome of law-making (i.e., substantive legal rules). I denote the entire set of efficiency issues related to these process choices as “efficient production of legal rules” or, more briefly, “process efficiency”. My fundamental claim is that we need to account for the relationship between “process” efficiency and “substantive” efficiency, if we want to better understand the manner in which the production of efficient legal rules can enhance social welfare. The investigation of this relationship identifies factors

---

147 In this context, the individuals subject to the law are the actors whose behaviour is regulated by the law (i.e. the “target” actors). Subsequently, in this chapter, I will provide a more complete definition of the concept of “people subject to law”.

148 As explained previously, the reasons behind this methodological attitude lie in the pretence of a value-free economic analysis of law. Here, I criticize the implicit assumption that economic efficiency arises exclusively from the structure of legal entitlements. This assumption has led legal and economic scholarship to devote little (or non-systematic) attention to the organization of the decision-making activities leading to the decision about the allocation of legal rights and how different organizational patterns influence the economic quality of the output. To use neoclassical language: a great deal of attention has been devoted to the downstream market of legal rules, while less systematic investigation has been conducted on the upstream market of law production.

149 As I will explain later, the sweeping notion of process efficiency includes a heterogeneous array of economic goods. For explanatory convenience, I group them in one broad category.
influencing the economic efficiency of legal rules that are not usually captured by the mainstream output-oriented approach.\textsuperscript{150}

2.1. Pareto Optimality and Process Choice

Buchanan raised a critique of PO that proves useful for economic analysis of the law-making process.\textsuperscript{151} Buchanan’s reasoning provides insightful arguments demonstrating that allocative efficiency cannot be investigated properly without including in the economic calculus the costs associated with the institutional process of generating legal rules. Buchanan remains in the analytical domain of allocative efficiency without embracing any notion of “process” efficiency. However, Buchanan criticizes the use of the Pareto criterion as a test for classifying the final allocative positions or results by maintaining that PO should be employed to capture the increases in efficiency that can be obtained by changing the institutions that shape the legal rules within which resource allocation takes place.

Buchanan contends that the constitutional choice of alternative law-making processes should be included in the scope of Pareto efficiency analysis. That is, the Pareto efficiency test should take on a two-pronged structure: (1) choice of optimal allocation under existing institutional constraints and, (2) choice of optimal allocation considering also those allocations unattainable under existing institutional constraints but attainable under alternative institutional frameworks agreed upon unanimously by the members of a community at the constitutional stage of choice.\textsuperscript{152}

The important implication of Buchanan’s critique is that once the costs of organizing the law-making process are factored into the efficiency calculus, the optimal rule might be the one that conduces to non-optimal allocative results according to the standard

\textsuperscript{150} Hereafter, I will use the expression “substantive-efficiency” and “output-efficiency” interchangeably. Conceptually, they are distinct. However, since substantive efficiency has been investigated by focusing on the output and by insulating it from its formation process, I will refer to substantive efficiency as the mainstream idea of allocative efficiency of law-making output.

\textsuperscript{151} See, Buchanan, “The Relevance” see supra note 142.

\textsuperscript{152} Ibid. at 346–347.
outcome-oriented notion of Pareto efficiency. Stated differently, once Paretian logic is extended to the constitutional stage of choice, the optimality normative criterion may conduce to inefficient allocative results in a number of cases. Buchanan’s explanation of this crucial point is not comprehensive as it remains limited to the allocative efficiency analytical domain. Consequently, in this study, I introduce the notion of process efficiency and provide various explanations of why process efficiency often entails outcome inefficiency.

2.1.1. Pareto Region, Pareto Optimality, and Rules of Game

The premise of the reasoning is that the rules that define the institutional constraints within which parties carry out their exchange (hereinafter, also “rules of the game”, or “ROGs”) affect the shape of both the Pareto Region and the set of optimal points. The Pareto Region is the “region within which the Pareto criterion for classifying positions is to be employed”, i.e. the set of possible allocations within which the Pareto efficiency criterion identifies a given unique optimum. Traditional welfare economists often tacitly assume that people operate within a set of mutually respected rules that enable them to cooperate and engage in reciprocally advantageous exchanges. According to Buchanan, this traditional use of the Pareto criterion is pointless in that it does not consider possible variations of the Pareto region induced by changing institutional constraints. Instead of making tacit assumptions about the institutional constraints within which people organize their activity, the rules of the game should be included explicitly in the scope of the Pareto efficiency analysis.

Following the idea of including in the efficiency analysis the points attainable through changes in the existing set of institutional constraints, the PO criterion should be applied to demonstrate that people might improve their cooperative surplus by changing the actual set of ROGs. Put differently, the Pareto test should be employed so as to verify that

---

153 Buchanan defines the “rules of the game” or, more briefly, “rules”, as the “standards of conduct applicable to all the members of the social group”, Ibid. at 342

154 Ibid. at 241

155 For example, the conventional model based on the Edgeworth Box, an analytical tool typically employed by economists to explain the concept of economic efficiency, employs the implicit assumption that individuals interact with each other in a legal environment that enforces contractual agreements, and provide sanctions against fraud, theft, and other types of merely redistributive appropriation.
there is room for Pareto-superior exchanges, not only within the set of points located along a *given* possibility frontier (i.e., as determined by implicit assumptions regarding the institutional constraints), but also by modifying the Pareto region through a change in the set of ROGs within which individuals conduct their trading activities.\(^{156}\)

To illustrate the implications of a change in the rules of the game on the size and shape of the Pareto Region, Buchanan offers a helpful example. Let us consider two individuals, A and B who produce and consume two goods x and y in a situation preceding the establishment of the legal order. The area defined by the Edgeworth Box represents the consumption choices of A and B for two goods x and y. Each point within that area represents a feasible combination of consumption goods; that is, a possible pair of bundles chosen respectively by individuals A and B. See figure 5.

\(^{156}\) In Buchanan’s words, the Pareto criterion could be applied through a two-step process: (“First, within the existing set of rules, the classification may be helpful in dispelling ignorance and irrationality […]. Secondly, the criterion may be used to classify points, unattainable under existing rules, but which, conceptually, may be attained through the adoption of rules which themselves can be agreed upon by all members of the group”) (*Ibid.* at 346).
Figure 5: Edgeworth Box with changing rules of the game
A is more skilled in producing y, while B is more skilled in producing x. So, the possible initial endowment of A contains a greater quantity of good y, whereas the possible initial endowment of B contains a greater quantity of good x. The two lines P - Pa and P - Pb represent the production possibility frontiers of A and B, respectively. Buchanan describes three different scenarios characterized by three different ROGs and shows the impact of a change in the ROGs on the shape and the extension of both the Pareto Region and the set of optimal results.

Scenario 1: rules prohibit trade between A and B.
Since A and B cannot make any exchange, the only way to improve their utility is to move toward the production possibility frontier and attain the point of tangency with the highest indifference curve. Hence, in this case, the Pareto region is the set of points within the area defined by the production possibility frontier; the Pareto Region of A is the triangle POaPa and that of B is the triangle PObPb. Once the Pareto region is identified, it is easy to recognize that the positions below the possibilities frontier are non-optimal. For example, the point (xa, ya1), which is below the possibilities frontier of A, is not an optimal choice for A that can improve his level of utility by producing more units of both goods x and y. A reaches the individually optimal production and consumption at (xa, ya) on the possibilities frontier, and, similarly, B reaches the individually optimal production and consumption at (xb, yb). Finally, from the viewpoint of joint production and consumption, the Pareto-optimal level of output under the given institutional constraint is given by (xa, ya) + (xb, yb).

Scenario 2: rules enforce contractual agreements.
Because of freedom of contract, the two individuals A and B can specialize in the production of goods in which they have a comparative advantage and buy the needed amount of the second good from the other individual. In this case, individual A specializes in the production of good y; individual B specializes in the production of good x; and through mutually advantageous exchanges, they can reach a new set of positions

---

157 For simplicity, I assume a constant marginal rate of technical substitution between the two goods so that frontier possibilities are depicted as straight lines. Notice that the different slopes of P - Pa and P - Pb depend on the skill differential between A and B in producing goods y and x.
within the area coloured in yellow (i.e., the “trade area”). Within the defined trade area, A and B will trade until the marginal rates of substitution of the two individuals are equal to each other.

In this scenario, as an effect of a change in the ROG, the Pareto region expands compared with that in scenario 1. It now includes all points delimited by the possibility frontiers, as in the previous case, but extends to all points within the area in which the two individuals can gain through mutually advantageous exchanges (i.e. trade area). At the same time, the set of optimal points changes; individuals will reach the optimum at the point of tangency of the two highest indifferent curves along the portion of the contract curve between $\alpha$ and $\beta$. This point will be located in the trade area.

It is important to emphasize that under the regime of enforceability of contractual agreements, the new expanded Pareto region includes all points of the Pareto region of scenario 1 in which trade is not allowed. Secondly, the set of Pareto-optimal points reachable through exchange does not intersect the set of optimal points in scenario 1.

**Scenario 3: there are no rules of game.**

In this case, people either cooperate through specialization and exchange (as in scenario 2) or they appropriate the resources of others through robbery, fraud, and other types of violent appropriative behaviour. In other words, there are no institutional or social constraints on individuals’ behaviours that define and limit the Pareto region. Therefore, the set of all possible attainable points within which the Pareto criterion is to be employed is defined by the entire Edgeworth Box. This scenario resembles the situation preceding the establishment of legal order, previously described in scenario 1. In the case where one party, say A, is stronger that the other, say B, the individual equilibrium will be close to Ob. In this case, the optimal point is the one in which A exploits all the resources at the expense of B. In this scenario, the set of Pareto-optimal positions is represented by the points lying along the locus $Oa\alpha\beta Ob$. It is clear, then, that the unconstrained setting is not meaningful for efficiency analysis because the Pareto region
is undefined, and the set of optimal points end up including those which are optimal in the second case as well.

The important lesson to be drawn from the analysis of the above three cases is that a change in the institutional constraints directly affect the size and shape of the Pareto region (i.e., all points attainable by the parties) as well as the set of Pareto-optimal points (i.e., points that parties actually choose to maximize their gains). It follows that the tacit assumption that the characteristics of the rules are exogenous to the model ends up confining the efficiency analysis to a given strictly confined Pareto region, and, ultimately, predetermines its findings. In other words, once it is assumed that actors are fully rational and fully informed, and once the Pareto region is narrowed, it is of little value to draw the logical conclusion that economic actors will maximize their expected utility by attaining the set of Pareto-optimal points. On the contrary, if we assume that

---

158 The Pareto-region is undefined in the sense that it is not distinguished from the area included in the Edgeworth Box.

159 According to Buchanan, within a given Pareto region any position voluntarily chosen by individuals will be Pareto-optimum, unless limited rationality or information precludes the identification of efficient allocation: (see Ibid. at 343: “Once a region is defined, the Pareto criterion may be employed to classify all points or positions into two sets, those that are optimal and those that are ‘non-optimal’. In a region so strictly confined, however, any position finally attained will be in the Pareto-optimal set of positions, except for ignorance and irrationality in the behaviour of some or all individuals in the group. No paradox is involved here when it is recognized that each individual, if rational and fully informed, will maximize his expected utility within the constraints imposed on his behaviour by the existing social rules. If each individual does this, there is no change, within the existing rules, that can make anyone better off. The Pareto criterion, in this situation, has relevance only insofar as it provides some assistance in dispelling areas of ignorance or in removing certain barriers to rational behaviour”)

Note that one might object that Buchanan is wrong in stating that the traditional Pareto criterion is pointless under the assumptions of unbounded rationality and perfect information. In fact—one might object—there are “market failures” that are independent from individuals’ bounded rationality and limited information. I would reply to this objection with the following two considerations. First, Buchanan clearly refers to the failures of bargaining between two individuals. In this bilateral setting—according to the generally accepted principles of law and economics theory of contract—“bargaining failures” that prevent parties from reaching an optimal allocation of resources are generated by limited information and bounded rationality. In this respect, Buchanan’s criticism seems to be right. Second, what is logically relevant for the purpose of my analysis is not the claim that traditional use of the Pareto criterion is pointless. Rather, the crucial point is that the Pareto criterion should include in its domain the efficiency advantages of alternative rules of the game. Thus, even if the pointlessness objection raised by Buchanan proved misplaced, the insight that I borrow from Buchanan’s reasoning would remain valid for the purpose of my analysis.

Finally, it is interesting to note that this line of critique of the Pareto criterion is quite similar to the one raised by Guido Calabresi, in Calabresi, “The Pointlessness” supra note 27 at 1216: (“Unfortunately, the set of Pareto superior changes which would make no one worse off and at least one person better off must
the institutional constraints are *endogenous* to the model, the Pareto notion of efficiency includes all points attainable through a change in the existing ROGs in its domain. From this perspective, *efficiency analysis identifies changes in ROGs that improve individuals’ cooperative surplus*. The Pareto test is employed to verify whether there is room for Pareto-superior exchanges, not only within the set of points as determined by implicit assumptions regarding the institutional constraints but also through changes in the ROGs (i.e., process choices).

### 2.1.2. Unanimity as Constitutional Constraint on Definition of Pareto Region

The second step of Buchanan’s reasoning is to identify the *changes* in the rules of the game that are relevant to the economic analysis. As noticed above, the assumption that *all* changes are possible leads to an *unconditional* Pareto region whose extension coincides with the entire area of the Edgeworth Box, which ultimately deprives the analysis of any analytical value. For this reason, Buchanan proposes to assume the *unanimity* criterion as the constitutional constraint that allows for the definition of a specific Pareto region.

At this point, a clarification is needed with respect to the distinction between “constitutional” and “institutional” stages of choice. Buchanan does not make this distinction clear; however, for our purposes, it is useful to distinguish the *constitutional* stage of choice in which the members of a community choose the mechanisms for producing legal rules (i.e. law-making institutions) and the *institutional* stage of choice in which members of the community choose (within the varying procedural constraints characterizing alternative law-making institutions) the legal rules governing individuals’

---

ex ante be a void set. For if strict or fanatical Pareto is the criterion, why wouldn't any change that belonged in the set have already been made? Since, by definition, no one would in any way be hurt by the change, why would anyone object? The existence of ex ante objectors itself must mean that there are some people who—rightly or wrongly—believe that they will lose something from the change. Putting it another way: if Pareto optimality means a place where no improvement can be made without ex ante creating the possibility that there will be some losers, then we are always there. And if the strict Pareto test is the appropriate definition of efficiency, then what is efficient, though it may be only one of many such arrangements. Moreover, it will remain efficient at least until a new and better arrangement becomes known and achievable without hurting anyone. Once that happens, we will go there forthwith, without the need for a lot of expensive lawyer-economists to tell us to do it. Why would we ever fail to move to what is unanimously acceptable if we had the requisite knowledge and ability to do it? Why would we need anyone to tell us that we should do it?”).
behaviours. Once the distinction between the constitutional and institutional stages of choice has been clarified, the unanimity criterion proposed by Buchanan can be understood easily. Under this criterion, the optimal ROGs are those that cannot be modified without the consent of all individuals that have to decide at the constitutional choice stage and those that are situated symmetrically.\textsuperscript{160}

Unanimity (at the constitutional choice stage) enables the members of a community to identify the Pareto-optimal rule that, in turn, defines a unique Pareto region. In other words, while the conventional use of PO is confined to output analysis, the proposed criterion includes in its scope the constitutional choice between alternative institutional mechanisms whereby individuals determine the rules organizing their economic activity (hereinafter: “process choice”). From this constitutional perspective, the optimal rule is the one that cannot be modified without the consent of all persons involved in the economic activity. Conversely, a change in the rules upon which there is no general agreement is non-optimal. Crucially, the unanimity criterion is applied at the constitutional stage of choice under the assumption that individuals are “symmetrically”

\textsuperscript{160}The idea of a symmetric setting is the subject of a long analytical tradition. It aims to ensure the impartiality of normative evaluation over the rules regulating the future interaction among the members of a community (e.g. legal rules, rules regulating the distribution of wealth, and so on). Classic contributions in the field are, JC Harsanyi, “Cardinal Welfare Individualistic Ethics, and Interpersonal Comparisons of Utility” (1955) 63 J. Pol. Econ. 315 (the social setting that satisfies the requirement of impersonality of preferences is that in which an individual may recognize what is good from the social viewpoint without regard for their own personal preferences); John Rawls, A Theory of Justice, (Cambridge, Mass.: Harvard U.P., 1971; Oxford: Clarendon Press, 1972) (elaborating the idea of an ideal setting where individuals are under a “veil of ignorance,” thus ensuring a disinterested perspective for the normative evaluation of the relative merits of possible social situations); John C Harsanyi, “Bayesian Decision Theory and Utilitarian Ethics.” (1978) 68.2 Am. Econ. Rev. 223. (developing the idea of a setting in which each actor has the same probability of finding himself or herself in different alternative situations). More recently, from a law and economics perspective, Kaplow et al. “Fairness Versus Welfare” supra note 128 at 57 has suggested that the normative evaluation of legal rules need not serve one particular set of interests but should “proceed from a disinterested perspective” that takes into account a sequence of different cases, thereby emphasizing the condition of symmetric setting and ensuring impartiality and impersonality of judgment. Finally, Parisi has discussed the importance of symmetric settings as a condition for the spontaneous emergence of social norms; see, Francesco Parisi, “Toward a Theory of Spontaneous Law” (1995) 6.3 Constitutional Political Economy 211 (where the Author distinguishes three alternative settings that ensure preference impersonality: (i) structural symmetry (i.e. perfect incentive alignment) (ii) stochastic symmetry (role reversibility) (iii) induced symmetry (i.e. reciprocity constraints)); see also, Francesco Parisi “The Cost of the Game: a Taxonomy of Social Interactions” (2000) 9.2. J. Eur. Law and Econ. 99 (emphasizing the role of symmetric setting conditions in strategic social interactions).
situated\textsuperscript{161}. The importance of this assumption will become clear shortly below. The great advantage of this interpretation of Pareto efficiency is that unanimity is not required for single final allocative results. Rather, unanimity is required in the allocation of law-making power to alternative law-making institutions charged with the production of rules governing a sequence of future cases. \textit{The switch of the unanimity criterion to the constitutional stage of choice changes the nature of efficiency assessment}, which then justifies allocative results that are not efficient according to the standard outcome-oriented notion of PO. Symmetrically situated individuals in the constitutional stage of choice will chooses social-surplus-maximizing rules considering the characteristics of the flow of cases to be regulated in the future and the relative costs of alternative law-making processes.

2.1.3. Optimal Rules of Game Generating Non-Optimal Legal Outcomes

Once Paretian logic is extended to the constitutional stage of choice, there arises the central issue of how the Pareto efficiency of process choice is related to the Pareto efficiency of the final allocative results, i.e. whether the optimal rules identified at the constitutional stage of choice (\textit{i.e.} those that cannot be changed by unanimous consent of individuals symmetrical situated) lead to the optimal final allocative results (that make at least one person better off without making anyone worse off). The analysis of this crucial point sheds light on the connection between the costs of determining the legal rules and the allocative efficiency of legal outcomes. It also reveals that once the costs of alternative law-making mechanisms rules are considered, economic analysis might justify the existence of rules that, in a number of instances, are \textit{non}-optimal in the terms of the standard outcome-oriented Pareto criterion.

There are two fundamental reasons why optimal legal rules might generate non-optimal allocative results. First, once the analysis encompasses the process choice, the scope of the efficiency criterion is not restricted to a specific, individual case; rather, it refers to a sequence of prospective cases. Efficiency analysis is then conducted from the aggregate

\textsuperscript{161}This is an important feature of process efficiency that is emphasized by Buchanan as well (Buchanan, “The Relevance”, see \textit{supra} note 142 at 349–350).
perspective, which requires consideration of the distribution of cases with respect to the variables relevant for regulation. Stated differently, considering the process of producing legal rules requires one to take into account the variance of situations to which the same rule is to be applied to.\textsuperscript{162} This requires estimating both the distribution of people’s preferences and the magnitude of the costs incurred by different individuals for the lack of optimality of single allocative results. I will call these latter costs “outcome inefficiency” to emphasize that they result from the departure of the single allocative result from the usual PO standard.

Second, the process choice depends on the costs associated with alternative law-making institutions, and on their relative abilities of yielding a efficient outcome. Hence, there could be discrepancy between the efficiency of the institutional decision-making process and the efficiency of the allocative outcomes. For example, a given law-making process capable of producing an efficient outcome might be too costly compared with the available alternatives. People might prefer a less costly process even if they anticipate that the final outcome might be inefficient in a number of instances. That is, people might

\textsuperscript{162} A clarification is needed. One could object that we are simply shifting the focus from “single-period” to “long-period” social welfare maximization. From this perspective, if one takes a long-period view then—as a logical consequence—single-period results are irrelevant for defining efficiency, i.e. once a long-period view has been adopted, only long-run efficiency matters, so that single-period outcomes are simply immaterial and out of the long-term efficiency domain, rather than “non-optimal”. However, it would be misleading to interpret this line of reasoning in terms of a shift of focus from single-period to multiple-period. The choice is \textit{between} the two perspectives, not \textit{within} either one or the other perspective; thus, it is perfectly logical to compare the efficiency features of a single-period result to a long-period outcome. Let me explain this point.

Buchanan is interested in comparing alternative law-making institutions characterized by alternative rule-producing mechanisms: some involving the \textit{ex ante} regulation of a sequence of cases and others proceeding on a case-by-case basis. Therefore, Buchanan shifts the focus from a static perspective focused exclusively on the final allocative results to a constitutional perspective that takes into account the law-making institutional process. This entails a shift of focus from the single outcome (“single” because related to a single specific case) to the possible choice of an \textit{ex ante} perspective regulating a “class” of cases. Thus, it is misleading to state that single-period outcomes are immaterial once a long-term perspective is adopted, because the constitutional choice is based on a comparison between (i) the optimal single-period outcome plus the costs of obtaining this single-period optimum in each prospective case, (ii) and an \textit{ex ante} regulation that might determine non-optimal single outcomes in a number of instances but might be far less expensive in terms of law-making costs. Thus, the constitutional calculus takes into account (1) the differential law-making costs between a case-by-case allocation (e.g. individual bargaining) and an \textit{ex ante} allocative mechanism, and (2) the differential in terms of output efficiency between two rule-producing mechanisms entailing alternatively a single-period outcome and a multiple-period outcome. In essence—it is worth repeating—the choice is between the two perspectives, not within either one or the other perspective.
prefer lack of optimality in final allocative results because the saving in production costs outweighs the advantages in terms of outcome efficiency. Therefore, the process choice entails comparing the differentials between the size of the law-making costs and the magnitude of outcome inefficiency of alternative law-making processes.\footnote{I am not claiming that there is always a trade-off between process efficiency and substantive efficiency. Whether or not a trade-off exists is very difficult to say at an abstract level. A theoretically sound definition of process efficiency and a careful case-by-case examination is needed to provide an answer to that question. In the next section, I propose a methodology that enables us to deal with these economics problems.}

To understand why optimal legal rules might generate non-optimal allocative outcomes, a brief examination of three different scenarios is useful.

\textit{Case 1: all members of the group always benefit from rule changes.}

In this scenario, the optimal legal rule produces optimal results in all instances, so that all members of the group benefit from the optimal rule. In this case, there is no mismatch between Pareto-optimal rules and the allocative efficient final results, i.e. the optimal rule produces optimal results because unanimous consensus on the rule coincides with the consensus on final allocative results. An example of this type of scenario might be the change from a rule that prohibits some voluntary agreements to a rule that allows private bargaining. If there is room for bargaining, there is no external effect on third parties, and there are no obstacles to private bargaining, then trade among private parties will take place and conduce to Pareto-optimal allocative results. However, in the real world, situations such as these are infrequent. Usually, the optimal rule (i.e., the one that can only be changed with the consensus of all group members) produces terminal results that are not Pareto-optimal in many instances.

\textit{Case 2: the optimal legal rule generates both optimal and non-optimal terminal results, but the aggregate surplus associated with the optimal results more than offsets the aggregate loss associated with the non-optimal results.}

In this scenario, in some cases, the optimal rules produce results that make some individuals worse off. However, when observed over a sequence of possible
heterogeneous results, the aggregate net benefit derived from the cases in which the rule leads to optimal outcomes outweighs the aggregate net loss associated with non-optimal results.

Let $B$ and $G$ denote, respectively, the aggregate benefit and costs associated with a rule $r$, and $C^r$ denote the costs of producing $r$. $B$ and $G$ are functions of $r$, and $C^r$ is a function of the lawmaking process $\rho$.\textsuperscript{164} Consider the following social welfare maximization function:

$$\max_{(r, \rho)} W = [B(r) - G(r)] - C^r(\rho)$$

(1)

This implies that social welfare maximization requires both (i) maximization of the aggregate surplus associated with final allocative results and (ii) minimization of the total costs associated with production of the law. That is, respectively:

$$\max_{(r)} [B^r(r) - G^r(r)]$$

(2)

and

$$\min_{(\rho)} C^r(\rho)$$

(3)

Let me focus first on (2) and assume $C^\rho$ as a constant. Let me assume also that $r$ generates optimal final allocative results in a first group of cases (i.e., for some individuals) while generates losses in a second group (i.e., impose losses on other individuals). I refer to the former group as the “optimal cases” and to the latter as the

\textsuperscript{164} In the next section of this chapter, I provide a critique of this way of looking at the social-welfare maximization problem. Here, I proceed simplistically by assuming that the allocative surplus is a function of the rule, and the law-making costs are functions of the process, because I want to show that even when anchored to the allocative efficiency perspective, the outcome-oriented perspective is inadequate for evaluating the efficiency of legal rules. In particular, I ignore the “outcome effects” of the law-making process, which constitute the object of investigation in this study.
“non-optimal cases”. The aggregate net benefit is a result of the sum of the aggregate surplus realized in optimal and non-optimal cases. That is:

\[ B(r) - G(r) = (B_o - G_o) + (B_{no} - G_{no}) \]  

(4)

where \( B_o \) and \( G_o \) denote respectively the benefit and cost associated with the optimal allocative results, and \( B_{no} \) and \( G_{no} \) denote respectively the benefit and cost associated with the non-optimal allocative results. For brevity, let me denote \( (B_o - G_o) = S_o \) and \( (B_{no} - G_{no}) = S_{no} \), so that \( (B_o - G_o) + (B_{no} - G_{no}) = (S_o + S_{no}) \).

The condition (2) is satisfied when

\[ \max_{(r)} [S_o(r) + S_{no}(r)] \]  

(5)

This signifies that when a set of possible rules in a given regulated environment produce at the same time optimal and non-optimal allocative results, the social welfare maximizing rule is the one that maximizes the sum of the surplus generated by optimal and non-optimal allocative results. The presence of losers does not undermine Pareto efficiency (in the sense proposed here) as long as (5) is satisfied. In essence, \textit{ex ante} maximization is compatible with the expected presence of losers in a given number of cases. Members of a community agree unanimously \textit{ex ante} on a rule that is expected to result \textit{ex post} in non-optimal final results when it maximizes their expected social welfare, as compared with the available rules.\textsuperscript{165} To clarify, there is a substantial difference between this efficiency criterion and KH efficiency. Pareto efficiency applied at the constitutional stage of choice requires the unanimous actual consensus of individuals; by contrast, KH is based on the presumption of consent through the interpersonal utility comparison.

\textsuperscript{165} Recall that we assume that the individuals choosing the ROG are located in a symmetrical setting, so that they are equally exposed to the risk of being future losers.
Buchanan provides two intuitive examples. The first is the rule that requires drivers to stop at a stop sign. In no-traffic situations, the rule generates non-optimal results; drivers would be better off, and no one else worse off, should the signal be removed. However, in all intense-traffic situations, the costs of the absence of rule are probably so high as to induce members of the community to accept the rule despite its outcome-inefficiency associated with no-traffic situations. Moreover, the costs of reaching a consensus on alternative arrangements (i.e. agreement among all members involved in the intense-traffic situation) would probably exceed the costs of producing the existing rule. For this reason, people at the constitutional choice stage will prefer a centralized body to establish the rules regulating car traffic rather than leave it to private bargaining.

Another example is the move from strict liability to the negligence rule in the context of automobile accident liability. Under a negligence rule, pedestrians incur higher costs of walking because they bear the costs of non-negligent accidents; by contrast, under strict liability, drivers incur higher costs because they are responsible for both negligent and non-negligent accidents. Assume, for simplicity, that the population is divided in two groups of people: those who walk (“walkers”) and those who drive (“drivers”). A move from the strict liability regime to the negligence regime increases the costs of walking and decreases the costs of driving. Thus, walkers are the losers and drivers are the winners; the move to negligence is Pareto non-optimal. However, at the constitutional stage of choice, people do not know whether they will be drivers or walkers (given the assumption of symmetric setting). Since ex post private bargaining in this case is too costly (the number of private parties is too large), people will agree to choose the negligence rule if they estimate that the “number of drivers times the gain they receive” from a move to a negligence rule outweighs the “number of walker times the costs they suffer” from abandoning the strict liability regime. In this case, given that private bargaining is too costly, in a constitutional symmetric setting, individuals will opt for the negligence rule, even if they anticipate that in many instances this will generate non-optimal results. In conclusion, the aggregate perspective entails that the optimality of the

166 This example is discussed in Buchanan, “The Relevance” see supra note 142 at 348.
167 This example is discussed in Bebchuck, “The Pursuit” supra note 119 at 674–675.
rule depends on the anticipated distribution of people (numbers of walkers and drivers), magnitude of gains and losses associated with the lack of optimality (how much do drivers gain and walkers loose?), and costs of alternative law-making processes (an *ex ante* rule of negligence is less expensive than private bargaining).

The central issue here is to identify the logical foundation that justifies the optimality of a rule producing non-optimal outcomes. Following Paretian logic, a community’s members regard a rule as “optimal” when it is not possible to change it without unanimous agreement among members. This raises the question of how it may be possible that members of a group agree to unanimously adopt a rule that yields non-optimal result - i.e., terminal allocations with respect to which some members of the group are losers and others are winners. The key point is that while, from the traditional output-oriented perspective, the presence of losers coincides with a departure from PO, from the process-efficiency aggregate perspective, the presence of losers does not preclude PO. The reason is that once (i) the costs of producing the law and (ii) the heterogeneity of the prospective cases to be regulated are incorporated into the efficiency calculus that underlies the choice of the rule, group members situated symmetrically at the constitutional choice stage might be willing, under certain circumstances, to accept departure from PO at the outcome level.

*Case 3: the optimal rule generates non-optimal results in a large number of cases; this is justified because the production costs associated with alternative rules that would generate optimal terminal results in those cases exceeds the costs associated with the departure from optimality of the allocative terminal results generated by the existing rule.*

Let me now relax the assumption of constant $c^0$. As clarified previously, the second condition required for social welfare maximization is (3) – i.e., the minimization of the total cost of producing $r$. Now, since (2) and (3) operate independently, the maximization problem indicated in (1) leads to a result that is obviously different from the result obtained in (2). It could be possible that the rule satisfying (3) does not satisfy (2). This might be due to the fact that the lawmaking institution producing the rule
satisfying (2) does not coincide with the one satisfying (3). This simply signifies that once the costs of producing legal rules are incorporated into the social welfare function the social-welfare maximizing legal rule does not necessarily coincide with the rule the maximize the surplus associated with final allocative results.

To explain, assume two rules $a$ and $b$. Let me denote the costs of producing $a$ and $b$ as, respectively, $C_a^0$ and $C_b^0$, and the aggregate surplus associated with $a$ and $b$ as, respectively, $B_a - G_a$ and $B_b - G_b$. Let me assume that rule $a$ is the cost-minimizing rule, while $b$ is the rule that maximizes the surplus associated with terminal allocative results. This implies that $C_a < C_b$ and $(B_b - G_b) > (B_a - G_a)$. Now, if $C_b - C_a > (B_b - G_b) - (B_a - G_a)$, then $a$ is the social welfare maximizing rule. Individuals situated symmetrically in the constitutional choice stage will unanimously chose $a$, despite the fact that $a$ does not maximise the surplus associated with final allocative results.

The preceding consideration suggests that the second condition for welfare maximization/optimality (i.e., law-making-costs minimization) is not incompatible with inefficient allocative terminal results. If the law-making cost differential between the cost-minimizing rules and the alternatives rules exceeds the aggregate net benefit differential between them, then the cost-minimizing rule is preferred over the available alternatives. In particular, if the law-making-cost differential is significant, the cost-minimizing rule might be regarded as the optimal rule despite the fact it produces non-optimal results in a large number (or, put to the extreme, the majority or even all) of cases to be regulated. The law-making process capable of maximizing the aggregate net-benefit is excessively costly for the community. Therefore, rather than incurring these high law-making costs, the group members prefer a cheaper rule that generates allocative inefficient terminal results.

An example illustrating this latter scenario is provided by the rules enforced by the business community to preserve competition among different firms, such as rules
prohibiting price agreements, quota agreements, or mergers among competing firms.\textsuperscript{168} These kinds of agreements are mutually beneficial to the participant firms. Thus, for these industries, the rule prohibiting such business practices is clearly non-optimal. The traditional justification is that these practices produce external effects in the form of costs to consumers. However, Buchanan points out that the standard argument based on the presence of external effects is not decisive: in many instances, agreements producing external effects on third parties are not prohibited. The decisive element is not the presence of external effects, but "asymmetry in the costs of organizing different voluntary agreements"\textsuperscript{169} (i.e., transaction costs). For the rule prohibiting some mergers, if transaction costs were equal to zero, consumers could establish a cooperative, thereby allowing them to exploit economies of scale and avoiding monopolistic extra profits. In other words, private bargaining among consumers would enable them to attain allocative efficiency, and the enactment of a rule prohibiting anticompetitive business practices would be unnecessary.

Thus, it is not the presence of external effects that justifies the rule prohibiting some anticompetitive agreements; rather, it is the circumstance that \textit{the costs of organizing voluntary arrangements are not equal for all possible affected parties}. Consider both the aggregate surplus associated with the rule prohibiting anticompetitive agreements and the absence of this rule. In the first scenario (enacted rule prohibiting mergers among firms), the social surplus is given by the aggregate benefits that consumers derive from the rule prohibiting voluntary agreements, minus the costs imposed on the firms subject to the prohibition of anticompetitive agreements, minus the costs of producing this rule. In the second scenario (no rule prohibiting mergers among firms), the social surplus is given by the aggregate benefits to the firms, minus the cost to consumers of monopolistic exploitation, minus the costs of organizing cooperative arrangements among consumers. The optimality of the rule that prohibits mergers depends largely on the relative costs of reaching an agreement regulating economic activity. Since the number of consumers is large, it is likely that the magnitude of the relative costs required to reach an agreement

\textsuperscript{168}This example is discussed in Buchanan, “The Relevance” see \textit{supra} note 142 at 348–350.
\textsuperscript{169}Ibid. at 349
among consumers is so large that it outweighs the benefits derived by consumers from a hypothetical cooperative solution. Therefore, in this case, the benefit to consumers would be maximized by a rule that prohibits mergers. That is, the externality problem is better solved through a ROG that relies on a centralized mechanism for the production of a rule prohibiting certain business practices, even if the resulting final allocative results violate the Pareto principle by making some parties worse off in a number of instances. The important conclusion is that the decision about the optimal ROG is not related only to the structure of costs and benefits of the final allocative outcomes; rather, the relative costs of producing alternative rules play a crucial role in the decision about the most efficient legal arrangement.

Finally, the discrepancy between unanimity at the constitutional stage of choice and allocative inefficient results is embedded in the functioning of all institutions for collective decision-making. Buchanan emphasizes the example of the constitutional rule prescribing that group decisions should be made by a determined representative proportion of the entire group (e.g., majority rule). When decisions at the institutional level are not based on a unanimous consensus but are taken by a smaller group, it is likely that the results will be non-optimal in many occasions. Only unanimity ensures that the outcomes of a collective decision-making process are Pareto-optimal. However, as in cases 2 and 3, when the costs of reaching agreement (i.e., costs of producing the rule) are included in the cost-benefit analysis, members of the relevant group may prefer a rule that leads to results which are non-optimal in many cases, rather than bearing the costs of a rule-producing mechanism that generates Pareto-optimal outcomes. Since it is expensive to reach unanimity in collective decisions, members of the social group may accept a less expensive majoritarian mechanism, even if it produces non-optimal results.

2.1.3. Conclusions

The lesson to draw from the foregoing discussion is that once Paretian logic is extended to the constitutional choice stage, then the Pareto Optimality of the rule does not ensure the Pareto Optimality of final allocative results. Once the entire set of costs for producing

---

170 Ibid. at 351–353.
the rules governing an expected sequence of cases is considered, the closest attainment to the Pareto frontier might be reached at a point that cannot be qualified as Pareto-optimal in the standard sense. I synthesize the line of reasoning through the following summarizing propositions:

(i) The ROGs determine the shape and dimensions of the Pareto Region. (ii) A Pareto efficiency analysis should include in its domain the gains of efficiency that could be attained through changes in the ROGs. (iii) A constitutional constraint requiring unanimity in the choice of the ROGs should be assumed to ensure that the efficiency analysis is meaningful. Under this constitutional principle, optimal rules are those that cannot be modified without the consent of all persons involved in the economic activity. (iv) Optimal rules may generate final allocative results that are non-optimal according to the standard outcome-oriented notion of Pareto efficiency but are efficient (i.e., social welfare maximizing) when applied to an indefinite number of cases.

2.2. Process Efficiency Explained through Individual Constitutional Choice Model

This section clarifies the nature of process choice through a simple formalization of individual constitutional choices. The formalization aims at emphasizing the limitations of the conventional optimization approach in capturing the nature of the trade-off associated with process choices. Subsections 2.2.1 and 2.2.2 focus on the cost and benefit function of an individual in a hypothetical constitutional setting. Thus, I use the term “efficiency” in relation to individual (not social) outcomes. Obviously, once the individual enters into the social contract his or her costs and benefits represent part of the social costs and benefits associated the production of legal rules. For simplicity, it is useful to begin the analysis by considering the problem of process efficiency in relation to the single individual; the nature of the problem the individual faces (i.e., the choice of the welfare maximizing lawmaking process) is in many respects similar to the problem that an hypothetical constitutional convention would confront.
Consider an elementary model of constitutional choice in which an individual decides whether to enter into a social contract with other individuals basing on a utility-maximizing calculus. By entering into the social contract, the person agrees to be subject to legal order and to obey the law, thereby foregoing the possibility of taking any action prohibited by the law; at the same time, other individuals join the social contract and become subject to the same obligations. In this manner, each individual saves the costs that otherwise would have been incurred to protect his or her property and personal security in a world without legal order. Alternatively, rejecting the social contract would result in a Hobbesian state of nature, in which no legal rules exist, and where violence may deprive individuals of their entire endowments and personal security.

The law is a set of legal rules that I denote as $R$. The individual utility function can be described as follows:

$$U_i(B_i^R, G_i^R, C_i^R) = (B_i^R - G_i^R) - C_i^R$$

where $B_i^R$ denotes the benefit that an individual receives by obeying rules $R$. It may consist in the reduction of negative externalities suffered in the Hobbesian state of nature, and imposed by other people’s behaviour. It may also include the benefits deriving from collective efforts that would otherwise not be achieved in the absence of legal rules that ensure the division of the cooperative surplus from collective actions. In brief, $B_i^R$ represents either the benefit from the reduction of negative externalities in the state of nature or the provision of positive externalities from collective action made possible by the legal order. To reduce the costs of living in the Hobbesian state of nature and benefit

---

from the advantages of collective action, the individual sacrifices his or her freedom to perform some activities in a system without legal rules. The value to the individual of the prohibited activities represents the opportunity-cost of obeying legal rules (i.e. the opportunity cost of entering into the social contract). In (1), I denote this opportunity cost as $G_i^R$. Since alternative contents of legal rules affect the cost and benefit functions of the individual, it is reasonable to assume that $B_i^R$ and $G_i^R$ are functions of $R$.\(^{173}\)

Let me set aside for a moment $C_i^R$. The individual will support the existence of the legal order only if s/he has the expectation that the benefit from being subjected to the legal rules that will be enacted, $R$, exceeds its opportunity cost, so that $B_i^R - G_i^R > 0$, which defines the condition of incentive compatibility for the acceptance of the constitutional contract. $B_i^R - G_i^R = S_i^R$ denotes the individual surplus associated with $R$. Hence, the incentive compatibility can be denoted as $S_i^R > 0$.

Having agreed to enter into the social contract, the individual supports the rules that maximize $S_i^R$. From the individual’s perspective, the maximization of $S_i^R$ corresponds to what I have previously defined as “substantive efficiency”, which I denote as $E_i^r$.\(^{174}\)

Now, the individual taking part in the collective decision-making process will support the optimal rule $\beta$ that maximizes $E_i^r$, where $\beta \in R$. Since $B_i^R$ and $G_i^R$ are functions of $R$, then $\beta$ satisfies the following condition:

---

\(^{172}\) Following one of the tenets of economic analysis of law that characterizes legal sanctions as the price for undertaking certain courses of action, one might say that the individual accepts to pay the price of being subject to the legal sanctions in case s/he decides to engage in activities prohibited by the law. In this sense, the individual accepts to incur increasing costs of undertaking certain courses of action.

\(^{173}\) For simplicity, I assume $U_i$ to be a linear utility function.

\(^{174}\) Notice that the substantive efficiency in this context corresponds to the maximization of individual surplus. However, once we drop the restrictive assumptions of this simplified model of constitutional choice, it must be recognized that the legal rules that are substantive-efficient are those that maximize the “social” surplus.
Identifying and producing legal rules is itself a costly collective activity that can be organized in many different ways and through alternative institutional arrangements, i.e. the costs associated with the creation of surplus-enhancing legal rules depends on the features of the law-making process. For the sake of clarity, I group these production costs into the sweeping category of decision-making costs.\footnote{In this section, I use the expression “decision-making costs” in a different and more generic way than in the remainder of the study. Here decision-making costs are the entire costs of taking part in the collective decision-making process. Subsequently, the expression will refer specifically to the costs associated with the aggregation of social preferences.} I denote the individual portion of these decision-making costs as $C_i^R$ and the set of alternative institutional law-making processes as $P$. Under these assumptions, to maximize $1$, the individual will support the cost-minimizing decision-making process $\rho \in P$ so that

$$\min_{\rho} C_i^R = C_i^R(\rho).$$

(3)

Here, I confront a crucial theoretical point.\footnote{To be precise, following the conventional optimization logic, it should be said that the individual will support the process that minimizes the decision-making costs for each given output level. As I explain subsequently in the text, it is difficult to apply this logic in the context of the production of legal rules, where the qualitative element prevails over the quantitative dimension of production.} Apparently, the individual faces a common optimization problem split into two pieces: minimization of the decision-making process cost (2), and maximization of substantive efficiency (3). Notice also that while process choice is related to the minimization of the decision-making cost, output choice is related to the maximization of substantive efficiency.
I contend that this representation is too simplistic. If we consider the functioning of the many different rule-producing mechanisms in the real world, we recognize that the choice of decision-making process (hereinafter “process choice”) cannot be entirely reduced to $\min_{\rho} C_i^R$. When analyzing process choices, there is a second important issue to consider: the choice of decision-making process influences the way in which the problem of maximizing substantive efficiency ($\max_{\beta} E_i^S$) is formulated, calculated, and resolved; therefore, the choice of decision-making process predetermines the qualities of the process output.

To explain how process choice relates to substantive efficiency, let us assume that the individual entering into the social contract knows $\beta$. That is, he or she knows the content of the substantive rule that maximizes his or her individual surplus in a given anticipated situation. The individual cannot directly choose $\beta$ for the simple reason that $\beta$ is a “legal” rule, a binding code of conduct for the entire community, and as such it entails collective decision-making. Put differently, the choice of legal rule presupposes a previous choice of the collective decision-making process through which legal rules are created and enacted. For this reason, the individual joins other members of the community in a constitutional arena to make collective decisions as to how to create the legal rule governing a case. The constitutional arena addresses the difficult task of deciding who decides the rule governing the case. In this context, the individual will

---

177 In general, I assume that the lawmaker does not know the social-welfare-maximizing rule; however, at this preliminary stage of analysis, the hypothetical scenario in which the individual knows the surplus-maximizing rule is analytically useful because that knowledge makes it easier to observe the welfare effects of the process choice.
support the choice of the decision-making process that conduces to legal outcomes that are as close as possible to satisfying (2) and (3). For example: according to changing circumstances, the individual might support the election of a collective representative body that is in charge of law-making decisions; he or she might support the institution of an impartial authority for solving disputes; or if circumstances allow for bargaining activity at low cost, he or she might call for a private bargaining solution.

The hypothetical scenario described above helps explain the causal relationship between process choice and substantive efficiency. When the demand for law arises, the demand for a law-making process is necessarily entailed. The demand for the law-making process raises the problem of “choosing the chooser”, which inevitably precedes the choice of the rule. Crucially, from the individual’s standpoint, the “choice of the chooser” impacts the “choice of the rule”; that is, the outcome resulting of the law-making process inevitably differs from the ideal surplus-maximizing rule.\(^{178}\) Let \( r \in R \) denote a real outcome produced by the law-making process \( \rho \), so that \( r = f(\rho) \). Once \( \beta \) is identified, the individual seeks to minimize the difference between the surplus gained through it, denoted as \( S^\beta_i \), and the surplus he or she anticipates will be associated with the \( r \), denoted as \( S^r_i \). This means that the individual’s problem of identifying the substantively efficient rule can be expressed as follows:

\[
\max E_i^r = \min \{ S^\beta_i (\beta) - S^r_i (r(\rho)) \} 
\]  

(4)

Crucially, the surplus associated with the ideal rule can be defined as follows:

\[
S^\beta_i(\beta) = B^\beta_i(\beta) - G^\beta_i(\beta)
\]  

(5)

\(^{178}\) The reason is that collective choices underlying the production of law are events in which control is distributed among many individuals; therefore, one individual does not have the control over the outcome. The idea that the collective production of legal rules generates outcomes that do not correspond to the individual’s first best is largely recognized in the literature and will be clearer as the discussion proceeds. See, e.g. Gordon Tullock, Private Wants, Public Means: an Economic Analysis of the Desirable Scope of Government (New York: Basic Books, 1970) at 97–99; Buchanan, “The Relevance” see supra note 142 at 348–352.
The law-making process will not generate this surplus; however, it represents an efficiency parameter against which the individual assesses the outcomes that are achievable by the available law-making processes. The surplus that the individual actually obtains is equal to the benefit the individual expects to receive from the real outcome minus the cost associated with obeying it minus the decision-making costs associated with the law-making process:

\[ S_i^r = B_i^r(r) - G_i^r(r) - C_i^r(p) \]  

(6)

Since \( r = f(p) \), \( S_i^R \) can be defined as follows:

\[ S_i^r = f(r(p)) = B_i^r(r(p)) - G_i^r(r(p)) - C_i^r(p) \]  

(7)

If we substitute (7) into (5), we obtain the following maximization problem:

\[
\max E_i^s = \min_p \{ S_i^\beta(\beta) - S_i^r(r(p)) \} = \{ S_i^\beta(\beta) - [B_i^r(r(p)) - G_i^r(r(p)) - C_i^r(p)] \}
\]  

(8)

That is, the process-choice produces a twofold effect: (i) it directly affects the magnitude of \( C_i^r \) and (ii) it impacts \( S_i^r \). This is not a trivial conclusion. It means that *substantive efficiency does not reduce to the choice of the ideal outcome but entails the choice of the decision-making process*.\(^{179}\) In the next two subsections, I expand on this crucial point.

\(^{179}\) Another way of looking at it: when process choice is factored into the efficiency calculus, the structure of the individual maximization problem changes significantly from the identification of the ideally efficient rule to minimization of the frictions associated with the collective decision-making process.
2.2.1. Process Choice as Cost-minimization Problem

I will now compare two hypothetical descriptions of process choice an individual faces in the constitutional stage. The first describes process choice as a cost-minimization problem, whereas the second connects process choice to the law-making process outcome quality.

**Hypothesis 1**: for any given level of $U_i(r, \rho)$, the individual prefers the decision-making process that minimizes the magnitude of $C_i(\rho)$; after the cost-minimizing process has been identified, the individual prefers the rule $\beta$ that maximizes his or her expected net benefit.

The representation of process choice as a cost-minimization problem is based on the logic of the neoclassical theory of production, which separates production behaviour analysis of a firm into two steps: cost-minimization and profit-maximization. First, the producer chooses the technology that allows the production of a given quantity at the minimum cost, given input prices; second, the firm chooses the quantity to be produced to maximize the difference between costs (at the minimum level derived in the first step) and revenues. If we import the same logic to the constitutional stage of choice, the problem confronted by the individual can be split in two separate steps: identification of the cost-minimizing decision-making process (input), and identification of the surplus-maximizing rule (output choice). That is, process choice is conceptualized as a cost-minimizing problem independent of output maximization. Here, I contend that the application of the cost-minimization logic to process choice (proposition 1) is based on three implicit assumptions that do not adequately capture the nature of the law-making process.

---

180 Recall: the individual wants to maximize $W_i = (B_i - G_i) - C_i$ assuming that he or she already knows the rule $R$ that maximizes $(B_i - G_i)$.

Let me first identify the assumptions underpinning the neoclassical theory of the firm and then assess their adequacy for describing properly the choice of law-making process. There are three implicit assumptions adopted by the neoclassical theory of the firm: (i) separation between input and output choices; (ii) separation between substantive and process efficiency; and (iii) substitutability of input-output bundles.

(i) Separation of Input and Output Choices

According to the neoclassical theory of the firm, the cost-minimization problem is reduced to a choice between alternative input bundles, while the profit-maximization problem entails identification of the output quantity to be produced. A graphical representation of the production and cost functions provided by the neoclassical two-input model\(^1\) (Figure 6) helps clarify the cost-minimization logic according to the neoclassical model and its mechanism involving the separation of process choice from output choice. The isoquants provide a graphical representation of the production function, which summarizes the technological constraint determined by the existing technology. The isocosts represent the cost-function, which summarizes the economic constraint determined by current input prices for any given quantity. The point of tangency of the lowest isocosts with the isoquants represents the cost-minimizing production-technique. This graphical representation depicts the cheapest possible production technique (i.e. located on the isocost closest to the axis origin) among the set

---

\(^1\) A discussion of firm theory organized around the clear distinction between one-input and two-input models is provided in Nechyba, “Microeconomics” supra note 181.

In the one-input model, the production-function is represented on a Cartesian plane: the input level is represented on the x-axis and the profit-level on the y-axis. Each point on the production curve represents the maximum quantity that can be produced for any given amount of input employed. The production function represents the technological efficiency of production because input prices are taken as given; additionally, the output level is a function of the selected input level.

Conversely, the two-input model represents the cost-function on a Cartesian plane: the first input is represented on the x-axis and the second input on the y-axis. This model presents two characteristics. First, the input choice is analytically separated from the output choice, so that the economic problem confronted by the firm is divided in two parts: the choice of the cost-minimizing input-bundle and the choice of the profit-maximizing output-level. Second, in this model, the production function (represented by the line formed by the points of tangency between the isoquants and the isocosts) summarizes both technological efficiency and allocative efficiency.
of technologically efficient production processes (i.e. isoquants) for each level of output quantity.\(^{183}\)

**Figure 6 Cost-minimization Problem in Neoclassical Theory of the Firm\(^{184}\)**

In this model, the cost-minimization problem amounts to an input-choice problem (i.e. the cheapest production technique represented by the cheapest input-bundle), while the profit-maximization problem amounts to an output-choice problem (i.e. the number of output units to produce).

---

\(^{183}\) To be clear, a production technique is technologically efficient when it does not waste resources (i.e. producing the maximum output with the lowest level of input). Technological efficiency does not take into account input prices. By contrast, a production technique is economically efficient when it is the cheapest way of producing a good. That is to say, input prices are taken into account to choose from among many technologically efficient production techniques. For each output level, isoquants and isocosts represent respectively the technologically and economically efficient input bundles. Once the cost-minimizing production technique has been identified, the producer chooses the profit-maximizing output level corresponding to the point at which the marginal revenue of an additional output unit equals its marginal cost.

\(^{184}\) I borrow Figure 6 from the following source: Professor Martin J Osbourne, *Tutorial on Theory of the Firm and Industry*, posted on the website: www.economics.utoronto.ca/osborne/2x3/tutorial/index.html
(ii) Separation between Substantive and Process Efficiency

Following the neoclassical economizing logic, the constitutional choice of the individual could be described as a two-step problem with a cost-minimizing piece (i.e. the choice of most cost-effective decision-making process) and a surplus-maximizing piece (i.e. the choice of the surplus-maximizing legal rule). Thus, the whole utility-maximizing problem of the individual at the constitutional choice stage can be described in terms of the standard marginal efficiency analysis. First the individual identifies the decision making process producing $R$ at the minimum cost, and then he or she identifies the surplus-maximizing rule. In essence, the neoclassical line of reasoning reduces process-efficiency to a cost-minimization problem, while output-efficiency constitutes a distinct surplus-maximization problem.

(iii) Substitutability of Input-Bundles

The third implicit assumption of the neoclassical theory of production is the perfect substitutability of input-bundles. According to this idea, it is possible to identify sets of input bundles that are equivalent with respect to a) the level of output produced (i.e. isoquants)\(^{185}\) and b) impact on production-costs (i.e. isocosts). Within these sets of alternative production techniques, the marginal revenue on a unit of output does not change, regardless of the combination of inputs the producer chooses to employ. Different input bundles can be substituted in place of each other without consequences for the quality and profitability of the output. The technical and economic consequences of the input choice occur only when the individual chooses among different sets, but not from within sets.

\(^{185}\) Many textbooks define isoquants as the set of input bundles that result in the production of the same amount of output. See, for example, the textbooks cited supra note 181. Usually, the variable considered for grouping the input bundles is the “quantity” produced. Quality does not enter the picture. Thus, the output quality is considered exogenous to the model. As I will clarify shortly, this is one of the most important shortcomings of applying the neoclassical theory of the firm to the production of law. When it comes to legal rules, quantity is not the most important element to be considered. Quality of the output cannot be disentangled from the quantity. Since assumptions of theoretical models should capture the most relevant characteristics of the analyzed situation, a model of the law-making process should endogenize the quality of legal rules.
The technical terminology itself is eloquent. Isoquants and isocosts are “indifference” curves: the combination of inputs representing the same level of output and costs, respectively. Each point on an isoquant represents a combination of inputs that produces (in a technologically efficient way) the same level of output, so that the producer is indifferent between alternative combinations of inputs lying on the same isoquant. Holding the output-prices constant, whichever point the producer chooses on the isoquant, he or she will obtain the same revenue. Analogously, each point on an isocost curve represents a combination of inputs associated with a constant level of production cost. The producer is, therefore, “indifferent” to the two different combinations of inputs that lie on the same line.

The central point is that these indifference curves cannot be assumed properly in the context of the law-making process. When it comes to the process through which law is formed, alternative production techniques cannot be assumed to produce the same level of output or to entail the same level of costs. When we move from the environment of the neoclassical theory of the firm to the context of production of law, the assumptions of the substitutability of the inputs and of the insulation of output from the production process features must be abandoned. Different production techniques cannot generate the same type of legal output nor can they be associated with the same level of cost or the same impact on social welfare. Concepts such as “isocost”, “isoquant”, or “isoprofit” are inapplicable to the context of the law-making process because they preclude explaining how the quality of outcome results from the law-making process features. Put bluntly, *adoption of the indifference curve as an explanatory tool of the process choice precludes explanation of the structural mechanisms generating the quality of the legal rules.* Once it is assumed that an individual maximizes his or her own utility through the choice of decision-making process, it would be wrong to proceed by ordering alternative law-making techniques on different level curves, representing indifferent law-making

---

186 Usually the expression “indifference curves” is used in the context of consumer theory to indicate the level curves representing the set of consumption bundles yielding the same level of utility to the consumer. However, indifference curves are not used only in the context of consumer theory; indifference curves are “level curves” such as isocosts, isoquants, and isoprofit used for representing a three-variable function in a two-dimensional graph, wherein the variable on the vertical axis is assumed constant. For this reason, I consider the use of the term “indifference curves” with respect to the theory of the firm to be correct.
techniques corresponding to the given levels of individual utility. Assuming the individual to be indifferent to alternative law-making processes generating the same level of utility would obscure the crucial fact the law-making processes are characterized by alternative efficiency advantages over a changing set of circumstances. The very nature of process choice is choice among institutions that yield efficiency advantages under different circumstances.

In conclusion, if the neoclassical description of the process choice, based on the above illustrated implicit assumptions, were adopted as a comprehensive model for economic analysis of the law-making process, the central claim of my entire study could be reduced to a cost-minimization problem. In practical terms, my claim could be summarized by the idea of enriching the standard legal cost–benefit analysis so that it includes law-making costs into the legal economic calculus. The reasoning could be developed as follows: if the cost–benefit analysis of legal rules does not take into account the decision-making process, it might result in a suggestion for adopting legal rules that, despite their optimizing impact on the surplus of the individual involved, will not be produced because would be too expensive to produce. That is, if $B_i^R - G_i^R < C_i$, the individual does not find it convenient to enter into the constitutional contract and will seek other, more cost-effective ways to reduce the cost of living in the Hobbesian state of nature. From this perspective, the process choice is understood exclusively as a problem of decision-making cost minimization; it is not related to the output quality.

I do not deny that there is some truth to the preceding description of process choice. However, it contains a serious limitation. Assuming separation between input and output choices misses an important point, i.e. *technological separation between inputs and outputs postulated by the neoclassical theory does not fit in the complexity of the production of law*. Since the three assumptions identified above are closely related, if assumption (i) is dropped, assumptions (ii) and (iii) will drop as well. In other terms, the neoclassical model of production neglects the issue of output quality and its relationship with the choice of production process. In this model, output quality simply does not play
any role in the choice between alternative input bundles. Input bundles having the same cost and producing the same output are perfect substitutes.

To summarize: hypothesis 1 does not consider that in terms of the production of law, the input–output relationship is much more complex than that postulated by the neoclassical theory of the firm. The next subsection demonstrates this point by providing a more adequate description of process choice.

2.2.2. Process Choice as a Determinant of Outcome Efficiency

Let me relax the assumption that the individual knows the ideal surplus-maximizing rule $\beta$, and introduce the complicating factor that alternative decision-making processes inevitably lead to legal rules that differ in content. More precisely, in contrast with previous section, I assume (i) that the individual wants to maximize the surplus but does not know the content of $\beta$; and (ii) that input bundles are not technologically and economically substitutable, so the individual is not indifferent towards alternative decision-making processes.

Based on these assumptions, let me elaborate an alternative description of the constitutional choice.

_Hypothesis 2:_ for any given level of $U_i$, the individual prefers the decision-making process $\rho$ that is most likely to produce a legal rule $r$ that is the closest possible to $\beta$ and minimizes $C'_i$. Both $r$ and $C'_i$ are functions of $\rho$.

Unlike hypothesis 1, this description depicts a more complex choice, with three distinctive features: (i$'$) the minimization of costs is just one of the two pieces of process choice; (ii$'$) the inclusion of the law-making process implies an inevitable “imperfection” of the legal outcome so that $S'_i \succ S'_i \beta$; and (iii$'$) utility maximization is attained through minimization of the imperfection in the real outcome (as summarized by (4)).
The comparison with the neoclassical description of the firm is useful once again. The neoclassical firm pursues profit maximization by taking two decisions: an output decision and an input decision. By contrast, in the context of the constitutional choice, the utility-maximizing individual cannot make either the output choice or the input choice. He or she merely understands that in order to maximize his or her utility, he or she must choose a decision-making process that is possibly capable of producing a legal rule yielding the maximum surplus at the minimum cost. As I have lastly assumed, the individual does not know the rule having these properties. The only thing the individual knows is the choice set of alternative decision-making processes.

The individual knows the alternative decision-making processes available and wants to choose one that maximizes his or her utility functions, but he or she ignores the nature of the efficient output (i.e. the efficient rule). Put differently, the individual has an objective function—he or she knows the independent variable to optimize (i.e., individual utility)—but he or she does not know the input–output relationship that constitutes the object of his or her optimizing choices. In such a situation, the can observe the characteristics of the available decision-making processes and identify the processes that are likely to lead to a rule that maximizes his or her social surplus. To do this, he or she needs to know how the process exerts influence on the output decision and, indirectly, on his or her surplus. Only after these characteristics are identified is the individual in a position to make a rational decision.

187 Recall, from our previous discussion, that the concept of efficiency is based on the means-ends separation and that the efficiency calculus requires an objective relationship between output and input levels. Along these lines, the optimization calculus requires a knowledgeable input-output relationship. In the context of constitutional choice, the efficiency calculus requires that the individual knows which rule maximizes his or her own utility (utility-maximizing output), and what is the level of input that produces that rule at the least cost (cost-minimizing input bundle). However, since the individual ignores the content of the utility-maximizing legal rule, he or she cannot identify the decision-making process leading to the utility-maximizing rule. In this sense, the individual cannot optimize the input-output trade-off so as to maximize his or her own utility. The optimization problem of the individual takes a different form here. The individual can only identify those features of the decision-making process that affect the characteristics of the legal rule in the most favourable direction.
In conclusion, at the constitutional stage of choice, the individual cannot optimize any input–output relationship for the simple reason that a well-defined input–output relationship is not available. He or she ignores the efficient outcome; he or she has knowledge only of the available alternative processes. As the analysis proceeds, it become clear that rather than selecting the efficient rule, the individual optimizes the tradeoffs among the alternative process values that are instrumental to output efficiency maximization.

2.2.3. Process Efficiency and Social Welfare

This subsection discusses process choice as a collective choice, i.e. a choice jointly made by individuals who have decided to participate in the production of legal rules.

Let me assume that the objective function of the community is to maximize social utility (or welfare), intended as the sum of the utility functions of \( n \) individuals.\(^{188}\) The advantage of adopting such a definition, in which individuals’ utility is a direct argument in the social welfare function, is that one can read and interpret it through the lens of the previous discussion on individual utility.\(^{189}\)

The problem confronted by the community at the constitutional choice stage can be described as follows: how best to construct a set of institutional arrangements capable of producing legal rules that are “substantively efficient” at the “minimum cost”. In algebraic terms:

---

\(^{188}\) See *supra* note 144

\(^{189}\) Previously, I have defined the utility function of the individual entering into the social contract as maximization of the expected utility, as obtained by acceptance of the social contract. From this definition, I derive the notion of a social utility function. At the moment, for my analysis, it is sufficient to adopt a thin definition of social welfare, i.e. the sum of individual utilities defined previously. Here, I am not interested in elaborating a more sophisticated social welfare function, because my analytical focus is on the process-output relationships, rather than on the welfare-maximizing criteria. Moreover, as I will demonstrate subsequently, the structure of the process-output relationship in my model does not vary with the notion of social welfare.
Similar to the “individual” constitutional choice, the “collective” process choice cannot be reduced to a mere cost-minimization problem. The choice of process does not concern merely the magnitude of the decision-making cost but also influences the amount of social surplus.

Since the law-making process is a collective activity involving a number of individuals with different utility functions, it entails the aggregation of the community members’ contrasting preferences (i.e., social choice problem). As assumed previously, the outcome of the process of aggregating individual preferences is inevitably different from the ideal outcome desired by each individual; so, if \( r \neq \beta \) then:

\[
\sum_{i=1}^{n} S'_i(r(p)) \neq \sum_{i=1}^{n} S'_i(\beta) \tag{10}
\]

This means that the social surplus ensured by any law-making process is unavoidably different from the sum of the individual surplus that each individual would gain under \( \beta \).

Finally, notice that \( n \) is an endogenous variable. This is because the structure of compliance costs is a function of the efficiency features of the legal rules, which, in turn, are a function of the structure of the law-making process. It follows that process choice is related to compliance levels. This point will be discussed at length in chapter 3.

\[190\] As I will clarify subsequently, \( r \neq R \) is due not only to the social-choice problem but also to the vast array of imperfections internal to the law-making process.

\[191\] This point is very important and will be explained in detail in the next section—I will use the Edgeworth Box model to show that optimal rules might be generated in many instances of non-optimal allocative results.
2.2.4. Process Efficiency versus Output Efficiency Analysis

The foregoing analysis facilitates a summary of the differences between the traditional outcome-oriented approach and my proposed process-oriented approach. The conventional output-oriented approach focuses on substantive efficiency, understood in terms of the efficient allocation of legal entitlements by means of surplus-maximizing legal rules.\textsuperscript{192} Its chief concern is the identification of legal rules that enable individuals to maximize the surplus they obtain by obeying the law. In contrast, the structural approach looks at substantive efficiency as a function of the law-making process. It does not eschew legal rules by any means, but sees them as a function of their formation process. This analytical perspective allows for the inclusion of three elements, which are usually overlooked by the conventional approach, into the legal economic analysis: (i) the magnitude of law-making cost, (ii) impact of the law-making process on the efficiency properties of the outcome, and (iii) impact of the law-making process on individuals’ decisions to comply with the law. Let me briefly clarify these three points that will be analysed in depth throughout the entire discussion in this study.

(i) The structural approach incorporates the law-making costs, which are traditionally overlooked by the outcome-oriented approach, into the efficiency calculus.\textsuperscript{193} The next section of this chapter develops a uniform taxonomy of law-making costs that enables us to systematically factor in the law-making cost functions into the efficiency calculus. Subsequently, part II and part III of this study will examine the institutional and structural variables influencing these law-making costs.

(ii) The structural approach inquires into how surplus maximization is related to the characteristics of the law-making process.\textsuperscript{194} More precisely, it enables us to explain how the imperfections of the law-making process cause legal rules to depart from the

\textsuperscript{192} The analytical focus of the output-oriented approach is well summarized as the maximization problem described in (9) in section 2.2.3.

\textsuperscript{193} This is reflected in (9), which includes the minimization of $C_i$ in the maximization problem.

\textsuperscript{194} This is reflected in (9), which describes economic surplus as a function of legal rules, which are, in turn, a function of the law-making process.
allocative efficiency parameter. As the discussion proceeds it will become clearer that the institutional features of law-making mechanisms significantly influence the efficiency properties of the outcome.

(iii) Conventional law and economics is predominantly based on the behavioural assumption that people obey the law. However, empirical observation suggests that legal compliance is, at least in part, induced institutionally. If law-making processes generate legal rules that are too difficult to be appreciated by citizens or that are too costly to comply with, then citizens have strong incentives to try circumventing the law. The structural approach recognizes these problems by assuming that individuals decide to comply with the law based on a utility-maximizing calculus. Under this assumption, the features of the law-making process influence the levels of legal compliance.\textsuperscript{195}

2.3. Definition of Process Efficiency

Section 2.1 clarifies that once a process choice is included in the efficiency calculus, Paretian logic leads to the identification of optimal legal rules that may not generate allocative-efficient results in a given number of cases. Section 2.2 explains that the relevance of process choice is not limited to the magnitude of the law-making costs; rather, the choice of law-making process significantly influences the efficiency properties of the outcome. The discussion in both sections demonstrates that once Paretian logic is applied at the constitutional stage choice, the problem of allocative efficiency of legal rules becomes one of minimization of the differential between the aggregate surplus that would be associated with the sum of Pareto-optimal terminal allocative results and the aggregate surplus associated with the optimal legal rule agreed upon by all members of the group. Moreover, I have explained that misalignment between optimal legal rules and optimal allocative results is due to the following two reasons. First, once Paretian logic is applied at the constitutional stage choice, allocative efficiency is assessed from an

\textsuperscript{195}This is reflected in (9), which defines social welfare as the sum of the surplus of n individuals. It is assumed that n is a function of the process. That is, the structure of the law-making process influences social welfare through its effects on individuals’ decisions to comply with the law.
aggregate perspective that necessarily accounts for the variance of the situations in which the rule is applied. Second, once law-making costs are factored into the efficiency calculus, there arises the possibility of a misalignment between cost-minimizing and surplus-maximizing law-making processes.

The preceding discussion has left the nature and cause of the law-making costs unexplained and has not provided any indication of those features of the law-making process that have the greatest effect on the efficiency properties of the outcome. Moreover, the discussion has not provided any operational notion of process efficiency based on which alternative law-making processes can be compared. Notice also that so far, the discussion has implicitly assumed a world with (i) perfect information, (ii) no agency problems in the relationships between lawmakers and beneficiaries of legal rules, (iii) no problems of intransitive collective preferences, and (iv) no problems of transitions from inefficient to efficient legal regimes.

In this section, I turn to the construction of a methodology for conducting process efficiency analysis that allows for the inclusion of the imperfections and limitations associated with the production of law into the efficiency calculus. I provide (i) an operational notion of process efficiency and (ii) a uniform taxonomy of law-making costs.

Under my formulation of the approach, a law-making process is efficient when it satisfies the following four conditions:

---

196 In identifying these four efficiency criteria, I build on Francesco Parisi, “Sources of Law and the Institutional Design of Law-making” (2001) 19 J. Pub. Fin. and Pub. Choice 2-3 at 95-122, George Mason Law and Economics Research Paper [Parisi, “Sources”]. Parisi identifies three criteria for the comparative analysis of alternative law-making processes: (a) minimization of agency problems; (b) minimization of rulemaking costs; (c) stability and transitivity of collective outcomes. My proposed definition of process efficiency largely draws on the framework of Parisi, but at the same time distinguishes in some respects. First, I provide a different definition of law-making costs. Second, I place greater emphasis on the importance of “adaptive” efficiency. Third, I organize concepts along the lines of the supply and demand of law. Fourth, I adopt partially different terminology and different organization of concepts; for example, I attempt to incorporate the different criteria identified for comparative assessment of the sources of law under the same notion of “process efficiency”.
(i) It minimizes the *information* costs of the law-making process and maximizes the number of individuals who decide to acquire legal information and comply with the law (i.e. productive efficiency);

(ii) it is responsive to the intensity of individual *preferences* and aggregates them into a rational, stable collective outcome (i.e. social choice efficiency);

(iii) it is responsive to the *interests* of the people who are affected by the law (i.e. agency efficiency);

(iv) it produces legal rules that are *adaptive* to changes in the regulated environment (i.e. adaptive efficiency).

Hereinafter, I will refer to these characteristics as the “*procedural determinants*” of substantive efficiency. These are both *positive* criteria for comparative institutional analysis and *normative* criteria for institutional choice. Let me briefly discuss each procedural determinant of substantive efficiency and provide a law-making-cost taxonomy.

**2.3.1. Productive Efficiency**

The first procedural determinant of substantive efficiency is the ability of the law-making process to minimize the law-making costs and maximize legal compliance. This definition raises the problem of the measure of technical efficiency. As we know from chapter 1, we need to identify inputs and outputs carefully to identify efficiency. In particular, in this context, I need to define the input and output of the law-making process, as well as the related costs of producing the law. The following two subsections are devoted to resolving these two definitional problems.

**2.3.1.1. Input and Output of Law-making**

Legal rules are theoretical constructs, and their formulation requires great knowledge, competence, and experience. In economic terms, legal rules are *informational* goods in a twofold sense in that they a) contain information and b) require the generation of robust information flows. Based on this, I assume that the most important production factor of

---

the law-making process is *information*,\(^{198}\) and that information costs are the principal cost-item to be considered when measuring technical efficiency. As the comparative analysis will demonstrate, the characteristics of information flow through the law-making process can be regarded as the hallmark of distinction among alternative law-making institutions: actors involved in the law-making process deal with informational problems and bear considerable costs for acquiring the necessary information.

While it is straightforward that legal rules are the output of the law-making process, it is more problematic to establish a *definition* and a measure of the output for formulating the production function. It would be completely misleading to measure the output level in terms of the *quantity* of legal rules.\(^{199}\) The lawmaker’s goal is to maximize social-welfare not the quantity of legal rules. However, in this study I will assume collective ignorance about the social welfare function, as I will explain subsequently. Thus, I need a unit of measurement that is *indirectly* welfare-conducive, and, at the same, relatively easy to measure.\(^ {200}\) Based on these requirements, I define the law-making output in terms of *legal compliance*. That is to say, the lawmaker maximizes social welfare by maximizing the number of people who obey the law. Crucially, the maximization of legal compliance requires the lawmaker to be able to appreciate individuals’ preferences and to produce legal rules that meet the demand for legal rules. For analytical convenience, it is

\(^{198}\) See Kaplow Louis, “Rules Versus Standards: an Economic Analysis” (1992) 42 Duke L. J. 557 at 568: (“the problem of creating the law can be interpreted as one involving the government acquisition and dissemination of information”).

\(^{199}\) Actually, the excess in the quantity of legal rules has been, for a long-time, identified as one of the main causes of the decay of law in many contemporary legal systems. See Leoni, *Freedom, supra* note 197 (criticizing the phenomenon of over-legislation affecting modern democracies and, with them, the rule of law and the principle of individual freedom), and Giovanni Sartori, *Democratic Theory*, (Detroit: Wayne State University Press, 1962) at 306–313 [Sartori, “Democratic Theory”]. The problem of defining the production function and measuring the output of the law-making process is exacerbated in the bureaucratic process where bureaucrats are not subject to either electorate or market discipline. This latter point will be clarified in chapter 6.

\(^{200}\) This point will be made more clear in chapter 4, when I clarify the whole set of assumptions underlying productive efficiency analysis. For the moment, I just want to emphasize that in my model, the law-making process is not assumed to directly maximize the social welfare function. Although such purpose certainly orients the functioning of the law-making process, we cannot measure with certainty whether it may or may not be achieved because the social welfare function is not objectively knowable. It is because I assume collective ignorance of the social welfare function that I focus on the process features rather than the output. Subsequently, I discuss in detail these assumptions underlying my explanatory model of the law-making process.
convenient to break down the analysis of productive efficiency into two pieces: maximization of legal compliance and minimization of information costs.

2.3.1.2. Information Costs

Hayek has long clarified that the production of law entails dramatic information problems.\(^{201}\) In this subsection, I define a typology of information costs that is organized along the lines of the supply of and demand for law. On the supply side, the lawmaker bears information costs when providing the content of the legal rules and when assigning legal entitlements. I call these costs “production-information costs”. They include both information-gathering costs and measurement costs.\(^{202}\) Information-gathering costs are the costs associated with activity of gathering the information required to produce legal content. This includes information about the preferences of the people subject to the law and people whose behaviour is regulated by legal rules. Moreover, it includes information about the factual conditions of the regulated environments. Measurement costs are those borne by the lawmaker when assessing the economic impact of legal rule on parties’ behaviour. These are the costs of performing a cost-benefit analysis of legal rules; they tend to be higher when dealing with a heterogeneous demand for law.

On the demand side, people who are subject to the law incur significant costs as well. I distinguish between the costs that individuals bear in becoming informed about the content of legal rules and the costs incurred by individuals in undertaking the courses of action required by the law. I call the first “compliance information costs” and the latter “compliance conduct costs”. This distinction is of great importance from the process efficiency analysis viewpoint. The compliance conduct costs are included in the “substantive efficiency” calculus (i.e. they are the focus of investigation in the conventional output-oriented cost-benefit analysis). As such, these costs are assumed to be exogenous to the proposed law-making process model. Unlike conduct costs, compliance information costs are significantly affected by the structure of the law-


making process: as the cost of knowing the legal rules increases, fewer people will tend to abide by them.

Individuals subject to the law are involved in the law-making process not just as consumers of legal rules but also as participants in the production process. Participation requires gathering significant amounts of information, which is a costly activity for individuals. I call the costs associated with the participation in the law-making process “participation information costs”. Let me succinctly provide some explanatory examples. In the political process, voters bear significant information costs towards participating in the election of their representatives; in the judicial process, litigants spend a significant amount of resources in collecting and elaborating the information essential to the adversarial process; finally, individuals in the market process bear information costs towards acquiring the information necessary to engage in market transactions (i.e. information on the quality of the economic goods, reliability of the counterpart to the transaction, and so forth).

In sum, law-making information costs include (i) information-gathering costs, (ii) measurement costs, (iii) compliance information costs, and (iv) participation information costs. The structure of these costs is a function of the characteristics of the law-making process and the regulatory environment.

2.3.2. Agency Efficiency

The second procedural determinant of substantive efficiency is the degree of responsiveness of the law-making process to the interests of the individuals subject to the...
The law-making process can be conceptualized as a principal-agent relationship between the people, i.e. beneficiaries of the legal rules, and the lawmaker, to whom the people delegate the power to produce legal rules. The law-making process should serve the interests of the people who are subject to law. However, in many cases, the incentives of the lawmakers and those of the people who are subject to law diverge. Additionally, the actions of regulators are not easily observable by individuals affected by regulation, so people who are subject to the law face considerable difficulties in monitoring the activities of the lawmakers and ensuring that legal rules serve their own interests rather than those of the rulers or organized minority interest groups. In sum, diverging incentives and difficulties in observing lawmakers’ activities create room for severe agency problems that affect the quality of the production of law.

---

206 Once ignorance of the social welfare function is assumed, an important theoretical problem is that of defining the “interests of the people subject to law”. In particular, the question arises of whether people are able (e.g. bounded rationality and limited information) to appreciate their interests, especially in cases where people’s interests are properly assessed in the long-term perspective. I will discuss these issues in chapter 6, in the context of an analysis of the political process.

207 The literature on this point is extensive. Providing a comprehensive analysis of the issue is beyond the scope of this study; I will limit the discussion to the role played by the agency problem in the definition of the law-production techniques and in the comparative-institutional analysis. See, Parisi, Sources, supra note 196 at 96.


Historically, the ability of organized minorities to influence political power at the expense of large non-organized majorities has been emphasized by representatives of the “elitist theory” originally developed by Robert Michel (developing “the iron law of oligarchy”), Vilfredo Pareto (emphasizing the role of the “political elites”), Gaetano Mosca (identifying the organizational superiority of an organized “ruling class” over the unorganized majority of the population). See, respectively, Robert Michels, On the Sociology of Political Parties in Modern Democracy: a Study on Oligarchic Tendencies in Political Aggregations, translated by Eden and Cedar Paul (English New York: Dover Publications, 1959); Gaetano Mosca, The Ruling Class. Elementi di Scienza Politica, (New York: McGraw-Hill Book Company, trans. HD Kahn, 1939); Pareto, Manuale, supra note 81.
As it will become clearer when comparing alternative law-making institutions, the agency problem, i.e. the various forms of shirking by agents, is present in every law-making process, whether political, bureaucratic, judicial, or spontaneous in nature. Since it is costly for the principals to monitor the use of inputs and the production of output by agents, problems such as moral hazards arise and generate divergence between the principal’s interests and the agent’s actions. In response to this and other problems of incentive misalignment between lawmakers and law beneficiaries, law-making institutions devise various institutional arrangements. Over the past few decades, legal theorists, political science scholars, and economists have investigated how rational actors can control the behaviours of the agents to whom they delegate law-making authority in various institutional settings. One of the major conclusions in this literature is that institutions evolve and survive to mitigate these problems.

The production of welfare-enhancing legal rules requires minimization of the agency problem. This can be achieved through two complementary strategies: (1) shaping the structure of the decision making process in such way as to align the incentives of the lawmakers with those of the people subject to the law; and (2) fostering processes for

---


Economic theory looks at firms, hierarchical relationships among owners, managers, and workers; see, for example, Armen A. Alchian and Harold Demsetz, “Production, Information Costs, and Economic Organization” (1972) The American Economic Review, 777 (Alchian and Desmetz, “Production Information Costs”); Bengt Hölstrom, “Moral Hazard and Observability” (1979) The Bell Journal of Economics, 74. In general, economic literature emphasizes the impossibility and even undesirability of perfect control owing to problems of asymmetric information, monitoring costs, and inherent uncertainty about future conditions.

monitoring lawmakers’ activity and accountability. The analysis of the supply side of law-making processes will illuminate various agency-cost-minimizing institutional strategies.

2.3.2.1. Agency Costs

For process efficiency analysis, “agency costs” are the costs generated by the lack of responsiveness of the law-making process to the interests of the individuals’ subject to the law due to the opportunistic behaviour on part of the lawmakers. It is useful to distinguish among three different categories of agency costs: (i) external pressure costs (or influence agency costs), (ii) internal pressure costs, (iii) outcome agency costs, and (iv) the costs of reducing the agency costs.

(i) In the first instance, departure from welfare-maximizing decision-making originates from inside the law-making arena due to the misalignment of the agent’s and the principal’s incentives. That is, the first source of agency cost is the internal pressure associated with the lawmaker’s moral hazards (i.e., opportunistic behaviour of lawmakers to the detriment of the principal’s interests). Internal pressure costs manifest in various forms, e.g. rent-extortion costs and transaction-augmented costs, in the political process.\(^{212}\)

(ii) External pressures are generated by incentive misalignment among different groups of principals – i.e. rent-seeking competition among various pressure groups leads to dissipation of resources that could be employed more productively elsewhere. This resource dissipation is referred to in the literature as “rent-seeking costs”.\(^{213}\) To summarize, external and internal pressures differ with respect to the source of the rent-seeking activity that affects the decision-making process. While internal pressure derives from rent-seeking efforts by officials operating within the law-making arena, external

\(^{212}\) For detailed discussion of rent-extortion costs and transaction-augmented costs see infra sub-section 5.3.2.

\(^{213}\) See infra sub-section 5.3.1.2.5.
pressure derives from various stakeholders’ efforts to extract rents from the law-making process.  

(iii) While external and internal pressure costs are independent of whether the rent-seeking efforts by agents and principals succeed, a third important category of agency costs is associated with successful rent-seeking. I will refer to these costs as to “outcome-agency costs”. Bluntly put, these are the costs of bad law originated due to both opportunistic behaviour of the lawmakers and rent-seeking pressure by various stakeholders. Internal pressure costs manifest in various forms, e.g. monopolistic-rent costs in the political process and agency slack by bureaucrats in technocratic law-making.

(iv) Finally, another important category of agency costs is associated with the institutional arrangements designed for minimizing the rent-seeking pressures. Think, for example, to the costs associated with institutional arrangements designed to minimize the lawmakers’ moral hazards in bureaucratic processes (i.e., the costs associated with ex ante and ex post control of bureaucratic action by politicians). These can be referred to as the “costs of reducing the agency costs”.

In sum, agency costs result from the sum of (i) internal pressure costs (including the costs associated with agency-cost-minimizing mechanisms), (ii) external pressure costs, (iii) outcome-agency costs, and (iv) the costs of reducing the agency costs.

2.3.3. Social Choice Efficiency

The third procedural determinant of substantive efficiency is the ability of the law-making process to minimize the costs associated with the social choice problem – i.e. to aggregate individual preference orderings so as to generate a rational, stable collective outcome and be responsive to individuals’ preference intensity. The social choice problem is independent of the agency problem: the lawmaker is confronted with it even

---

214 See Moe, “Political Institutions” supra note 208 at 233 (“Citizens commonly have two enemies to fear in politics: other citizens and public officials”).
assuming zero agency costs.\textsuperscript{215} This does not exclude the possibility that the social choice problem may exacerbate the agency costs associated with law-making.\textsuperscript{216}

The first aspect to be considered is that the structure of the decision-making process might distort (rather than reflect) social preferences. With individual preferences held constant, alternative decision-making patterns employing different aggregating criteria are likely to produce different legal outcomes. That is, institutional choice plays the twofold role of being concurrently the subject and the object of social choice. It is the subject because the result of the institutional choice largely affects the outcome of the law-making process; it is the object because the choice of law-making institution is the result of society’s prevailing preferences and interests.

Second, public choice literature has long demonstrated the possibility that the optimum equilibrium does not exist under collective legal decision-making processes; that is to say, the outcomes of collective decision-making might often not be entirely rational.\textsuperscript{217} In the political process, this is related to the difficulties of aggregating individual preferences on a variety of social issues (i.e. cycling preferences, difficulty in capturing the intensity of individual preferences, issue-bundling problem, and so on). Despite the focus of public-choice scholars on the political process, many studies demonstrate that other law-making institutions, too, deal with the problem of non-rational outcomes. Think, for example, of the presence of strong path-dependence effects in judicial law-making, or, of the so-called “empty core” among private actors.\textsuperscript{218} I will discuss these problems in part III of this study. For the moment, it is important to emphasize that the law-making process—indeed of its institutional design—is plagued by the problem of incoherent results.

\textsuperscript{215} See, Parisi, “Sources”, supra note 196 at 8: [“even if we contemplate a world of perfect incentive alignment between political representatives and the represented citizens (i.e., even if we assume away agency problems in political representation), there is no assurance that the mechanisms of law creation are responsive to the underlying preferences of individuals in society”].

\textsuperscript{216} See infra sub-sections 5.2.5.2.2 and 5.3.2.1.1.


\textsuperscript{218} The problem of path dependence in adjudication is discussed in section 7.3., while the problem associated with private bargaining lacking a “core” is discussed in section 8.2.
As I will explain subsequently, irrational outcomes are not a problem *per se*, but to the extent they compromise the perceived *legitimacy* of law-making. Path independency with respect to a decision-making process ensures that the outcome quality does not depend on arbitrary or strategically selected choice paths, but on the *intrinsic* quality of the outcomes. A related concern is to avoid the risk of *cyclical* decision-making, i.e. those situations where an alternative preferred outcome is always available so that the decision-making process never reaches a stable outcome. This is generally referred to in the literature as “intransitive social preferences”. Cyclical decision-making is highly problematic for two reasons: (i) it impairs outcome predictability and (ii) it poses the problem of agenda control. One of the tasks of process efficiency analysis is identifying the institutional design that minimizes this negative impact of the social choice problem on the law-making process.

### 2.3.3.1. Decision-Making Costs

The process of comparing and aggregating individual preferences is operated by a variety of institutional mechanisms that ensure responsiveness of the decision-making process to the preferences of the people subject to law. The functioning of these mechanisms often entails *bargaining*, 219 which, in turn, entails strategic behaviours. I call the costs associated with participation in the decision-making process as “decision-making costs”, which include the costs of bargaining and strategic behaviours associated with the decision-making process. Consider, for example, logrolling or vote-trading within the legislature, the bargaining component of the strategic interactions between judges in panel court, or bargaining activities associated with private transactions. The structure of the decision-making costs affects the behaviour of individuals within the law-making arena and, consequently, influences significantly the functioning of the law-making process. Finally, I view the social choice problem as related to the *supply* side of the lawmaking process. In private legal ordering, the supply and demand side are characterized by the same set of individuals, therefore in this case the decision-making costs can be viewed as affecting both the demand and supply sides.

---

Similar to agency efficiency, social choice efficiency has a significant *outcome* component. In particular, an involuntary redistribution of wealth is inherent in any collective decision-making process based on non-unanimous decision criteria. Non-unanimity ensures that the legal outcome is the source of utility loss for some participants in the decision-making process. For our purposes, these costs can be called “outcome social choice costs”; these are forms of institutionally induced externalities.

**2.3.4. Adaptive Efficiency**

The fourth procedural determinant of legal efficiency lies in the *ability of the law-making process to produce legal rules that are adaptive to the changes in the regulated environment*. Stated differently, the law-making process maximizes social welfare by optimizing the trade-off between legal certainty and legal change. Legal certainty is a foundational goal of any welfare-maximizing law-making process. At the same time, however, while legal certainty is essential to law-making, so is the efficient adaptation of legal rules to the evolution of the social and economic environments.

The optimization of the trade-off between legal certainty and legal adaptation can be understood usefully from the perspective of individuals’ attitudes to legal risk. In particular, the common assumption in economics that people are risk-averse proves valuable when imported into the context of legal risks. In fact, to some extent, people are willing to sacrifice a certain amount of outcome average efficiency in order to be able to easily predict the legal consequences of their behaviour. The reason for this behaviour is that individuals gain high expected benefit from the predictability of the law. To benefit from higher predictability, individuals are often willing to give up optimal allocations in a number of instances in exchange of a stable set of legal rules. In addition to this demand for legal predictability, individuals demand some degree of legal adaptation to exogenous social and economic changes. Hence, there is a need for counterbalancing the demands for legal certainty and legal change.

---

220 See on this point *infra* sub-section 5.2.2
A further related issue is the relationship between the adaptive efficiency of law-making and the dynamic efficiency of economic environments. That is, the propensity of law-making to incentivize (rather than obstruct) economic innovation produces important consequences for long-term economic development.\textsuperscript{221} To maximize social welfare, law-making processes must adapt the production of law according to the rate of (economic and technological) change: more dynamic environments expose legal rules to greater obsolescence, thereby entailing higher legal adaptation; more stable environments entail lower legal change and are consequently compatible with more rigid types of regulation.

As the discussion progresses, it will become clear that various forms of inertia and path-dependence affect all law-making processes.\textsuperscript{222} However, both sources of inertia and mechanisms of legal change vary across law-making institutions and regulatory environments. On one hand, the legal system tends to \textit{preserve} the informational capital accumulated from the past; on the other hand, it incentivizes \textit{evolution} through a vast array of institutional arrangements. These two driving forces of law-making are differently shaped depending on the features of institutional law-making design. For example, judge-made law resembles the evolutionary model: the conservative tendency takes the form of judicial decisions’ path-dependence, while the innovative tendency is characterized by decisions overruling legal precedents. Legal change originated by the political process is more similar to the innovative process generated by private markets, characterized by higher discontinuity compared with the evolutionary model.

In conclusion, the features of the supply and the demand side determine the structure of \textit{adaptive costs}. Alternative sources of law are characterized by different mechanisms for balancing legal certainty and legal change. The important question that process efficiency analysis raises is that of the factual situations that allow for relatively higher adaptive

\textsuperscript{221} See, Douglas North, \textit{Institutions, Institutional Change, and Economic Performance} (Cambridge: Cambridge University Press, 1990) at 80 [North, “Institutions”]: (“Adaptive efficiency […] is concerned with the kinds of rules that shape the way an economy evolves through time. It is also concerned with the willingness of a society to acquire knowledge and learning, to induce innovation, to undertake risk and creative activity of all sorts, as well as to resolve problems and bottlenecks of the society through time”).

\textsuperscript{222} On path dependency, see North, “Institutions” supra note 221; Avner Greif, “Cultural Belief and the Organization of Society” (1994) J. Pol. Econ. 102 912; and Peyton H Young “The Economics of Convention” (1996) 10 J. Econ. Perspectives 105.
efficiency of one law-making institution compared with those of the feasible institutional alternatives.

2.3.4.1. Adaptive Costs

It is useful to distinguish between different types of adaptation costs. (i) The *adaptation transaction* costs are the costs incurred by individuals to appreciate the efficiency advantages of a superior legal regime and to coordinate simultaneous, mutually advantageous migration towards the new legal regime. Think, for example, about the costs of appreciating the advantages of a new social norm or the adoption of a new technology by members of all groups.

(ii) The *adjustment* costs are the costs individuals incur to transit to the new legal regime and change their behaviour as required by the law. These are conduct costs in nature, that is, they are related to the behavioural change imposed on individuals by the legal change.

(iii) *Resistance* costs include the cost increment borne by individuals towards the maintenance of the old behaviour. Legal change is often associated with resistance from those individuals who suffer losses due to the new legal rules (i.e., losers).²²³ Resistance comes also from individuals who are slow to appreciate the advantages of legal change. Both losers and individuals that do not appreciate the efficiency advantages of legal change tend to stick to the “old” norm, thereby incurring resistance costs. In general, the individual chooses to adjust to a new legal regime up to the point where the present value of an increment in adjustment costs equals the present value of an additional increment in resistance cost.

(iv) Finally, adaptive efficiency has an important *outcome* dimension. That is, inefficient adaptation of the production of law to exogenous changes in the demand for law generates substantive inefficient norms. I refer to the costs generated by substantively inefficient legal rules resulting from the inefficient adaptation of law-making to changes.

in the demand for law as *maladaptation* costs. In the event of a change in the regulatory environment, legal rules that were previously efficient may become inefficient. Mismatch between the supervening inefficient rule and the changing context determine the magnitude of the maladaptation costs.

The structure of the adaptation costs depends on the differential between the “*rate of adaptation*” of the law and the “*rate of change*” in the regulatory environment, as well as variations in said differential across alternative law-making processes. Another important factor is the heterogeneity of individual *preferences* about the desirable legal change. In general, it is possible to identify two groups of individuals: (i) legal *entrepreneurs* (i.e., those favouring legal change, and (ii) opposing groups (i.e., those resisting legal change). Legal change will depend upon whether (a) the influence power is concentrated in the hands of legal entrepreneurs or in the hands of opposing groups (this, in turn, depends on the ability of legal entrepreneurs to coalesce effectively—organizational cost—and on the strength of the opposing groups), (b) legal entrepreneurs are rent-seeking or efficiency-seeking, and (c) the expected benefit of the new legal regime more than compensates for the adjustment costs.

An example might help to better understand the rationale underlying this taxonomy of the adaptation costs. Think, for example, about the introduction into the market at period $t_0$ of a new technology that increases safety in car usage. Assume that efficiency requires the installation of this new equipment by all drivers. Imagine that, in period $t_1$ (after $t_0$), courts react to this technological innovation by ruling that in cases of car accidents, where a car in which the new safety equipment is yet installed is involved, negligence on part of the driver of the “non-adapted” car is presumed. So, in period $t_0$, driver negligence should be proved, whereas in period $t_1$, the driver is presumed to be negligent. The change in the liability regime generates all four categories of adaptive costs previously identified. In the period $t_1 - t_0$, the liability regime remains unchanged. There is no incentive for car drivers to install the new equipment, so a number of accidents occur that could have been avoided by a more timely change in the liability regime the by

---

224 On maladaptation costs, see also *infra* sub-section 3.1.4.3.
courts (maladaptation costs). In response to the new legal regime, car users have to buy the new technology, install it in their car, and obtain a security certificate attesting the regularity of the new equipment. These costs are the adjustment costs incurred by individuals in transitioning to a new legal regime. As an effect of the new legal regime, driving a car with the old technology is more expensive because the expected costs of car accidents increase. These increased costs of driving a car with the old technology are the resistance costs, i.e. the costs of maintaining the old behaviour. Finally, the costs incurred by drivers to appreciate the existence of the new safety-increasing technology and to coordinate with each other to transition simultaneously to the new technology are the adaptive information costs. If adaptive transaction costs are low, individuals identify the existence a more efficient technology and transition spontaneously to the new technology. That is, people recognize the superiority of the new technology, and coordinate to comply simultaneously with the new legal regime, well before the establishment of the new rule by the courts. On the contrary, the higher the adaptive transaction costs, the less likely are people to recognize the existence of a superior technology, and, consequently, it would be more unlikely that they will spontaneously coordinate to transition to the new legal regime.

2.4. Process Efficiency Logic

This section clarifies the test employed in this thesis to compare the efficiency features of alternative law-making processes. It then examines the logical underpinnings of process efficiency analysis: (i) the assumption of collective ignorance of social welfare, (ii) shift of focus from the regulation of incentives to incentives for regulation, and (iii) concept of demand and supply of law.

2.4.1. Efficiency Test

*An efficient law-making process is shaped such that it better facilitates the production of welfare-enhancing legal rules in comparison with the feasible institutional alternatives.* Let me illustrate three important points: i) the proposed notion of process efficiency meets the requirements for a sound theoretical definition of efficiency; ii) process efficiency incorporates the *unavoidable* law-making costs and results in a *comparison*
among imperfect institutions; and iii) process efficiency is based on consensus and not vice versa.

First, process efficiency is consistent with the requirements for avoiding logical circularity, as identified in section 1.1.3. (a) The output of the law-making process is defined in terms of the production of welfare-maximizing legal rules. (b) The concept of process efficiency is used as a comparative criterion on a set of available alternatives identified on the basis of defined structural criteria that I will define subsequently. (c) The interests that are relevant to the identification of the notion of efficiency are defined. This latter point is subtler than it might appear. It is not always easy to define the interests at stake in the law-making process. In general terms, I define an efficient process as the one that maximizes the incentive to produce law that maximize the economic welfare of the people subject to the law. From this perspective, the interests that matter for the definition of process efficiency are the individuals and groups whose behaviours are regulated by legal rules (i.e., target actors) and those whose interests are served by the law (i.e., norm beneficiaries). I will consider more delicate issues (i.e., the impact of the time horizon on the definition of the relevant interests) as the discussion proceeds.

Second, process efficiency does not arise out of an assessment against the ideal standard of a zero-transaction-cost world; rather, it is based on the premise that each source of law is highly imperfect and exhibits a significant level of internal transaction costs. The capacity of alternative law-making institutions to yield a welfare-maximizing outcome

---

225 The importance of identifying the set of peoples whose interest is relevant for selecting efficient institutions is a point of fundamental importance for a theoretically sound comparative efficiency analysis. Buchanan explains this point through the example of the prisoner’s dilemma, i.e. a setting where two persons are induced to confess by preventing them from cooperating through mutual agreement. An incentive structure such as the one embedded in the prisoner’s dilemma setting might be considered “inefficient” when evaluated from the viewpoint of the prisoners themselves; it might otherwise be defined as an “efficient” institution for inducing prisoners to confess when the set of relevant interests includes those of individuals members of the society as whole. See, James Buchanan, “Rights, Efficiency, and Exchange: The Irrelevance of Transaction Costs” in “Anspruche, Eigentums und Verfungungsrechte, [“Liberty, Market And State: Political Economy in the 1980s”], (Berlin, Duncker und Humblot, 9–24, 1984), reprinted in James M, Buchanan, Liberty, Market, and State: Political Economy in the 1980s, (Brighton, England: Wheatsheaf Books, University Press, 1985) at 92–107 [Buchanan, “Rights”].
reflects the relative magnitude of internal frictions. That is, process efficiency incorporates the set of unavoidable costs of law-making into the calculus. As we have seen, the taxonomy of the law-making costs reflects a spectrum of institutional imperfections related to institutional design. From this perspective, imperfection is not synonymous with inefficiency. A law-making process is inefficient not because it is imperfect, but because there are less imperfect institutions capable of better accomplishing the task of producing welfare-maximizing legal rules in a given regulatory framework. Process efficiency is the result of a comparative assessment of relative imperfections, rather than a consequence of the elimination of transaction costs.

Third, process efficiency implies a reversal of the relationship between efficiency and consensus, which is traditionally assumed by the output-oriented law and economics approach. Under the orthodox methodology, consensus on legal rules is based on efficiency. Thus, for example, one of the tenets of law and economics contract theory is the criterion of hypothetical bargaining. Theorists hypothesize that the parties to a hypothetical bargain would voluntarily consent to the rule that generates the most efficient outcome. On closer examination, this assumption is logically rooted in the idea that the efficiency of an allocative outcome should be evaluated independently of the characteristics of the allocative process. By contrast, the structural approach presumes that efficiency is based on consensus. Efficiency is not a property of the outcome independent of the process; rather, efficiency depends upon the process embodying general consensus on the institutional rules. Because efficiency is shaped by the decision-making mechanism, the structure of the process itself is the object of efficiency

---

226 See, Komesar, Imperfect Alternatives, supra note 7 at 5 (“Institutional choice is difficult as well as essential. The choice is always a choice among highly imperfect alternatives. The strengths and weaknesses of one institution versus another vary from one set of circumstances to another”).

assessment; furthermore, as explained previously, the possibility that the final outcome might be allocatively inefficient in a number of instances is contemplated.

2.4.2. Presumption of Collective Ignorance

An investigation of process efficiency requires a preliminary assumption of the goals of the law-making process.\textsuperscript{228} As mentioned above, I assume that the ultimate goal of the law-making process is to produce legal norms aimed at maximizing social welfare. In this section, I briefly discuss this assumption, which I construct on the basis of a subjective conception of individual preferences.

The traditional approach to law and economics is based on the idea that evaluation of the outcomes of the law-making process is carried out based on cost-benefit analysis, which is capable of measuring the allocative efficiency of legal rules. This methodology rests on the methodological presumption that an external observer can hypothesize about individual preferences: the economist, social scientist, and policymaker are assumed to be capable of recognizing an increase in social welfare and predicting the choice that an individual would make from among a set of feasible alternatives. In this view, the social welfare function is assumed to be something that the external observer can penetrate and read through: the very idea of allocative efficiency would not be logically consistent without assuming the possibility of knowing the scale of social collective preferences. In relation, the economic process is conceptualized in terms of maximization decision-making and relies on the premise that the ends of the economic activities in question are given and externally knowable.\textsuperscript{229}

\textsuperscript{228} Efficiency is a relative concept. As such, it logically presupposes a clear definition of the relevant output (i.e., what has to be efficient about). With specific regard to the law-making process, the construction of a logically consistent notion of efficiency of law-making presupposes the antecedent specification of what the rule-producing mechanism has to be efficient about.

The structural approach negates the possibility of an acknowledgeable, objectively definable social welfare function. Individual utility becomes inherently subjective and cannot be measured by an external observer based on objective criteria. Thus, economic efficiency cannot be used as an objective social choice criterion; only individuals can define their own subjective ends, so that no external mechanism is able to aggregate individual preferences on the ground of a social scale of values formed independently of individual preferences. As a practical matter, the identification of the most desirable allocation of resources should be based on individuals’ preferences as revealed by their actual choices. In other words, the process of aggregating individual preferences itself becomes the object of the efficiency standard. Since no outcome-criteria can be developed for social decisions, the efficiency evaluation of legal production must be grounded in the way that people agree to generate enforceable legal rules to improve social welfare. This is a crucial point for understanding the logic of the structural approach. The efficiency of the outcome is not directly measured; rather, it is inferred from the structure of the process whereby it is generated.230

In conclusion, the process efficiency analysis focuses on the rules that shape the structure of the decision-making process. The presumption of an external observer capable of constructing a social welfare function is dropped, and the collective ignorance of outcome efficiency is assumed. In this sense, process efficiency measures the capacity of individuals as collective decision makers.231

---

230 See Buchanan, “Rights” supra note 225 at 14 (Agreement on a change in the rules within which exchanges are allowed to take place would be a signal that patterns of outcomes reached or predicted under the previously-existing set of rules are less preferred or valued than the patterns expected to be generated under the rule-as-changed. Hence, the new rule is deemed more efficient than the old. The discussion and agreement on the change in the rule here is analogous to the trade that takes place between ordinary traders in the simple exchanges made under postulated rules”).
2.4.3. Regulation of Incentives vs. Incentives for Regulation

One of the logical consequences of the shift in focus from the outcome to the process is the analytical shift from the “individuals’ incentive provided by the law” to the “individuals’ incentive to produce efficient law”. The structural approach focuses on incentives to engage in the production of efficient legal rules, rather than on the incentives provided by efficient legal rules.

The methodological advantage of this approach can be understood when one considers the informational difficulties associated with the traditional cost-benefit analysis of legal rules. Identifying an efficient incentive structure for producing law requires an economic evaluation that is different in nature from that required for measuring the relative efficiency of alternative legal rules. First, traditional cost-benefit analysis often relies on simplistic assumptions about the subjective preferences of private parties. By contrast, the structural approach formulates hypotheses to determine the institutional path that better facilitates the production of legal rules responsive to the preferences of the parties impacted by the law. Thus, while traditional cost-benefit analysis requires the measurement of the allocative efficiency of legal rules by making simplifying assumptions about the structure of individuals’ preferences (i.e. based on the willingness to pay), the structural approach proceeds by making inferences on the efficiency features of the legal rules by analyzing the incentive structure embedded in the institutional law-making design. When the utilitarian calculus of allocative efficiency proves difficult (i.e. when the willingness to pay is not a good proxy for individual preferences), inferring preferences from the structure of the law-making process proves useful. Second, cost-benefit analysis often leaves unanswered the question of technical feasibility of the efficient suggested normative solutions. By contrast, the structural approach allows

---

232 An example might better clarify this point. Law and economics practitioners usually study the economics of product safety by focusing on the incentives for undertaking risk-analysis and investing in the safety precautions provided to firms by alternative legal arrangements. A structural approach to the economics of product safety focuses on questions such as whether private parties have the incentive to subject themselves to a regime of tort liability; whether public provision of legal rules (by courts, agencies, or legislatures) is well suited for achieving the goal of overcoming market failure generated by information asymmetry; what is the mix of alternative sources of law that optimizes alternative comparative advantages in terms of law-making efficiency, and so forth.
inquiry into whether the incentive structure embedded in the law-making process can realistically be expected to produce substantively efficient legal rule.

In conclusion, the structural approach affords two advantages: (i) efficiency of legal rules is inferred from the characteristics of the law-making process, rather than directly measured on the basis of simplifying assumptions about individual preferences; and (ii) it allows one to assess the technical feasibility of the efficient legal rules identified by the conventional law and economics approach.

2.4.4. Supply of and Demand for Legal Rules

Let me now provide an overview of the model of law-making that I will construct in the chapters that follow. I organize the analysis of the causal factors determining the structure of the law-making costs along the lines of a supply–demand model. The supply-side summarizes the institutional features of the law-making process that are predicted to be the most relevant in terms of their effect on the structure of law-making costs (i.e., institutional variables). The demand-side summarizes two orders of variables. On the one hand, it includes the incentives, constraints, and preferences of the people subject to the law; these include both the people whose behaviour falls into the scope of application of the law (i.e., norm beneficiaries) and those whose interests are protected by the law (i.e., target actors). On the other hand, the demand for law is significantly influenced by the features of the regulated environment from which the need for law arises (i.e., structural variables). The discussion in the remaining chapters of this thesis focuses on identifying both the institutional and structural variables of the law-making process and explaining how they relate to each other. The interaction between the demand and supply sides determines the structure of law-making costs and explains the comparative advantages in terms of efficiency of alternative law-making institutions across a variety of regulatory environments.
Let me finally emphasize the characteristics that distinguish demand and supply in the law-making context from conventional demand and supply in private markets.\textsuperscript{233} First, while in private markets supply and demand interact with each other through the price mechanism, here, supply and demand influence each other \textit{directly} without price intermediation. The individuals subject to the law can act concurrently on the demand and supply side. For example, the people subject to the law influence the outcome of the law-making process by electing their representatives to legislative bodies or by voting in referenda consultations. For this reason, as I will discuss at length in chapter 5, the \textit{distance} between the two sides of the law-making process significantly influences the structure of law-making costs. Conversely, the demand for law can also be significantly influenced by the supply side of the process. For example, as I will explain in chapter 3, the legal definition of the scope of the law provided by a lawmaker influences the \textit{aggregate} frequency and the frequency \textit{per actor}. In conclusion, the individuals on the demand side directly influence the production of law; likewise, law-making institutions directly shape some important features of the demand for law.

\textbf{2.4.5. Law-making as a “Second-Order” Collective Good Problem}

The demand for legal rules emerges from interactions among economic actors in situations where individual incentive structures prevent people from cooperating or from engaging in efficient bargaining.\textsuperscript{234} These situations are referred to as “social dilemmas” in sociological literature.\textsuperscript{235} Essentially, legal rules serve to improve the social welfare of norm beneficiaries by enabling them to overcome social dilemma situations. Social dilemmas entail coordination of the collective action required to produce the collective good from which every community member benefits.

\textsuperscript{233} From a different perspective, but confronting with a similar methodological problem, see Dari-Mattiacci and Deffains, “Uncertainty”, supra note 6. Although they adopt different assumptions regarding the object of the law-making process (for them, it is the reduction of law’s uncertainty), they correctly point out the peculiarity of using the concepts of demand and supply in the context of the production of law: “The relationship between litigation and uncertainty may be loosely analogized to a demand-supply model for the specific good of our concern: certainty of law. This analogy is only suggestive and may help us interpret the results in a graphical manner, \textit{but it should not be understood as carrying the usual meaning in terms of prices and quantities exchanged in a market} [emphasis is mine]”.

\textsuperscript{234} See Buchanan, “The Reason” supra note 171.

For the purposes of this study, it is important to identify the twofold nature of social dilemmas by distinguishing between the collective action problem that generates the need for law from the collective action problem that generates the need for law-making. These are referred to in the sociological literature as “first-order” and “second-order” collective good problems.\footnote{See, Pamela Oliver, “Reward and Punishments as Selective Incentives for Collective Action” (1980) Am. J. Sociol. 85 1356 [Oliver, “Rewards and Punishments”]; Douglas D Heckathorn, “Collective Action and Second-Order Free Rider Problem” Rationality and Society 1 78; James S Coleman, \textit{Foundations of Social Theory} (Cambridge, Mass.: Harvard University Press, Belknap Press) [Coleman, “Foundations of Social Theory”].} The first-order problem is to obtain conformity of every community member to the norm, so as to obtain each individual’s cooperative contribution towards the production of the collective good. Briefly, the first-order problem entails overcoming social dilemma by ensuring collective action through a norm requiring cooperation. However, maintaining and administering a system of collective punishment that sanctions norm deviators appropriately is itself a public good. In fact, producing and administering sanctions is itself a costly activity whose benefit is non-excludable. Consequently, many individuals may choose to enjoy the benefit of others’ participation in the collective punishment, regardless of their contribution to its provision. Hence, there is scope for a significant free-rider problem undermining the possibility of an effective enforcement system. This has been termed the “second-order collective action problem”. Law-making processes can be usefully conceptualized as institutional arrangements evolving in response to the second-order collective action problems inherent in the production of law. The output-oriented approach focuses on the first-order collective action problem, while overlooking the costs associated with the second-order collective action problem. Process efficiency analysis aims to investigate the relationship between first-order and second-order collective action problems.

Conclusions

Let me recapitulate the main findings of this chapter through the following summarizing propositions.
(i) Once process choice is included into the efficiency calculus, Paretian logic leads to identification of optimal legal rules that may not generate allocative efficiency results in a given number of cases. In particular, once Paretian logic is applied at the constitutional stage choice, the problem of allocative efficiency of legal rules becomes one of minimization of the differential between the aggregate surplus that would be associated with the sum of Pareto-optimal terminal allocative results and the aggregate surplus associated with the optimal legal rule agreed upon by all members in the group. Misalignment between optimal legal rules and optimal allocative results arises owing to the following two reasons. First, once Paretian logic is applied at the constitutional stage choice, allocative efficiency is assessed from an aggregate perspective that necessarily takes into account the variance of situations that the rule is to be applied to. Second, once the law-making costs are factored into the efficiency calculus, there arises the possibility of misalignment between cost-minimizing and surplus-maximizing law-making processes.

(ii) The traditional outcome-oriented approach loses sight of the three ways whereby process choice influences social-welfare maximization: (a) minimization of the individual costs of taking part in the production of law; (b) maximization of the number of people who decide to enter into the social contract and obey the law; and (c) minimization of the problems internal to the law-making process that hinder the production of substantively efficient legal rules.

(iii) The efficiency test for comparing alternative law-making institutions can be summarized as follows: a law-making process is efficient if there is no other alternative institution that does better, across the circumstances in which it actually operates, in the production of welfare-enhancing legal rules. This is assessed on the basis of four characteristics of the law-making process: (a) productive efficiency; (b) agency efficiency; (c) social-choice efficiency; and (d) adaptive efficiency.

(iv) I assume that the goal of the law-making process is the maximization of the social welfare of individuals. However, I presume collective ignorance about the scale of social
preferences. The logical consequence of the collective ignorance presumption is the analytical shift from individuals’ incentive as shaped by efficient law to individuals’ incentive to produce efficient law.

(v) The focus on incentives embedded in the law-making process has two informational advantages: (a) efficiency of legal rules is inferred from the characteristics of the law-making process rather than directly measured; and (b) analysis of the efficiency of legal rules is integrated with the assessment of the technical feasibility of the normative solutions suggested by the conventional law and economics approach, i.e. assessment of the ability of the available law-making institutions to generate legal rules identified as efficient through output oriented analysis.
Part II

FUNDAMENTAL LAW-MAKING DIMENSIONS

In Part I, I (i) emphasized the highly problematic nature of allocative efficiency in the legal context; (ii) proposed an alternative notion of process efficiency that helps explain and assess alternative lawmaking processes; and (iii) identified a uniform taxonomy of lawmaking costs. The discussion of these points prepared the methodological groundwork for a systematic analysis of the relationship between the institutional features of alternative lawmaking processes and the efficiency properties of the outcome. That is, the notions of process efficiency and the taxonomy of lawmaking costs identify the elements that must be considered to understand the relationships between the characteristics of the institutional lawmaking design and the efficiency of the law.

In Part II, I identify the most relevant institutional and structural variables affecting process efficiency. In particular, for each component of process efficiency (i.e. productive, social-choice, agency and adaptive efficiency), I discuss the institutional features of lawmaking processes and the conditions of the regulated environments that most affect the structure of lawmaking costs (identified in the taxonomy provided in Chapter 2).

I focus on four variables, which I refer to as the ‘fundamental’ dimensions of lawmaking. As shown by the discussion, these factors are fundamental because they crucially affect the structure of lawmaking costs. Specifically, I contend that lawmaking costs are a function of two institutional dimensions of lawmaking: (i) ex ante / ex post lawmaking and (ii) centralized / decentralized lawmaking. These two institutional dimensions are sources of comparative advantages and disadvantages in terms of process efficiency depending on the structural conditions of the regulated environment. With respect to the structural variables, I focus the analysis on the (i) degree of homogeneity and (ii) the degree of frequency of the cases to be regulated.
The four-dimensional framework developed in Part II allows the formation of hypotheses regarding the core tendencies underlying the production of law, which will be verified in Part III. Part III will examine four alternative lawmaking institutions (i.e., politics, bureaucracy, adjudication and spontaneous lawmaking). Each of these institutions is characterized by alternative combinations of \textit{ex ante} / \textit{ex post} and \textit{centralized} / \textit{decentralized} lawmaking as well as additional institutional features (e.g., political representation, bureaucratic insulation, doctrine of precedent, and so on). The comparative institutional analysis conducted in Part III will enable testing of the hypothesis formulated in Part II and identify some core tendencies that characterize the production of law.
Chapter 3

EX-ANTE VERSUS EX-POST LAW-MAKING

The village authorities were seriously considering at the time the advisability of using this car instead of a mule to call every day for the mail; by so doing they would set up a regular service for travelers coming or going by bus. But because a man’s time and labor were unimportant in these parts and cost practically nothing, there was quite a difference in the expense, and perhaps, also various family and social relationships had to be taken into account; in any case the question was postponed from day to day, and when I left it was still unsettled.

Carlo Levi, Christ Stopped at Eboli, 1945

This chapter examines the trade-offs associated with the choice between ex-ante and ex-post law-making. The main part of the discussion is focused on productive efficiency and is divided into two optimization problems: a) the minimization of production-information costs and b) the maximization of legal compliance. The last three sections examine briefly the other components of process efficiency (i.e., social-choice, agency and adaptive efficiency).

The analysis of the ex-ante versus ex-post choice, coupled with the discussion (developed in the next chapter) of alternative levels of centralization, leads to the formulation of hypothesis that will be tested in a discussion of alternative law-making institutions, which is provided in Part III.

---

237 This chapter makes extensive use of the analytical framework provided by the neoclassical theory, properly adapted to the context of the law-making. I cannot overemphasize that the neoclassical cost-minimizing model is not in itself sufficient to capture the complexity of the process-choice: a full understanding of the set of economic problems that I have called “process efficiency” requires further analysis, along with the relaxation of the simplifying assumptions of the neoclassical model. Nonetheless, the cost-minimization model offers a useful starting point and provides a good understanding of some of the fundamental trade-offs entailed by alternative law-making activities. In the context of the comparative analysis of alternative law-making institutions, I will adopt other economic approaches (e.g., public-choice, transaction-cost economics, evolutionary economics, and so on), based on different assumptions (e.g., the self-interested nature of the lawmaker; or the rule-following rationality of the individual subjects to law will be of most use).
Introduction: Assumptions

Let me clarify the underlying assumptions of the analytical framework regarding both the demand and supply sides of law-making. First, I assume that individual compliance depends exclusively on the costs and benefits of compliance (information and conduct). Individuals who have full appreciation for the content of law choose to comply with the legal rules, unless either the cost of the legal conduct exceeds the expected benefits or the information-compliance costs outweigh the expected benefits. The information-compliance costs are the costs associated with gaining appreciation of the content of the law. Hence, I exclude from the scope of the analysis the vast array of non-economic factors that might influence the behaviour of individuals and induce them either to obey the law when compliance is economically inefficient or not to obey it when compliance is economically efficient.

Second, I assume—as a logical consequence of the previous assumption—that the lawmaker is able to influence people’s behaviour under the two conditions that it devises rules that are not too difficult to be appreciated and that the private benefits of the required conduct exceed its costs. If the sum of the compliance-conduct costs and compliance-information costs exceeds the private benefit from compliance, then the individual will decide not to abide by the rule. Because my focus is on compliance-information costs, I assume that the compliance-conduct costs are fixed and exogenously determined. That is, I do not consider the substantive efficiency of legal rules and how it affects the behaviour of people, which is the object of a traditional, outcome-oriented analysis.238

---

238 To clarify, I assume that people comply with the law as long as legal rules are substantively efficient (i.e. the costs of legal compliance do not exceed the benefits in terms of avoided legal sanction), and the costs of becoming informed about the law and complying with the law do not exceed the benefits (in terms of avoided legal sanctions). The first condition is exogenous to the model, while the second condition is related to the features of the law-making process (i.e. ex ante versus ex post specification of the legal content of the rule). Obviously, this is a simplifying assumption that does not consider the complexity of the phenomenon of legal compliance. On this latter point, see generally: Tom R. Tyler, Why People Obey the Law (Princeton: Princeton University Press, 2006); Stephen Shavell and Louis Kaplow, “Why is it Socially Desirable That People Obey the Law?” (2010) 682 Harvard Law, Economics and Business Discussion Paper; Robert Cooter, “The Intrinsic Value of Obeying a Law: Economic Analysis of the Internal Viewpoint” (2006) 75 Fordham Law Review 1275.
It is also necessary to simplify some assumptions with respect to the supply side. First, I assume that only one law-making institution has the *monopoly* of the production of legal rules. In addition, I assume that the lawmaker produces a *single output* (i.e., legal norm) by using alternative combinations of *two inputs* (i.e., *ex ante* and *ex post* specification of legal content) and uses a production function to describe the production possibilities of the lawmaker. The production process is structured in two stages: i) the promulgation and ii) the adjudication of legal norms. In this chapter, I do not conceptualize them as distinct processes operated by different law-making institutions. Instead, I consider promulgation and adjudication parts of the same production process, which are operated by the same producer. The process choice here is between different input-bundles (i.e., alternative combinations of *ex ante* and *ex post* law-making) *within* the same production-chain, rather than alternative law-making processes. I conceptualize the lawmaker as a *cost-minimizing actor* and examine its production behaviour by investigating the properties of the cost function of law-making. Coherent with the single-input assumption, the only costs that I take into account are those in acquiring all the information needed to produce legal rules. Subsequently, the analysis of the costs curves is followed by a discussion of the choice of the social-welfare maximizing level of production.

Second, I assume substantive efficiency as given and exogenous to the model; that is, once individuals have applied legal rules, the individual-surplus maximization is achieved. Thus, the only way the lawmaker can maximize social welfare is by maximizing the number of people who decide to abide by the rules. The lawmaker can influence the level of legal compliance by choosing the appropriate level of specificity of the legal command. Like any other firm, the lawmaker also faces a *technological*

---

239 Subsequently, I relax this assumption when analyzing alternative models of law-making processes characterized by different comparative advantages in terms of process efficiency.

240 To clarify, because I assume that the magnitude of the economic surplus is exogenous to the model (i.e., the individual economic surplus generated by legal compliance is given), the only way the lawmaker can increase the aggregate social surplus is by increasing the number of people who decide to obey the law and undertake the course of action as commanded by the law. I have previously assumed that the only reason individuals do not comply with the law is that the costs of becoming informed by the law exceed the expected benefits. Therefore, the critical variable determining the level of legal compliance is the cost of legal information, which in turn is a function of some characteristics of the law-making process. This
constraint (i.e., the production function) and a *market* constraint (i.e., the number of people that are willing to abide by the rules depending on the costs of compliance).

### 3.1. Productive Efficiency

I organize the discussion as follows. In section 3.1.1, I model the individual demand for *ex ante* legal specificity as a function of the expected net benefit from legal compliance. In section 3.1.2, I analyze the *technological* constraint confronting the lawmaker (i.e., the cost of the information required to devise efficient legal rules), and I attempt to identify the conditions for cost-minimization. In section 3.1.3, I discuss the *economic* constraint on the production of law (i.e., the demand for legal specificity) and describe the optimization problem confronted by the lawmaker. Finally, in sections 3.1.4 and 3.1.5, I deepen the analysis of legal frequency and heterogeneity, which represent the two conditions for realizing the cost-saving effect of *ex ante* law-making.

#### 3.1.1. The Demand for *Ex Ante* Legal Specificity

The *ex ante–ex post* choice affects the following: (i) the amount of compliance-information costs (*demand-side costs effect*); (ii) the individuals’ decision to comply with the law (*behavioural effect*); (iii) the structure of the production-information costs (*supply-side costs effect*).\(^{241}\) In this section, I analyze the causal mechanisms determining the demand-side cost effects and the related behavioural effect. In the next section, I will analyze the supply-side cost effect.

I conceptualize the choice of the optimal level of *ex ante* precision as an *output* choice. In this view, the lawmaker seeking to maximize social welfare undertakes two choices: first, it decides *how much* legal content to produce; and, second, it decides how much of it should be produced *ex ante* (promulgation stage) or *ex post* (adjudication phase). That is, the lawmaker chooses first the optimal *quantity* of legal content and second the optimal *allocation* to either the *ex-ante* or the *ex-post* dimension of law-making. Whereas the first

---

choice (i.e., optimal specificity of the legal rule) is exogenous to the model, the second choice constitutes the object of inquiry of the process-efficiency analysis.\(^\text{242}\)

Let me begin by considering the issue of how the choice between \textit{ex ante} or \textit{ex post} creation of law affects the behaviour of people through its effect on compliance information costs.\(^\text{243}\) To understand this point, it is useful to break the causal sequence into two distinct segments: (i) \textit{ex ante}–\textit{ex post} choice → compliance-information costs, and (ii) compliance-information costs → individuals’ behaviour. In subsection 3.1.1.1. –

---

\(^{242}\) To refer to the choice of the level of ex ante specificity of the legal command, I will use the following expression interchangeably: “ex ante-post choice”, “choice between ex ante or ex post creation of law”, “choice of the optimal level of ex ante legal precision”, or “choice of the optimal level of ex ante detail (or specificity, or precision) of legal rules”. By using these expressions, I distinguish the choice of the optimal level of precision of the legal command, from the problem of determining, for each level of precision, the optimal level of precision to be allocated to the “\textit{ex ante}” stage of the production of law. On this latter point, which is usually described as the choice between rules and standards, there has been a large debate in the law and economic literature. It is not my intention here to make an original contribution to this debate. I will largely draw from this important strand of literature in order to gain insights into my object of analysis, focused on the law-making information costs.


\(^{243}\) Before proceeding, let me clarify that the abstraction level at this stage of the analysis excludes all the specific institutional features of alternative law-making mechanisms (e.g., legislatures, judicial process, agency and bureaucracies, and private legal orderings). This simplifying assumption is justified by the consideration that the economic effects of the degree of \textit{ex ante} precision of the legal command operate independently from other institutional characteristics of the law production process (i.e., \textit{ex ante} law-making is carried out not only by centralized institutions, such as legislatures and agencies, but also by decentralized processes, such as private contracting). In brief, the \textit{ex-ante} component of law-making entails some economic trade-offs regardless of the specific institutional context in which the lawmaker operates.
3.1.1.4., I discuss segment (ii) and explain how the structure of compliance information costs influences individuals’ decisions to comply with the legal rules. In subsection 3.1.2., I then focus on segment (i) and explain how the choice between the *ex ante* versus *ex post* creation of law affects the structure of the compliance information costs.

### 3.1.1.1. The Compliance-Information Costs

**Assumption 1:** *I assume that legal rules are substantively efficient, and that in the absence of the cost of appreciating the law, individuals decide to abide by the rules.*

Suppose that an individual has to decide whether to comply with a legal command that imposes an obligation of engaging in a certain course of action. It is useful to conceptualize the decision-making process of the individual as a two-period game. In the first period, the individual decides whether to acquire information about the content of the law; in the second period, when the individual has become aware of the content of the law, he/she decides whether to comply with the law.

The individual’s decision to comply with the law depends upon the surplus generated by the substantive features of the legal rule. The individual decides to obey the law if the expected net benefit from compliance exceeds its cost; conversely, when the expected net benefit from compliance is negative, the individual decides not to comply with the legal rule. The individual’s expected net benefit from compliance is equal to the difference between the value of avoiding the legal sanction in case of violation and the costs of undertaking the behaviour required by the law. I assume that—after becoming acquainted with the content of the legal command—the individual is *certain* that the law applies to him or her (i.e., he or she falls within the scope of the law) but is *uncertain* as to whether the law will actually be enforced in case of violation.244

---

244 As I will subsequently clarify, I draw the distinction between the probability that the law applies to the individual’s behavior, denoted as δ, and the probability of enforcement, denoted as π. The individual weighs the two probabilistic factors at different stages of the game. Before acquiring knowledge of the content of the legal rules, the individual is uncertain as to whether or not his or her behavior falls within the scope of the law. To eliminate (or minimize) this uncertainty, he or she decides to bear the compliance-information costs and purchase legal information. At this point, the enforcement probability comes into play: in order to evaluate the expected net benefit from compliance, the individual weighs the net-benefit from compliance with the probability to be detected in case of violation.
Let $c_c$ denote the compliance-conduct costs, $b_g$ the benefit from compliance, and $\pi$ the probability that the individual will be detected and sanctioned in case of violation of the law. The expected net benefit of the individual is:

$$\pi b_g - c_c$$

I assume that both $\pi b_g$ and $c_c$ are positive and *exogenously* determined, and that

$$\pi b_g \geq c_c$$

or equivalently,

$$b_g \geq \frac{c_c}{\pi}$$

I will refer to (3) as “the assumption of substantive efficiency”; this is an analytically convenient assumption that allows focusing on compliance *information* costs without considering the effects on individual behaviour of the substantive features of the legal rule. Another convenient way to express (3) is that the expected net benefit from the conduct $b_n = (\pi b_g - c_c)$ is positive.

**Proposition 1:** For each level of substantive efficiency, there is a threshold level of expected net benefit from legal information below which it is inefficient for an individual to acquire legal information. If the magnitude of compliance information costs exceeds that threshold level, people subject to law are induced to remain rationally ignorant of the content of the law, despite the substantive efficiency of the law.

---

245 The probability of the enforcement of legal rules is an important element for the purpose of the process-efficiency analysis. The ability to ensure an efficient and effective enforcement is one of the distinctive features of alternative law-making institutions. However, for the sake of simplicity, in this paper I prefer not to analyze how $\pi$ is affected by the institutional choice of the law-making process. I leave the analysis of this point to subsequent comparative studies of the alternative sources of law.
Compliance information costs enter the analysis when the individual’s choice of becoming informed about the content of the law (i.e., the choice of whether to acquire legal information or not) is considered. At first, the individual ignores the content of the law and, consequently, is uncertain on whether the law applies to him or not. Namely, he or she knows that a law is in force, but he or she does not know whether his or her behaviour will fall into the scope of application of the law. In this situation, he or she decides to bear the cost of becoming informed about the legal command if the expected net benefit from compliance, weighed by the probability that the law applies to him or her, exceeds the compliance information costs. The expected net benefit of becoming informed of the legal command and abiding by the law is given by \( b_n \), weighted by the probability that the law applies to the individual that I denote \( \delta \).

If the individual decides to bear \( c_l \) and acquire legal information, then the expected net benefit is

\[
 b_i = \delta (\pi b_g - c_c) - c_l \tag{4}
\]

which can be equivalently expressed as

\[
 b_i = \delta (b_n) - c_l \tag{5}
\]

Thus, an individual will not acquire legal information if (5) is negative and will acquire information if (5) is non-negative. If we set (5) \( \geq 0 \), we obtain the condition under which the individual is willing to acquire legal information and behave accordingly:

\[
 b_i = \delta (b_n) - c_l \geq 0 \tag{6}
\]

for which the following condition is true:

\[
 b_n \geq \frac{c_l}{\delta} \tag{7}
\]

\textsuperscript{246} The probability that the law applies to the individual varies across individuals. However, for simplicity, I assume that this value is constant.
I will refer equivalently to (6) and (7) as the “condition of individual incentive compatibility”, and I will denote the threshold level for which (8) holds as equality as $b_n$. When the condition of individual incentive compatibility is met, so that the expected net benefit from legal information is greater than or equal to $b_n$, then an individual has an economic incentive to comply; on the contrary, when the condition is not met, then an individual will not acquire legal information, since compliance with the legal rule generates an economic loss. This point contains a fundamental implication for the purpose of the process-efficiency analysis: if ascertaining the law is too costly, the substantive efficiency of legal rules vanishes because the excess of compliance-information costs induces people to remain rationally ignorant of the content of the law. This implies that when creating legal rules, the lawmaker should not only consider the efficiency implications deriving from the substantive aspects of the legal rules but also the impact of the rule upon compliance-information costs.

3.1.1.2. The Expected Benefit from Compliance

In the previous section, I emphasized that the positive welfare-effect of output-efficiency might be negated by the magnitude of the compliance information costs. I now turn to the explanation of the causal relationship between the ex ante–ex post choice, and the structure of the compliance information costs. Since this issue is usually discussed in the context of the comparison between rules and standards, I will use the dichotomies ex ante vs. ex posts and rules vs. standards interchangeably.

Assumption 2: There are two types of individuals: Kaplow-type individuals who strive for legal precision and Ayres-type individuals who are satisfied with a rough idea of the content of the law.

---

247 In reality, it is likely that the individual who decides not to acquire legal information “makes a guess” on the content of the legal command. However, for the sake of simplicity, I prefer, assuming that once the individual has decided not to acquire legal information, that he or she will not undertake the course of action established by the law. Thus, I assume only the two following possibilities: (1) acquire → obedience to the law, or (2) not acquire → violation of the law.

248 Whereas rules are characterized by a greater level of detail in advance, with respect to the acts of individuals, standards leave the specification of the content of the law to the ex post law-making process. For the related law and economics literature, see supra note 12.
Whether rules or standards minimize the compliance information costs is an empirical question that depends on a number of different factors. In the law and economics literature, scholars have argued in one direction or the other, with plausible arguments in both cases. Hereinafter, I will indicate the two opposing perspectives (on the cost implications of rules and standards) by referring to the assumptions formulated, respectively, by Kaplow and Ayres.

**Kaplow assumption.** According to the Kaplow approach, rules facilitate compliance with the law because they reduce the costs of becoming informed about the law.\(^{249}\)

**Ayres assumption.** According to the Ayers approach, in many cases, standards facilitate compliance with the law because they reduce the information compliance costs for individuals who do not strive for legal precision. In many cases, individuals are not willing to invest many resources in becoming informed of the content of the law, and they are satisfied with having a less precise (but cheaper) idea of the content of the legal rules.\(^{250}\)

The Kaplow and Ayres assumptions are both plausible, and they do not contradict each other. The difference between the two perspectives is not in the assumption regarding the shape of the compliance-information costs; instead, the difference lies in the assumption...

\(^{249}\) See Kaplow, Rules, supra note 198 at 596: (“Individuals will find it cheaper to learn how rules would apply to their circumstances than to learn how standards would apply, because the former will have already been given content whereas the latter will require predicting the content that a later decision-maker will provide. As a result, rules tend to be preferable with regard to individual behavior, because individuals will expend fewer resources learning about the law and will learn more under rules and thus behave more in accordance with the law.”); Parisi and Vincy Fon, “The Economics” see supra note 242 at 12 puts forward a similar argument: (“In our analysis we take the value of the law as a function of legal precision. Rules advance certainty, consistency, and predictability to private parties and promote judicial economies by minimizing the need for a detailed consideration of facts and circumstances each time a law is applied. […] Given the greater accessibility and predictability of detailed rules, more individuals are likely to become informed in a regime dominated by rules than in a regime dominated by standards. This is a value of law’s specificity. Under rules, individuals are more likely to adjust their conduct to the percepts of the law”).

\(^{250}\) See Ian Ayres, “Preliminary Thought on Optimal Tailoring of Contractual Rules” (1994) Southern California Interdisciplinary Law Journal 8 (“it may be that standards give individuals a less expensive way of gaining a rough idea of the law's content”) [Ayres, “Preliminary]. A similar argument is made by Richard A Posner, How Judges Think, supra note 123 at 177 (“Standards are more likely to conform to lay understandings – which means that despite their greater vagueness they may provide better guidance to compliance with the law”).
of individual preferences on the level of *ex ante* legal precision. Kaplow seems to assume implicitly that *all* individuals have the same (high) incentive to obtain appreciation of the law’s content; in contrast, Ayres recognizes the existence of different utility functions for legal knowledge among the people subject to law.\(^{251}\)

The difficult point here lies in identifying the economic principle to distinguish cases in which people subject to law demand more legal precision from cases in which people are satisfied with rough knowledge of the legal command. In this respect, I argue that jointly read, the Kaplow and Ayres assumptions implicitly establish the following relation between the expected net benefits from information and compliance information costs: *individuals are willing to bear additional compliance information costs to the point at which the marginal expected benefits of increasing the knowledge of the legal rule equal the marginal costs of acquiring an additional amount of information.* Thus, it is from the individual marginal cost and marginal benefit functions that we have to start to understand individual behavioural response to the *ex ante-ex post* choice.

**Proposition 2:** *The higher the expected net benefit value from compliance is, the greater the amount of resources that the individual is willing to use to increase his or her knowledge of the law.*

Here I am interested in understanding the structure of the individual demand for legal specificity (i.e., the amount of legal specificity that individuals demand in each level of expected net benefit from compliance).

\(^{251}\) The implication of Kaplow’s assumption is that individuals who strive for legal precision prefer detailed rules (that provide precise guidance in advance) to standards (that generally give rise to a variety of judicial interpretations). Ayres does not contradict this argument; he simply says that there are many individuals that do not strive for legal precision that are satisfied with more vague, but also more cheaply understandable standards. Put differently, Ayres does not say that standards are less costly to be interpreted compared to detailed rules (which would constitute the contrary of Kaplow’s assumption); he says that in some cases individuals do not have an incentive to bear the cost of having a precise knowledge of the law’s content. In these cases, when the individual is satisfied with a “rough” idea of the legal content, standards might be better because they provide at less cost a “rough” idea of the content of the legal command. Hence, the two arguments do not contradict each other; rather, they are the logical consequence of two different factual assumptions about individuals’ utility function. Kaplow constructs the value of law as a function of legal precision while Ayres constructs the value of legal precision as a function of the value of law. See my considerations on the position taken by Parisi and Vincy Fon, “The Economics” see *supra* note 242.
The condition of incentive compatibility tells us that in order for the individual to comply with the law, the difference between $\delta(b_n)$ and $c_i$ must be $\geq 0$. This condition implies that an individual is willing to bear a greater amount of $c_i$ as $\delta(b_n)$ increases. In other words, the higher the expected net benefit from compliance is, the higher the amount of resources the individual is willing to invest to increase knowledge of legal rules without incurring an economic loss. In contrast, when the value of the expected net benefit from legal compliance is low, then the individual can afford to devote a smaller amount of resources to increase his or her knowledge of the legal rule. This principle is compatible with the (apparently contradictory) assumptions of Kaplow and Ayres. Kaplow’s assumption describes the individual demand for law at high levels of expected benefits, whereas Ayres’ assumption describes the individual demand at low levels of expected benefit.

As we will see, Proposition 2 bears important implications for the choice of the efficient process. However, it does not provide any economic explanation for the individual behaviour at different levels of $b_n$. In the following step, I attempt to provide this explanation.

**Proposition 3:** When the expected net benefit from compliance is low, it is likely that the level of ex ante legal specificity of a detailed rule exceeds the individually optimal level of specificity; in this case, a standard incentivizes legal compliance. When the expected net benefit from compliance is high, it is likely that the level of ex ante legal specificity of a standard lies below the individual’s optimal level of specificity; in this case, a detailed rule incentivizes legal compliance.

Based on the identity $b_i = \delta(b_n) - c_i$, since we assume $\delta$ as constant and exogenous to the model, we can define the expected net benefit from legal information as a function of both the expected net benefit from compliance and the cost of legal information:

$$b_i = f(b_n, c_i) = \delta(b_n) - c_i$$ (8)
As previously noted, I keep $b_n$ as exogenously determined and consider $c_l$ the only dependent variable. Thus, (8) expresses the obvious principle that in order to maximize the benefits from legal information—and to incentivize individual compliance—the cost of legal information should be minimized. Now to make a further explanatory step, let me bring into the analysis the level of precision of the legal rule and then formulate explanatory hypotheses concerning the structure of the individual cost and benefit functions of legal precision.

I assume $c_l$ to be a function of the specificity of the legal rule, which I denote as $\omega$:

$$c_l = g(\omega)$$  \hspace{1cm} (9)

where $c_l' > 0$ and $c_l'' > 0$. The meaning of this assumption is that $c_l$ is an increasing function of $\omega$, with an increasing marginal cost for any additional increment of ex ante legal specificity: that is, as the degree of ex ante legal specificity increases, the cost of interpreting the rule and becoming acquainted with the content of the legal command becomes increasingly higher.\(^{252}\) If we substitute (9) into (8), we obtain $b_l$ as a function of the specificity of the legal rule:

$$b_l = f(g(\omega)) = \delta(b_n) - g(\omega)$$  \hspace{1cm} (10)

where $b_l' > 0$ and $b_l'' < 0$. This means that $b_l$ is an increasing function of $\omega$ with a decreasing marginal benefit function: that is, as the level of ex ante legal precision increases (i.e., the degree of specificity), the cost of interpreting the rule and becoming acquainted with the content of the legal command becomes increasingly higher.\(^{253}\)

\(^{252}\) As I will explain below, the opposite assumption (i.e., less legal precision increases the costs of legal information) is plausible as well, which constitutes a complicating factor for the analysis of the optimal degree of legal precision.

\(^{253}\) The intuition behind this expression is well expressed by the following statement contained in Ayres Preliminary supra note 250, at 8: “[...] it may be that standards give individuals a less expensive way of gaining a rough idea of the law's content. Thus, although the reasonable price default or the lost profit damage default may be more expensive to interpret precisely than a more explicit ex ante rule, it might give market participants a rough approximation of damages cheaply. Anyone who has tried to work
An important guideline for the lawmaker is derived from the foregoing considerations: to maximize the level of compliance with the law, the lawmaker should have to resort to higher levels of *ex ante* specificity (rules) as the individual expected benefits from legal information increases, and to higher levels of *ex post* specificity (standards) when the expected benefit from legal information is low.

### 3.1.1.3. Probability and Frequency per Actor

In this subsection, I introduce two factors influencing the individual’s incentives to acquire legal information: the probability that the individual falls within the scope of application of the law; and the frequency with which individuals undertake the course of action covered by the law.

**Proposition 4**: The probability of falling into the scope of application of the law is a determinant of the expected net benefit from legal information.

The level of $\delta$ is an argument of the function of the expected net benefit from information that emerges from (5). Here I take a further step: I contend that the probability of legal application depends upon the degree to which the legal definition of the regulated behaviour is specified. The more the objective and subjective features of the targeted activity are *ex ante* specified by the law, the greater the probability of falling into the scope of application of the law tends to be concentrated upon a specific set of individuals. Consequently (since, by definition, the level of $b_i$ is an increasing function

---

254 Objective features are those regarding the regulated behavior. Subjective features are those regarding the individual or entity to be subjected to the law.

255 For instance, let me consider tort law or, generally, the regulation of activities producing externalities. In some cases, tort law is addressed to the generality of people, such as the general principles of extra-contractual liability provided by civil codes in many civil law traditions. In other instances, tort law is addressed to a restricted set of risky activities, like for example the regulation of product liability or the regulation of polluting activities. In both cases, the general economic principle behind tort regulation is that of inducing individuals to internalize the external effects or their actions. However, the distribution of probability of law application is different in the two cases. When the regulation is addressed to the vast generality of people subject to law and to an indefinitely large set of activities, then the probability for the single individual to commit an act that is subject to the law is relatively low and is equally distributed among the people. This is the reason why, for example, in civil-law countries the civil code contains a
of $\delta$), when the regulated activity is particularly qualified, the individual expected net benefit from legal information is higher, compared to unqualified activities, so the individual is willing to invest more to acquire legal information. The positive relation between $\delta$ and the individual’s willingness to invest in legal information constitutes the economic reason that, for example, the general principles of tort law are generally formulated through legal standards, while the regulation of qualified risky activities is often formulated through more detailed legal rules. Namely, individuals performing qualified economic activities are willing to invest more resources in understanding the content of the legal rules regulating those activities because their behaviour is more likely to fall within the scope of application of the law.

**Proposition 5:** *Frequency per actor is a determinant of the expected net benefit from legal information.*

Another variable that is particularly important in determining the magnitude of the expected net benefit from legal information is the frequency with which an individual falls within the scope of application of the law, which I call *frequency per actor.*

Since the law applies when the individual undertakes the regulated conduct, frequency per actor can be defined as the number of times an individual undertakes the course of action covered by the law. The higher the frequency per actor is, the higher is the expected net benefit from legal information, consequently, the higher the level of $c_i$ that the individual is willing to sustain to acquire legal information.

Let $b_u$ denote the individual expected net benefit from applying the law in a single case, so $b_n = \lambda b_u$. If I substitute $b_u$ into (6), I obtain

---

[256] This expression is used in Ayres, “Preliminary” supra note 250 at 9 to distinguish aggregate frequency from frequency per actor. Note that frequency per actor is one of the determinants of the behavioral effect, while aggregate frequency (i.e., the total number of cases to which the law is applied) is one of the determinants of the level of average information-cost.
\[ b_i = \delta(\lambda b_{ui}) - c_i \]  

(11)

which expresses the idea that—holding all other variables constant—the higher the frequency per actor is, the higher the expected net benefit from legal information.

3.1.1.4. Conclusions on Individual Legal Demand

I have demonstrated that substantive efficiency is only one part of the complex structure of the incentives for making individual compliance decisions. Other variables that are related to the features of the law-making process combine to determine the expected net benefit from legal compliance. Let me briefly summarize the factors influencing the causal relationship between the ex ante–ex post choice and the decisions of individuals to acquire legal information: (i) the lawmaker can influence the magnitude of compliance-information costs through the choice of the level of ex ante legal specificity of legal rules; (ii) the individual demand for ex ante legal specificity is a function of the expected net benefit from legal compliance; (iii) when choosing the level of specificity of the legal rule, the lawmaker should respect the condition of incentive compatibility; an excess of compliance-information costs—caused by an ex ante excess of legal specificity—incentivizes individuals to remain rationally ignorant about the content of legal rules, despite their substantive efficiency; (iv) two important determinants of the expected net benefit from legal information are the probability of the law’s application (probability that the undertaken conduct falls within the scope of the law) and the frequency per actor (the number of times the individual undertakes the course of action covered by the law).

3.1.2. The Supply of Ex-Ante Legal Specificity

In this section, I discuss the issue of the lawmaker’s optimal informational effort in determining the content of the law. It is analytically convenient to break the lawmaker’s optimization problem into two parts. First, the lawmaker engages in the problem of how to minimize the costs of producing any given level of ex ante legal specificity (i.e., cost-minimization problem). Second, once the cost-minimizing technique for each level of production has been identified, the lawmaker chooses the level of ex ante legal specificity
to be produced in order to maximize the number of people who are willing to abide by the law. In subsection 3.1.2.1, I describe the cost-function of the lawmaker; in subsection 3.1.2.2, I examine the problem of identifying the cost-minimizing law-making technique. Subsequently, in section 3.1.3., I discuss the problem of maximizing legal compliance.

In general, the main cost-advantage of *ex ante* law-making is the presence of economies of scale. Hence, a crucial analytical point concerns identifying the structural conditions that generate scale effects in law-making activity. In this respect, the existing literature has long recognized that the law-making process is characterized by significant economies of scale when the regulated behaviour has the characteristics of *frequency* and *homogeneity*. I now turn to these two features of the demand-side to provide some

---

257 The importance of frequency and homogeneity as conditions for the economies of scale effect have already been pointed out by Werner Z. Hirsch, “Reducing Law’s Uncertainty and Complexity” (1973) 21 UCLA L Rev 1233 at 1241 [Hirsch, “Reducing”] (“The factors germane to a comparison of benefits and costs can be related to the question of scale economies. The greater the complexity, diversity, and changing conditions are, the fewer the cases where the law would likely to be applicable. Hence, smaller groups, i.e., courts or private groups through contracts, may be more efficient than universal statutes. Within such a framework one might want to focus on the expected uniformity of the situation (both in occurrence and determination of occurrence), number of repetitions of the situation, and relative costs of having smaller rather than larger groups act [my emphasis].”)

With specific respect to legal frequency, Hirsch underlines its importance in choosing the level of ex ante legal precision, ibidem: “Cases that have a high probability of occurring frequently and of being fairly uniform would warrant the government's providing statutes that are relatively precise and detailed. Cases with a low probability of occurrence would rely on vague laws—often contracts—or no laws at all [my emphasis]”. More recently, Kaplow pointed out the importance of frequency of law’s application in determining the structure of the law-making costs.

Kaplow, Rules, *supra* note 198, at at 573 (“[...] one must take into account the frequency with which the two types of costs will be incurred. Promulgation costs are incurred once. In contrast, enforcement costs may be incurred repeatedly or never. 5 On the one hand, a law may apply to an activity that is undertaken by many individuals: some federal income tax provisions apply to millions of individuals and billions of trans- actions.' In such instances, rules tend to be preferable. Even if the promulgation cost differential significantly exceeds the enforcement cost differential in applying standards, rules may be much cheaper.’ On the other hand, a law —or, as is often relevant, a particular component of a law, possibly a highly detailed one—may have a small likelihood of applying to any activity; consider the example of myriad unique accident scenarios. Then, standards tend to be preferable. Even if they are extremely costly to apply, the significant likelihood that the particular application will never arise may make standards much cheaper”); at 577 (“[...] the greater the frequency with which a legal command will apply, the more desirable rules tend to be relative to standards. This arises because promulgation costs are borne only once, whereas efforts to comply with and action to enforce the law may occur rarely or often. Rules cost more to promulgate, standards cost more to enforce.’); at 585: (“Whether the ideal time to acquire and disseminate information is ex ante or ex post depends, most importantly, on the frequency with which the information will be used. The savings from a single ex ante investigation will be great when the use of the results will be frequent, but will be negligible when the use of the results will be unlikely. In addition, to the extent there are economies of scale in information acquisition, ex ante wholesale investments may be
preliminary guidelines for the optimal choice between the *ex-ante* and *ex-post* creation of law.

### 3.1.2.1. Cost-Functions of the Lawmaker

In this subsection, I articulate the discussion around four key propositions concerning, respectively: i) the economies of scale effect in the case of homogeneity, ii) the relation between cost-elasticity and *ex ante* legal specificity, iii) the cost-curves of *ex post* law-making, and iv) the impact of heterogeneity on the cost-curves.

**Assumption 1**: *the law-making production-information costs are a function of the level of specificity of the legal rules and the number of cases to be regulated.*

Before discussing the structure of the law-making costs, let me clarify some assumptions that I believe well describe a large percentage of actual law-making cases. First, I assume that the amount of production-information costs borne by the lawmaker is a function of the level of detail embedded in the legal norm and of the number of cases to be regulated:

\[
    c_p = f(\omega, N)
\]  

(12)

superior. However, if there are advantages in delay because information will be easier to acquire at the time individuals act or cases are adjudicated, ex post investments would tend to be preferable; and finally at 621 (“The central factor influencing the desirability of rules and standards is the frequency with which a law will govern conduct. If conduct will be frequent, the additional costs of designing rules that are borne once are likely to be exceeded by the savings realized each time the rule is applied. Thus, rules involve a wholesale approach to an information problem, that of determining the law's appropriate content. Standards instead require adjudicators to undertake this effort, which may have to be done repeatedly (unless the standard is transformed into a rule through precedent)”.

Regarding legal heterogeneity, see Kaplow, Rules, *supra* note 198, at 595: (“When one makes a single pronouncement that will govern many (perhaps millions) of cases, it is worthwhile to undertake greater investigation into the relevance of additional factors and to expend more effort fine-tuning the weight accorded to each. Thus, when rules are to be applicable to frequent behavior with recurring characteristics, there is a systematic tendency for rule systems to be more complex than the content that would actually be given to standards covering the same activity. In contrast, when the behavior to be regulated by law is infrequent, or when each instance (no one very likely to occur) is unique in important ways, substantial ex ante analysis for each conceivable contingency would be a poor investment, whereas ex post determinations under standards are made with the knowledge that the scenario has indeed arisen” [emphasis is mine]); Ibidem at 621: “Of particular relevance are laws for which behavior varies greatly, so that most relevant scenarios are unlikely ever to occur. Determining the appropriate content of the law for all such contingencies would be expensive, and most of the expense would be wasted. It would be preferable to wait until particular circumstances arise”.
where $c_p$ represents the production-information costs, and $N$ represents the magnitude of aggregate frequency.\textsuperscript{258} Second, each level of $\omega$ corresponds with a given level of $c_p$ determined by the sum of the costs of \textit{ex ante} and \textit{ex post} legal specificity. I call the cost of producing law \textit{ex ante} “promulgation-information costs” $p$, and the cost of specifying the legal command \textit{ex post} (on a case-by-case basis) as “adjudication-information costs” $j$:

$$c_p = p + j$$  \hspace{1cm} (13)

From the economic standpoint, the latter distinction is important because while promulgation costs are \textit{fixed}, adjudication costs are \textit{variable}.\textsuperscript{259} Based on the existing literature\textsuperscript{260}, I assume that both $p$ and $j$ are a function of the legal specificity of the legal command, denoted as $P(\omega)$, and $J (\omega, N)$, respectively. Because $J$ describes the variable component of the law-making costs, it is also a function of $N$. Thus, (17) can be rewritten as follows:

$$c_p = f(\omega, N) = P(\omega), + J (\omega, N)$$  \hspace{1cm} (14)

Because it is obvious that a greater level of detail requires more information, it is straightforward to assume that promulgation costs are an increasing function of the \textit{ex ante} legal precision of the legal command ($p > 0$). More problematic is the assumption

\textsuperscript{258} As I explained in note 256, the expression “aggregate” frequency of law indicates the total number of cases in which a legal rule is applied to the generality of the people subject to law. It distinguishes from the frequency “per actor” that indicates the frequency of law’s application to a single individual. In my model, aggregate frequency influences the cost function of the lawmaker (supply side), while frequency per actor is relevant for the analysis of the benefit function of the individual in purchasing legal information and applying the law (demand side).

\textsuperscript{259} On this point, I follow Parisi and Vincy Fon, “The Economics” see supra note 242 at 13, which distinguishes between “fixed promulgation costs” and “variable adjudication costs”. However, while Parisi and Vincy Fon seem to assume implicitly that fixed promulgation cost are borne by legislatures and variable adjudication costs by courts, I prefer to connect the typology of production information costs only to the moment of production (ex ante/ex post) and not to the type of institution producing the law. In practical terms, this means that, in my model, promulgation costs include not only legislative promulgation costs but also the costs borne by the courts when introducing a new principle of law (i.e. overruling) and adopting a forward-looking perspective. The reason for this different use of the category of costs lies in a different level of abstraction of the analysis.

\textsuperscript{260} See supra note 257.
concerning the behaviour of the variable component of the law-making costs, which depends upon the degree of homogeneity of the legal demand and on the shape of the marginal cost function. Let me clarify this point:

**Proposition 1**: In cases of homogeneity of the legal demand, the ex-ante law-making activity is characterized by significant economies of scale; both the marginal cost and the average variable costs are decreasing functions of ex ante legal specificity and aggregate legal frequency.

A fundamental principle of the economics of law-making is that the frequency of the regulated behaviour affects the variable component of the cost function. Precisely, when the frequency of legal application increases, the average production information cost decreases (*economies of scale effect*). This principle is based on the observation that in many cases, law-making activity entails high fixed costs and, to the extent that the regulated behaviour occurs frequently, low (and sometimes decreasing) marginal costs. The economic implication of this characteristic of the *ex ante* law-making is best explained by comparing the cost curves here described to the more traditional U-shaped representation of the production cost curve.

Microeconomic theory usually depicts the average total cost curve of the firm (hereinafter, “average cost curve”) as initially sloping down until the point of intersection with the marginal cost curve. The intersection between the average and the marginal cost curves corresponds to the cost-minimizing output level at which the economies of scale

---

261 This point was first discussed by Kaplow, Rules, supra note 198. From a somewhat similar perspective, Parisi and Vincy Fon, “The Economics” (see supra note 242) has provided a discussion of the frequency of the application of the law as one of the determinants of the optimal degree of specificity of legal rules. Precisely, Parisi and Vincy Fon build a model in which the adjudication costs constitute the variable component of the law-making costs; in this perspective, the aggregate frequency of the application of the law is one of the factors explaining economies of scale in adjudication.

262 The average and the margin are mathematically related. For a textbook discussion of the relation between marginal costs and average costs in the context of the theory of the firm. See Hal R Varian, *Intermediate Microeconomics. A Modern Approach* 8th ed. (New York: W. W. Norton and Company, 2010) 378-382. The law-making process does not constitute an exception to this principle: the shape of the average law-making cost-curve is determined by the trend of the marginal cost-curve. However, unlike the most common representation of the production cost-function as a U-shaped curve, in the case of homogeneous demand for law, the law-making cost curve has an increasing negative slope along the entire interval of production.
effect is optimized (Figure 7). Once the point of intersection between the average and marginal cost curves has been reached, the average costs start to exceed the marginal costs, and further increments in the scale of production increase the average cost of production (i.e., *diseconomies of scale*).

**Figure 7 U-Shaped Cost Curves**

The economic explanation of this U-shape cost curve lies in the twofold assumption of (i) **initial high-fixed costs** and (ii) **diminishing returns of productive factors**. Under this assumption, the economy of scale effect endures until the point at which the average-cost curve is intersected from below by the marginal cost-curve. In brief, in the case of U-shaped cost-curves, the *economies of scale occur to the extent that the average-cost curve lies above the marginal-cost curve.*

Characterizing the law-making cost-functions in the cases of the homogeneous demand for law is the different shape of the marginal-product curve, which, in turn, determines a different shape of the average-product curve. The explanation for the different shape of the law-making marginal cost curve is that when the demand for law is homogeneous, the law of diminishing marginal returns does not apply to the productive-factors involved in

---

263 I borrow Figure 7 from professor Martin J Osbourne, *Tutorial on Theory of the Firm and Industry* at [www.economics.utoronto.ca/osborne/2x3/tutorial/index.html](http://www.economics.utoronto.ca/osborne/2x3/tutorial/index.html)
the law-making process. In this respect, from the economic standpoint, I cannot overemphasize the importance of the informational nature of the productive factors used in the law-making process. The adjudication activity benefits from the experience accumulated in the application of legal rules to previous similar cases. The accumulated information produces an increasing stock of legal information that enhances the marginal productivity of law-making. Therefore, it is reasonable to assume that to the extent that legal cases are relatively similar and do not require a complex case-specific analysis (homogeneity), the curve of the marginal adjudication cost is a decreasing function of ex ante legal specificity \((J' < 0)\). In brief, in contrast to the production process as usually described by micro-economic textbooks, the law-making process is characterized by increasing returns and, relatedly, by a decreasing marginal cost curve. The promulgation of detailed rules increases fixed promulgation costs; however, when higher initial fixed promulgation costs have been borne, rules are applicable to an indefinite number of cases at a lower informational adjudication cost. Thus, as ex ante legal specificity increases, the level of promulgation information costs becomes higher, while the level of adjudication information costs becomes lower. Conversely, when ex ante specificity decreases, the level of promulgation costs becomes lower, and the level of adjudication information costs becomes higher. Figure 8 shows the difference between the more common U-shaped cost-curves and the downward-sloping law-making cost-curves. Unlike the curves depicted in Figure 7, here the average cost curve does not

---

264 The idea that the body of legal precedents can be analogized to a capital-stock helping the resolution of new cases has been first introduced by William M Landes and Richard A Posner “Legal Precedent: A Theoretical and Empirical Analysis” (1976) Journal of Law and Economics 249.

265 In reality, some costs are independent of the (increasing) legal knowledge of the adjudicator, and from the complexity of the case. In fact, even with the assumption that the marginal adjudication information cost approximates zero, each time the law is applied, some costs (i.e., the administrative costs of the trial) are fixed and do not benefit in the short run from economies of scale. However, at this level of abstraction, these costs might be considered elastic to the degree of homogeneity; for example, it is plausible that the adjudication process is organized by lawmaker in such a way to minimizes the variable (non-informational) adjudication costs for those cases which are frequent and homogenous. If this assumption is plausible, these non-information costs can be temporarily left out of the picture, without affecting the validity of the discussion.

266 Both assumptions are consistent with Parisi and Vincy Fon, “The Economics” see supra note 242 at 13-14.
intersect the marginal cost curve, and the scale effect endures for the whole interval of law production.\(^{267}\)

**Figure 8 The Marginal and Average Production-Information Costs Curves**

Having clarified that the homogeneity of the demand for law allows for significant economies of scale effect, the next step is describing the relationship between cost-elasticity and \textit{ex ante} legal specificity.

**Proposition 2**: In cases of homogeneity, as the level of \((\omega)\) increases, the economies of scale effect associated with \textit{ex ante} law-making becomes increasingly significant.

Thus far, I have assumed that the level of specificity as constant. I now investigate the behaviour of the average cost-function as \(\omega\) varies, while holding \(N\) constant. The

\(^{267}\) Let me denote the average production information cost as \(c_{ap} = \frac{c_P}{N}\). By substituting (18) into the denominator of the latter expression, the average cost-function of the lawmaker can be rewritten as \(c_{ap} = \frac{P(\omega) + f(\omega, N)}{N}\). This algebraic representation of the average production-information costs clarify that for each level of \(\omega\), as \(N\) increases \(c_{ap}\) becomes lower. In fact, holding \(\omega\) constant, as \(N\) increases, the numerator increases by a decreasing rate, while the denominator is constantly increasing; thus, the value of \(c_{ap}\) decreases. In other words, the average production-cost of legal complexity decreases with an increase of the number of cases covered by the norm.
anticipated conclusion is that the more that the level of specificity increases, the more significant becomes the economies of scale effect. To understand this point, I introduce the concept of cost elasticity with respect to aggregate frequency, which is defined as the ratio of the relative variation of $c_p$ and the relative variation of $N$:

$$E_c = \frac{\Delta c_p}{c_p} \div \frac{\Delta N}{N}$$  \hspace{1cm} (15)$$

If we substitute (17) into (18), the definition of cost-elasticity can be rewritten as follows:

$$E_c = \frac{\Delta [P(\omega) + f(\omega, N)]}{P(\omega) + f(\omega, N)} \div \frac{\Delta N}{N}$$  \hspace{1cm} (16)$$

Thus, as $\omega$ increases, the numerator (i.e., relative variation of $c_p$) decreases, so $E_c$ becomes smaller. Hence, given $\frac{\Delta N}{N}$, as the level of specificity of the legal command increases, the elasticity of the production-information costs with respect to the frequency of application of the legal rules increases (i.e., approaches zero). The underlying intuition is that the more specific the rule is, the more its cost function is elastic with respect to an increase in the frequency of its application and consequently the more significant the economies of scale effect is.

**Proposition 3**: In cases of homogeneity, the ex post production of law is likely to increase the total amount of production information costs for each level of $\omega$, when compared to the ex ante production process (diseconomies of scale). The increase in the total production information costs is the result of two distinct effects: (1) the magnitude of variable costs increase because part of the law-making activity is performed ex post rather than ex ante (temporal effect); (2) part of the law-making activity is repeatedly performed so that fixed cost increases (repetition effect).

Let me now turn to the ex post law-making process, which is well represented by the case of regulation through standards. In particular, I want to investigate how the allocations of a given level of legal specificity to the ex post dimension changes the structure of the
production-information costs. To clarify, both the level of specificity of the legal norm and the assumption of legal homogeneity of the demand for law hold constant; however, the allocation of the legal specificity changes from the ex ante to the ex post regulation. I demonstrate that the shift to the ex post perspective produces two distinct effects: one effect is on the variable component of the production-information cost, and the other effect is on the fixed component.

The first obvious consideration is that the ex post law-making entails a lower level of promulgation costs and a higher level of adjudication costs. However, this description is too simplistic because it does not capture the complexity of the ex ante–ex post choice, which derives from the fact that the change in the timing of the production of law coincides with a change in the production technique. In addition to a different allocation of the law-making costs, the ex post perspective determines a change in the structure of the production-information costs. In particular, the shift to an ex post dimension has a twofold effect on production-information costs: (i) it “transforms” part of the fixed promulgation costs into variable adjudication costs, and (ii) it increases the total amount of production information costs. Let me explain these two points.

Ex-ante law-making produces legal principles that are applicable to the whole class of cases covered by the legal rule; therefore, the (fixed) costs of creating legal principles are borne only once. In contrast, from an ex post perspective, each adjudicated case is somewhat different from previous cases; in addition, in the adjudication stage, the individuals operating within the adjudicating arena tend to change case by case. In this context, the adjudicating authorities often replicate part of the law-making activity (with the associated fixed costs) required to establish legal principles that are common to different cases. In other words, an effect of the shift to the ex post perspective is that part of the law-making activity is repeated many times, instead of being performed only once. The frequency of repetition may vary depending on many factors; however,

\[268 I assume here that there is no stare decisis and that courts are not bound by the previous cases.\]

\[269 On this point see Kaplow, Rules, supra note 198 at 596: ([...] there may be economies of scale in making the inquiries at the promulgation stage: Determining the appropriate treatment for each of the possible acts may cost little more than for one act (because the same investigation may yield most of the\]
the relevant point is that the \textit{ex post} law-making reiterates parts of the fixed costs that from the \textit{ex ante} perspective are borne only once. In this sense, I argue that a part of the fixed costs takes the form of variable adjudication costs.\footnote{\textsuperscript{271}} Therefore, the curve of the average total cost does \textit{not} decrease at the same rate as it does in the \textit{ex ante} dimension.

In the context of \textit{ex post} law-making, the slope of the average total cost curve depends upon the magnitude of the fixed costs that are “transformed” into \textit{ex post} variable adjudication costs. If judges “start from zero” each time they consider new cases, the cost of establishing a legal principle for a class of homogeneous cases multiplies by the number of legal cases. Hence, for each level of $\omega$, the total amount of $p$ increases by an amount equal to the cost of the activity \textit{repeatedly} performed multiplied by $N$. In this hypothetical situation, as $N$ increases, $P(\omega)$ also increases, and a diseconomy of scale occurs. However, since it is unrealistic to assume that the adjudication authority starts each time from zero,\footnote{\textsuperscript{272}} let me suppose that the number of cases in which the adjudicating authority starts from zero is \textit{variable}. Given a certain value of $N$, let $\phi$ denote the percentage of cases in which judges depart from previous decisions by reconsidering the case from zero (i.e., the percentage of variability), with $(1 - \phi)$ as the percentage of cases in which authorities adhere to previous decisions. Under these assumptions, fixed costs $p$
increase of $\phi Np$. This means that, in the case of the \textit{ex post} production of law, the $\phi Np$ amount of production-information costs changes its nature and is "\textit{transformed}" into a variable component of the total production-information cost function. In summary, compared to \textit{ex ante} law-making, the \textit{ex post} production of law entails an increase of fixed costs equal to the amount of fixed costs ($\phi Np$) that are repeatedly borne over the \textit{ex post} stage of the production process and that would have not been borne from an \textit{ex ante} perspective.

In conclusion, the allocation of a given amount of legal specificity to the \textit{ex post} perspective generates two related effects: (1) part of the law-making activity shifts from the \textit{ex ante} to the \textit{ex post} dimension thereby increasing the \textit{ex post} adjudication costs; and (2) part of the law-making activity is repeated many times thereby increasing (and partially re-allocating \textit{ex post}) the fixed costs. Henceforth, I will refer to these two effects respectively as the "\textit{temporal effect}" and the "\textit{repetition effect}". The important conclusion to bear in mind is that the total effect of shifting part of the law-making efforts to the \textit{ex post} perspective results in an increase of the production-information costs by an amount equal to $\phi Np$.\footnote{I have assumed that in $\phi N$ cases, the adjudicating authority starts from zero: that is, the adjudicating authority considers the case in all the elements that bear legal relevance. A more sophisticated analysis could introduce to the model the fact that courts might consider only some aspects of the legally relevant cases, depending on the strength of the path-dependence mechanisms that are at work in the adjudicating phase. One method for considering this most complicated scenario might be that of measuring the average value of the percentage of the legal relevant elements that are reconsidered case by case. This average value could be considered the measure of the strength of path-dependence mechanisms. Additionally, if we denote this average value as, say, $x$, then, the value of $\Delta c_p$ determined by the shift \textit{ex post} of the law-making activity would be $x\phi Np$, where $\phi$ is the percentage of cases that are reconsidered, and $x$ is the average percentage of the elements that are reconsidered.}

\textbf{Proposition 4}: \textit{In cases of heterogeneity, no economies of scale effects occur, and the \textit{ex post} production of law benefit from an informational advantage.}

As we have previously seen, the relative advantage of \textit{ex ante} law-making in cases of homogeneous legal demand derives from the decreasing marginal cost curve (which reflects the increasing returns of legal information) that determines a decreasing average cost curve. The heterogeneity of legal cases precludes the \textit{ex ante} generalization of legal principles and prevents the possibility of spreading fixed costs over a sequence of
homogeneous cases. Consequently, when the demand for law is heterogeneous, no scale effects occur. Furthermore, because heterogeneity is unpredictable, ex post law-making (which is case-specific by nature) benefits from an informational advantage.

3.1.2.2. The Cost-Minimizing Supply of Law

At this point, it is useful to summarize the foregoing discussion about the ex ante–ex post choice through the following synthesizing propositions.274

*Ex Ante Production of Law*

(i) In the case of the frequency and homogeneity of cases covered by the law, ex ante law-making benefits from economies of scale; (ii) frequency affects the average production information costs; (iii) homogeneity affects the shape of marginal production information-costs and amplifies the economies of scale effect; and (iv) heterogeneity makes the marginal informational-adjudication costs vary case by case, so there is no scale effect.

*Ex Post Production of Law*

(i) In the case of homogeneity, ex post law-making suffers from diseconomies of scale because of the “repetition effect”; (ii) when facts are heterogeneous, ex post law-making benefits from of an efficiency advantage relative to the ex ante perspective because of the reduced fixed costs of law-making (i.e., the magnitude of the “repetition effect” deriving from the shift towards the ex post law-making is smaller than the magnitude of the effect of heterogeneity on ex ante fixed costs).

I now want to operationalize these economic principles and draw some practical guidelines for identifying the cost-minimizing institutional response to the demand for law. First, I identify four different configurations of the demand for law, depending upon

---

274 Recall that the discussion on production information costs is based on the assumption that rules and standards are compared by holding constant the level of precision of the legal rules.
different levels of homogeneity and frequency (Figure 9). I then discuss each possible scenario from the perspective of the *ex ante-ex post* choice (Figure 10).

**Figure 9 Four Types of the Demand for Law**

<table>
<thead>
<tr>
<th></th>
<th>High Frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Homogeneity</td>
<td>Case (1)</td>
<td>Case (3)</td>
</tr>
<tr>
<td></td>
<td>High Frequency</td>
<td>Low Frequency</td>
</tr>
<tr>
<td></td>
<td>High Homogeneity</td>
<td>High Homogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Homogeneity (Heterogeneity)</td>
<td>Case (4)</td>
<td>Case (2)</td>
</tr>
<tr>
<td></td>
<td>High Frequency</td>
<td>Low Frequency</td>
</tr>
<tr>
<td></td>
<td>Low Homogeneity</td>
<td>Low Homogeneity</td>
</tr>
</tbody>
</table>

275 In approaching this discussion, the reader should bear in mind two assumptions. First, I treat homogeneity and frequency as an all-or-nothing issue, rather than a matter of degree. This assumption will be relaxed in the next sections when I will attempt to capture more of the complex relationship between the supply and demand of legal rules and the production-information costs. Second, I consider only a single feature of the supply (*ex ante-ex post*). Subsequently, I will turn back to this chart and bring into the picture the choice between centralized and decentralized law-making processes.
Case 1: High Frequency, High Homogeneity (High-High)

The high level of frequency, coupled with the high homogeneity of prospective legal cases, militates toward the supply of a high level of *ex ante* legal specificity.

Case 2: Low Frequency, Low Homogeneity (Low-Low)

The low levels of both frequency and homogeneity negate any scale effects. The high fixed costs of producing *ex ante* precise legal rules cannot be distributed over sufficiently numerous sequences of legal cases for two reasons: cases are not numerous enough (low frequency), and cases are not homogeneous enough to be included in a sufficiently precise legal classification.

Case 3: Low Frequency, High Homogeneity (Low-High)

The homogeneity of the legal demand allows regulating prospective legal cases through a high level of *ex ante* legal precision in order to exploit the scale effect. However, the low level of frequency negates the potential economizing effect of homogeneity and militates towards the *ex post* creation of law. When legal rules are applied to a small number of cases, *ex ante* regulation would be inefficient because of the high fixed cost associated with the high level of precision of the legal rules. In brief, in this type of scenario, the high level of homogeneity calls for an *ex ante* production of law, while the low frequency of the future cases calls for *ex post* law-making. Simply put, despite the homogeneity of the legal demand, the low level of frequency impairs the potential economizing effect of the *ex ante* production of law.

Case 4: High Frequency, Low Homogeneity (High-Low)

The high level of frequency would call for an *ex ante* production process in order to distribute the high fixed costs of *ex ante* legal precision over a high number of legal cases. However, given the heterogeneity of the cases, an *ex-ante* law-making process would not benefit from economies of scale because the low degree of homogeneity would preclude the lawmaker from regulating the prospective legal cases through a sufficiently
precise\textsuperscript{276} legal rule. In summary, the high frequency of law application would call for the \textit{ex ante} production of law, whereas the heterogeneity of the facts of law would call for \textit{ex post} law-making. Under these circumstances, technical efficiency would result from the relative magnitude of the contrasting effects of high frequency and low homogeneity. As in case 3, in this case, homogeneity and frequency run in opposing directions; however, unlike case 3, in this case, the low level of homogeneity hampers the potential economizing effect of the \textit{ex ante} production of law.

Figure 10 The Cost-Minimizing Supply of Law

<table>
<thead>
<tr>
<th></th>
<th>High Frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Homogeneity</td>
<td>(\text{1) Ex ante}) Economies of scale</td>
<td>(\text{3) ?}) No economies of scale because of low frequency</td>
</tr>
<tr>
<td>Low Homogeneity</td>
<td>(\text{4) ?}) No economies of scale because of low homogeneity</td>
<td>(\text{2) Ex post}) (No economies of scale).</td>
</tr>
</tbody>
</table>

The chart shown in Figure 10 summarizes the economic findings of cost-minimization. We know that the frequency and homogeneity of the demand of law are two conditions for the operation of the economy of scale effect. In cases 1 and 2, the cost effects of frequency and homogeneity run in the same direction, so it is easy to draw a practical conclusion regarding the best production technique. In cases 3 and 4, the effects of frequency and homogeneity counterbalance each other; there is no straightforward implication for the cost-minimizing supply of law. Thus, in the latter two cases, technical efficiency depends on the relative magnitude of these two effects.

\textsuperscript{276} To clarify, the expression “sufficiently precise” refers to the magnitude of the scale of effect triggered by the level of \textit{ex ante} complexity of the rule.
Although these findings are a useful starting point, they provide a limited understanding of the choice of the lawmaker for two reasons. First, they leave substantially unanswered the question of technical efficiency in cases 3 and 4. Second, these findings provide insights only on the supply side of the cost effect; they leave the demand side of the cost effect (and the related behavioural effect) out of the picture. In other words, they concern only the cost-minimizing aspect of the lawmaker optimization problem, and they do not provide any indication of the degree of specificity that the lawmaker should produce in order to maximize its objective function. In the next sections, I will attempt to overcome the limitations in the predictive capability of the model by moving in two directions. In section 3.1.3, I will consider the second part of the law-making problem, which is the maximization of legal compliance. In sections 3.1.4 and 3.1.5, I will relax the assumption of frequency and homogeneity as an all-or-nothing issue, instead, I will treat them both as a matter of degree, and I will attempt to advance the economic understanding of these two important features of legal demand.

3.1.3. The Maximization of Legal Compliance

Having discussed the cost-minimizing problem, I extend the analysis to the choice of the social welfare-maximizing level of ex ante legal specificity. Because the maximization problem confronted by the lawmaker shares some similarities with those of the monopolist seeking to maximize its economic profits, I will provide a discussion of the maximization problem according to the theory of monopoly.

---

277 As previously assumed, the lawmaker maximizes social welfare by maximizing the number of people abiding by the rules (i.e., choosing the compliance-maximizing level of ex ante legal complexity).

278 As the monopolist, the lawmaker faces two different constraints: the technological constraint summarized by the cost function (telling us the total cost of producing any quantity of \( \omega \)), and the economic constraint represented by the demand for law (telling us how much legal specificity individuals are willing to purchase). We know from economic theory that unlike the competitive firm facing a flat demand curve, the monopolist’s profit maximization strategy is characterized by a downward-sloping demand curve. Additionally, we know that the elasticity of the market demand curve in a monopolistic market with respect to the output price, determines the power of the monopolist to influence the price and the quantity of the output-units sold in the market. In particular, the elasticity of the demand curve determines the trade-off between two contrasting effects of the price choice: the “quantity effect” and the “price effect”. For instance, when the monopolist decides to increase the output-price, two effects take place; on one hand, as an effect of the increased marginal revenues on all the units sold, the total revenue raises (price effects); on the other hand, as a consequence of the increase of the price, the amount of units sold decreases, which has a negative effect on the total revenue (quantity effect). As explained in the text, things work in part differently for the lawmaker.
As defined here, social welfare is equal to the vertical sum\textsuperscript{279} of individual utility functions \textit{minus} the cost of producing the law. Based on this definition, the lawmaker’s problem can be described in terms of the maximization of the difference between the sum of individual utilities and the total production-information costs:

\[ \max W = \sum_{i=1}^{n} [b_i]_i - c_P \]  

(17)

where \( I \) represents the individual subject to the law. If we substitute (10) and (14) into (17), we derive a definition of the maximization problem as a function of the level of specificity of legal rules:

\[ \max_{\omega} W = \sum_{i=1}^{n} [\delta(b_i) - g(\omega)]_i - [P(\omega) + J(\omega)] \]  

(18)

The first term represents the sum of individual utilities expressed as a function of \textit{ex ante} legal specificity. The second term represents the total production-information cost for each level of specificity. Based on this definition of the lawmaker’s objective function, we can easily recognize the effects of the choice of the level of \( \omega \) on \( W \). It is convenient to analyze this point by starting from the simplest scenario and then add complicating factors.

Let me start by assuming that the demand for law is highly homogeneous and frequent.\textsuperscript{280} For a moment, let me assume also that all the individuals subject to law have a downward-sloping individual demand function and that all individuals who decide to remain ignorant do not obey the law. Under these simplifying assumptions, the lawmaker’s decision to increase the level of the \textit{ex ante} legal specificity of an amount equal to \( \Delta \omega \) generates two effects. First, there is a cost effect on the supply side: the total cost function of the lawmaker increases by the amount of \( \frac{\delta(X_{\omega}N)}{\delta\omega} \), which represents the

\textsuperscript{279} The summation is vertical: because law is a public good characterized by non-rival consumption, the social benefit from a legal rule is the sum of the marginal benefit to each individual belonging to the community to which the law applies. Therefore, to obtain aggregate welfare, we need to use a vertical summation of the individual marginal benefits for each level of ex ante legal specificity. See \textit{supra} note 145.

\textsuperscript{280} I also assume a short-term perspective, so the fixed costs of law-making are constant.
additional cost of the increment of specificity. Second, an increase in the level of specificity of $\Delta \omega$ produces consequences for the aggregate level of utility. This second effect is more complex than the first is, and it might be divided into two distinct components: i) a cost effect on the demand side and ii) a related behavioural effect. It is worth elaborating these two demand-side effects because they illuminate the peculiarity of the monopolistic position of the lawmaker.

Let us take a closer look at the algebraic expression representing the aggregate benefit of compliance $\sum_{i=1}^{n} [\delta(b_i) - g(\omega)]$, and read it through the lens of our previous discussion about the optimizing behaviour of individuals. The first obvious consequence for the demand side of an increase in the value of $\omega$ is that each individual cost function changes from $g(\omega)$ to $g(\omega + \Delta \omega)$, so that $W$ decreases by the amount of

$$\sum_{i=1}^{n}[g(\omega + \Delta \omega) - g(\omega)]$$

which represents the demand side of the cost effect. The second effect on aggregate utility is represented by a diminishing value of $n$ by an amount equal to $\Delta n$; if the individual expected net benefit from compliance decreases, then the number of people that decide to comply with the law also decreases, and vice versa (i.e., behavioural effect). The magnitude of this effect depends on many factors. Under the assumption made here, some individuals will find that the increased level of specificity is too costly and that they cannot afford the increased amount of compliance information costs; consequently, they will prefer not to acquire legal information, remain rationally ignorant, and not obey the law. Thus, as an effect of an increase of the individual incentive to comply, the number of persons that comply with the law decreases by an amount equal to $\Delta n$; furthermore, the decreasing value of $n$ reinforces the above effect on the social welfare function deriving from an increase in $g(\omega)$.

Algebraically, an increase of $\Delta \omega$ in the level of legal specificity decreases the level of $W$ by the amount of
\[ \sum_{i=1}^{\Delta n} \left( \delta(b_i) - [g(\omega + \Delta \omega) - g(\omega)] \right) = \left[ P(\omega) + J(\omega) \right] \] (20)

which represents the loss of welfare determined by sum of the cost effects on the supply side, the cost effect on the demand side, and the behavioural effect. In brief, an increase in the level of \textit{ex ante} legal specificity produces two effects on the demand side: a) it reduces the individual expected net benefit from compliance of the whole set of individuals subject to the law, and b) shrinks the same set of individuals as a consequence of the behavioural impact of the modified incentives to comply. In this respect, we can see the peculiarity of the law production monopoly. Whereas, in the conventional model, the monopolist faces a trade-off between price effect and quantity effect, under the above restrictive assumptions\footnote{To clarify, I refer to the following two assumptions: (i) all the people subject to law have a downward-sloping individual demand function, and (ii) all individuals who decide to remain ignorant do not obey the law. These assumptions parallel those of the conventional model of the monopoly. In fact, in the simplest version of the standard model of monopolistic market, all the consumers have the same preferences, and the market demand is such that an increase in the output-price causes some consumers to not to buy the product. It is worth pointing out that the peculiarity of the situation of the lawmaker does not derive from the assumption of my model; but, as I explain in the text, it is due to the nature of the lawmaker function.}, the lawmaker does not face a similar trade-off because the two effects on the demand side run in the same direction with respect to the objective function of the lawmaker.

If I introduce into the model the possibility of both Kaplow and Ayres types of individual preferences, the effect on the demand side of an increase in the \textit{ex ante} legal specificity may move in two possible, opposite directions. For those individuals whose preferences are of the type described by the Kaplow assumption, an increase in the level of \textit{ex ante} legal specificity increases the net benefit from compliance and consequently \textit{increases} the incentives to comply with the law. On the contrary, in the case of the Ayres type of individual preferences, an increase in the level of \textit{ex ante} legal specificity decreases the net benefit from compliance and reduces the incentives to comply with the law.\footnote{In the case of the Kaplow type of preferences, the legal content has the feature of a Giffen good, which increases in demand when its price increases. In the case of those individuals whose preferences are of the type described by the Ayres assumption, the legal content is a normal good that decreases in demand as the price rises.} Thus, the behavioural net-effect of an increased level of specificity depends on the relative magnitude of the portion of people with either the Kaplow-type or the Ayres-type utility preferences.
function and on the relative elasticity of the two aggregate demand functions expressed by the two groups. Let me summarize the two different scenarios.

**Kaplow Assumption.** If we adopt the Kaplow-perspective (i.e., $b_l$ is higher in rules than in standards for actors seeking higher legal precision), both $b_n$ and $c_p$ will be economized by a higher degree of specificity. In other words, rules that are more detailed will maximize the individual’s expected net benefit from compliance (because of the Kaplow assumption) and economize in the production information costs (because of the frequency and homogeneity assumption, which make applicable hypotheses 2 and 3 of section 2.4.). Thus, in this case, rules that are more detailed will reinforce both the cost effect and the behavioural effect.

**Ayres Assumption.** If we adopt the Ayres-perspective (i.e. $b_l$ is higher in standards for those actors seeking only a rough idea of the legal content), then the $b_n$ curve will be negatively inclined, and the $c_p$ cost curves will obviously decrease. Thus, an increase in the level of legal specificity will decrease the individual incentives to comply with legal rules, while economizing on the promulgation costs. In this case, the choice between rules and standards will depend on the relative magnitude of the cost effect and the behavioural effect. To predict which of the two effects is likely to prevail, we need to be able to calculate the elasticity of $b_n$ and $c_p$ with respect to a variation in $\omega$.

Given the presence on the demand side of both types of preferences, the lawmaker is confronted with a trade-off between the effects of $\Delta \omega$ on Kaplow-type people and Ayres-type people. This trade-off is related to the relative degree of the elasticity of the demands for legal rules emerging from the two types. Depending on the prevailing elasticity, the lawmaker faces a different optimization problem. In the case of the prevailing elasticity of the Kaplow-type demand, the lawmaker is likely to increase the level of *ex ante* legal precision in order to exploit the economies of scale with the behavioural effect (thus, there is no trade-off).
In contrast, in the case of the prevailing Ayres-type demand, the lawmaker faces a trade-off between the scale-effect and the behavioural effect. Rules that are more detailed allow for distributing fixed-costs over a larger number of cases (assuming homogeneity and frequency), but a number of individuals might decide to not obey the law because acquiring legal knowledge is too costly. Less-detailed rules might increase legal compliance, but at the same time, they would increase the variable costs in the \textit{ex post} adjudication phase. In these latter cases, the trade-off depends upon the relative degree of elasticity of the lawmaker’s cost-curve and the aggregate demand curve. In general, the content of the Ayres assumption indicates that it seems reasonable to assume that when the Ayres-type of preference prevails, the \textit{ex-ante} specificity demand curve is highly \textit{inelastic} because the interests at stake for the individual are, by assumption, very low.\textsuperscript{283}

In these circumstances, a small increase in $\omega$ might undermine the individual incentives to comply; consequently, the behavioural effect would prevail over the scale effect. However, logically, I recognize that the converse might also be argued. Since, by assumption, the levels of $N$ and $(1 - \phi)$ are high, it is equally reasonable to assume that the elasticity of the production cost curve will also be high (proposition 3 in the previous section), so an increase in $\omega$ could be efficient because of the high level of cost elasticity.

The solution to this puzzle should be sought by a conducting a deeper investigation of the \textit{contents} of homogeneity and frequency. In particular, the following two issues need to be investigated more deeply: (i) what are the elements whose \textit{frequency} affects the \textit{average} information costs? (ii) what are the elements whose \textit{homogeneity} influences the \textit{marginal} information costs?\textsuperscript{284} This inquiry will allow for a better understanding of the optimizing-choice in the most controversial cases (i.e., cases 3 and 4). For this reason, the two

\textsuperscript{283} Recall that, under the Ayres assumption, individuals are already at a low level of $b_i$, and that it is reasonable to assume that at low levels of $b_i$ individuals’ net benefit is not only lower but also decreasing more rapidly with the additional increase in $\omega$.

\textsuperscript{284} In other words, it seems that in the legal economic scholarship, the concepts of frequency and homogeneity have exerted the function of black boxes; that is, they have been used to model the costs of the law-making process, without providing a real explanation as to the elements whose frequency or heterogeneity affects the efficiency of the \textit{ex ante} versus \textit{ex post} production of law. In the present and following sub-sections, I attempt to “open” the black boxes and take a closer look at the conditions of frequency and homogeneity.
remaining sections of this chapter will be devoted to discussing legal homogeneity and frequency.

3.1.4. Legal Homogeneity

The importance of legal heterogeneity in affecting the structure of the law-making costs has been recognized in the literature on law and economics. However, many questions remain unanswered or uninvestigated. In particular, the following issues have received little attention: (i) the definition of legal homogeneity; (ii) the explanation of the mechanism through which homogeneity influences the legal production function; and (iii) the measurement of the degree of heterogeneity for the purpose of institutional choice. In the following subsections, I attempt to further the economic understanding of these unresolved questions. I intend neither to exhaust the complexity of these issues nor to provide the only possible explanatory model of legal heterogeneity. However, I hope to encourage academic reflection beyond the generic recognition of the importance of heterogeneity.

3.1.4.1. Legal Homogeneity as Uniform Demand for Legal Protection

We already know that legal cases can be regulated through *ex ante* law-making to the extent that they are homogeneous. Without homogeneity, *ex ante* law-making is not viable. However, we need to clarify exactly how homogeneity counts in terms of efficiency. In particular, we need a criterion for establishing in which respects cases should be homogenous in order to be suitable for *ex ante* regulation. This is a delicate point that requires some thought about the process of the attribution of legal significance to the facts of law. I will call this the “process of legalization”. I focus on three components of legalization: (i) the facts of the law from which the economic need for

---

285 See *supra* note 257.

286 We must be careful in considering the type of homogeneity that is relevant for the purpose of process efficiency analysis. Homogeneity does not presuppose “identity” of the legal cases. Identity is not the condition for legal cases to be regulated *ex ante*; this is good news because otherwise no *ex ante* law-making would ever be possible, since perfect identity of cases does not exist. That would make the creation of legal rules (i.e., legal order) prohibitively expensive for humanity! Fortunately, homogeneity presupposes only that legal cases have some characteristics in common, that is, they are similar in some respects.
coordination arises (hereinafter “facts of law”)\textsuperscript{287}; (ii) the generalizing process entailed by the \textit{ex ante} perspective; (iii) the variety of circumstances surrounding the facts of the law. In the remaining part of this section, I will explain how these elements are related and why homogeneity concerns the process of legalization.

The demand for law originates from the need for coordination among economic actors. In this sense, facts have “legal relevance” (i.e., become “facts of law”) in producing the need for a legal mechanism to ensure economic cooperation.\textsuperscript{288} The crucial point here is that the legal order changes the quality of its regulatory response, depending on the structure of the coordination problem. Because different regulatory needs require different regulatory responses, the process of legalization takes \textit{different} forms according to the different features of the facts of law.\textsuperscript{289} However, the \textit{differentiation} of the legal response is costly and cannot be provided without limits, that is, the differentiation of the legal supply mechanisms is a \textit{scarce good}. Such scarcity calls for cost savings mechanisms in law-making.

Homogeneity is important: it allows the lawmaker to adopt costs \textit{savings} techniques to reduce the cost of providing a legal response to the growing need for legal rules in an infinite variety of circumstances. In particular, homogeneity allows for organizing the supply of a \textit{single} legal response in an indefinite number of facts of law. If a sequence of potential cases that give rise to a specific economic need are \textit{similar in the way they could be redressed by the law}, then the lawmaker can shape the same institutional legal response to all cases. That is, the lawmaker can economize effort when the elements that determine the need for a legal response occur in a \textit{similar} fashion. In brief, from the standpoint of process efficiency, the homogeneity that counts is that regarding the elements of the facts of law that trigger the demand for a \textit{specific} type of regulation. When homogeneity covers \textit{the elements of the facts of law upon which the most efficient

\textsuperscript{287} Examples of different facts of law are externalities, asymmetries of information, hold-up and free-rider problems, common-pool resources, and so on.
\textsuperscript{288} When some obstacles to voluntary cooperation among individuals (i.e., transaction costs) arise, the legal order supplies cost-savings mechanisms under the form of legal rules in order to facilitate the solution to the coordination problem.
\textsuperscript{289} This point is the subject of the contribution by Anthony Ogus, \textit{Legal Form and Economic Theory} (Oxford: Hart Publishing, 2004)
legal response depends, then the lawmaker is able to organize the supply of legal rules in order to exploit the economies of scale.

Let me elaborate this point through an example. A classical case generating the demand for law is that of negative externalities. Common elements in all negative externalities are the following: (i) a minimum of two economic actors; (ii) external effects; and (iii) a dynamic effect in terms of ex ante incentives to undertake the activity affected by externalities. A typical response from the legal order to the externalities problem is to provide tort law rules that establish legal liability for the person who commits the tort. Tort law rules attribute legal significance precisely to those elements that are relevant for legal protection. The legal definition of “tort” requires (i) the presence of two subjects (the injurer and the victim), (ii) the damage suffered by the victims, and (iii) the causal relation between the injurer’s conduct and the damage. The information costs of the ex ante regulation of the whole set of possible externalities characterized by the above three elements is prohibitive. Therefore, in general, the lawmaker makes use of low levels of ex ante precision in establishing the general principles informing the law of torts. In many legal systems, the fundamental principles of tort law are established in the form of legal standards (regardless of whether they are established thorough enacted law, judge made-law or administrative regulation).

However, within the indefinitely large set of possible negative externalities defined by the three constitutive elements of tort, there are subsets of cases that have in common some additional factual elements that trigger the demand for a more specific legal protection. These latter cases are homogeneous, not only with respect to the three constitutive elements of tort but also with respect to further elements that qualify more specifically the demand for legal rules. For example, one important subset of externalities

---

290 In other words, if cases are similar in that they express the need of the same type of regulatory response, they can be regulated by the same supply of law, and, in that manner, allow for economizing the costs of the supply of law.

291 The foundational principles of tort law are applicable to an indefinite (and potentially infinite) variety of legal cases. However, given the unpredictable heterogeneity of the cases of externalities, since the level of detail is so low, almost all of the costs of the law-making process are allocated at the adjudication phase. We are in this case corresponding to Case 4 of the table contained in Figure 9: low fixed costs, high variable costs, nonlinear marginal cost curve, and, consequently, and no economies of scale.
occurs in the context of a *contractual* relationship between the injurer and the victim. The presence of this additional “transactional” element triggers a more qualified demand for legal protection in comparison to other cases of externalities.\(^{292}\) In particular, the presence of an asymmetry of information between the manufacturer and the consumers of risky products leave consumers imperfectly informed about product-related risk, so firms do not always have the incentive to undertake a full risk analysis. This transactional problem triggers the demand for a specific legal protection that is appropriate regulate this subset of product liability cases. In addition to the generic obligation of restoring damages,\(^{293}\) the efficient legal response to the asymmetry of information is to facilitate the flow of *information* that the unregulated market is not able to provide. Additionally, the structure of the liability obligation, in cases of product safety litigation, has specific features, compared to generic damages. For instance, many scholars maintain that in these cases, the amount of damages should be equivalent to the value of the *expected cost* imposed on consumers by risky products and should be *consistent* with consumers’ attitudes towards risk. In essence, the transactional element is common to a subset of externalities cases and distinguishes them from the generically defined set of tort cases; the specific distinguishing feature is the element that generates the demand for a specific legal protection. From the perspective of law-making costs, the *homogeneity* of product liability cases, with respect to the demand for a specific legal protection, justifies economically the production of more specific *ex ante* regulation. Specifically, higher homogeneity increases aggregate frequency, which then generates the economizing effect of the production costs. In the discussion that follows, I will expand on this latter point.


\(^{293}\) Damages are common to all cases of tort law and correspond to the need to provide ex ante incentives to take precautions, and to engage in the optimal level of risky activity.
3.1.4.2. Legal Homogeneity as Productivity of Ex Ante Law-making

As previously emphasized, homogeneity allows for the adoption of cost-saving techniques and the related reduction of law-making costs. Instead of providing a single specific supply of legal rules for each single episode requiring legal regulation, the lawmaker proceeds *ex ante* by identifying “classes” of homogeneous cases and providing, *once and for all*, an efficient legal response for all the cases belonging to the same class. In particular, the lawmaker economizes on the production-information costs by regulating *ex ante* “classes” of facts of law that are homogeneous with respect to the type of legal intervention required for overcoming the obstacles to cooperation. To elaborate this point, I need to borrow from legal theorists some concepts that describe the functioning mechanisms of the previously labelled legalization process\(^{294}\).

Legal rules are *prescriptive* rules. They are designed to prescribe conduct to orient the behaviour of people. Any rule prescribing a given conduct can be divided into two components. First, the rule defines “the factual conditions triggering the application of the rule”.\(^{295}\) This component has a *hypothetical* structure. It has been labelled in different ways; I will call it a “factual predicate”, thereby emphasizing the function of describing the *facts* that are relevant for the law. The second component prescribes “what is to happen when the conditions specified in the factual predicate obtain”.\(^{296}\) I will call this second component the “*consequent*”.\(^{297}\) It describes the features of the legal intervention.

The factual predicate contains a *generalization* that selects the properties of the facts that are relevant for the law, that is, the facts having the properties described by the generalization are included in the scope of the legal norm; the facts that do not possess these properties are not regulated. From the economic standpoint, the most important problem is the identification of the *efficient generalization*—that is, shaping the factual

\(^{294}\) This point is subject to a long and overwhelmingly rich tradition of philosophical and legal theory studies. Here, I borrow the main concepts on the structure of legal rules from Frederick F Schauer, *Playing by the Rules. A Philosophical Examination of Rule-Based Decision Making in Law and in Life* (Oxford: Oxford University Press, Clarendon Law Series, 1991).

\(^{295}\) *Ibid* at 23.

\(^{296}\) *Ibid* at 23.

\(^{297}\) *Ibid* at 23.
predicate to include an economically efficient number of potential cases within the scope of the legal norm. The same problem can also be defined as the optimal scope (or breadth) of the legal norm. The breadth of a legal norm is efficient when the difference between the total benefit and cost of defining the scope of a legal provision is maximized. Let me assume that the benefit function of broadening the scope of the norm is defined in terms of aggregate social surplus. In this case, the surplus-maximizing breadth of the legal norm is reached at the point where the marginal benefit equals the marginal costs of the incremental effort in detailing the factual predicate. At this point of optimality, the number of cases (encompassed within the scope of the norm by means of a given level of detail of the factual predicate) maximizes the social surplus.

The reader might ask what the relation is between the problem of the optimal scope of the norm and legal homogeneity. Homogeneity is related to the optimal scope of the legal norm: it influences the marginal productivity of the ex ante regulation. Here I define productivity as the number of cases the lawmaker is able to regulate, whereby inclusion within the scope is defined by the factual predicate. In relation, marginal productivity is the productivity of an additional element of specificity of the factual predicate. I define the functioning mechanism of homogeneity as the causal relation between the degree of homogeneity of the demand for law and the productivity of the ex ante regulation.

By previous assumption, the task of the lawmaker is to maximize the number of persons that obey the law \( n \). For a person to obey the law, it is first necessary that the law apply to

---

298 The problem of the optimal scope of the legal norm is distinct from, but closely related to, the problem of the optimal ex ante specificity of legal rules. I might say that they are two different ways of looking at the same mechanism. On the one hand, the problem of ex ante precision is focused on information costs; on the other hand, the problem of the optimal breadth is in relation with the average productivity of the ex ante regulation. The concept of productivity of ex ante regulation will be discussed below.

299 This brings up the delicate issue of how to define and measure the incremental effort of the lawmaker in detailing the factual predicate. One proxy of the marginal increment of ex ante specificity might be that of the number of additional legal cases regulated by the ex ante legal provision. Included within the scope of the legal provision, an additional class of legal cases is costly for the lawmaker, that is, it entails an informational effort to measure the social cost and benefit of ex ante regulating an additional class of cases; at the same, an additional increment of ex ante legal precision generates benefits in terms of both law-making costs savings and social welfare impact.

300 An alternative, more refined (and difficult) analysis might relate the productivity of the ex ante law-making to the number of legal issues (rather than legal cases) included in the ex ante definition of the scope of the law.
him or her. That is, compliance with the law presupposes “inclusiveness” within the scope of the law. In other terms, compliance depends upon the frequency of the law’s application. For simplicity, assume for a moment that individuals always comply with the law when it applies to them and that the application of the law is homogeneously distributed among the individuals subject to law. Under these simplified assumptions, the lawmaker increases compliance (i.e., the number of people complying with the law) by increasing the number of people falling into the scope of application of the law. That is, the lawmaker increases compliance by choosing the optimal level of specificity of the legal rules that maximizes the aggregate frequency of law. Thus, there is a functional relation between the ex ante precision of the law, the aggregate frequency of law, and legal compliance. By manipulating the level of specificity of legal rules, the lawmaker can expand or reduce the number of cases falling within the scope of the law, thereby affecting the number of people whose behaviour will be affected by the legal rule. It is important to distinguish this mechanism from that described in section 1. In the latter context, ex ante legal specificity influences the individual’s compliance choices through its effects on compliance information costs. Here, ex ante legal specificity influences the individual’s compliance choices through its effects on the inclusiveness of the legal norm. By increasing inclusiveness, the lawmaker increases the number of people whose behaviour is affected by the ex ante regulation.

Under these simplified assumptions and holding homogeneity constant, I can construct the production function of the lawmaker as the relation between legal frequency and the level of ex ante legal precision: \( N = f(\omega) \). Accordingly, the marginal product can be defined as the change in frequency determined by a marginal variation of \( \omega \).

---

301 That is, I put aside for a moment the influence of both compliance-information costs and expected benefits from compliance on individuals’ decisions to abide by the law.

302 Two clarifications on the nature of this function are needed. First, here \( \omega \) is the production factor and \( N \) is the output of the law-making process; that is to say, \( \omega \) is the independent variable determining the value of \( N \). By contrast, in previous sections, I have assumed the legal specificity as the dependent variable. Second, here I do not embark in the analysis of the behavior of the function. Theoretically, it might be argued in different directions. It might be said that greater specificity reduces frequency, because more specificity reduces the number of cases that meet all the conditions of applicability of the norm. However, it might also be the case that adding more elements to the factual predicate produces an increase of the number of cases included in the scope of application (think, for example, of the casuistic structure of some factual predicates: more detail might mean more cases included in the range of
My understanding of legal homogeneity is related to the marginal productivity of *ex ante* regulation. I assume that the marginal productivity of *ex ante* law-making is an increasing function of the degree of homogeneity of prospective legal cases, which I denote as $\theta$:

$$\frac{dN}{d\omega} = f(\theta)$$

The more that the demand of law is *homogeneous*, the more that legal frequency becomes responsive to marginal changes in the breadth of the scope of application of the law, as determined by marginal changes in the level of *ex ante* legal precision.

In conclusion, holding constant the level of *ex ante* legal specificity, higher homogeneity increases the marginal productivity of *ex ante* law-making. The practical consequence is that at high levels of homogeneity, it is economically efficient to anticipate the law-making activity at the promulgation phase; on the contrary, at low levels of homogeneity, given the low productivity of *ex ante* law-making, it is economically efficient to leave the specification of the legal content to the adjudication phase. This is not a surprising conclusion. We already determined in section 3.2 that homogeneity calls for *ex ante* law-making. Relevant here is the different explanation for the underlying economic mechanism. In section 2, the comparative advantage of *ex ante* law-making was related to the cost elasticity of law-making; here the efficiency advantage of law-making is related to the elasticity of the marginal productivity of law-making, with respect to changes in the level of the *ex ante* precision of the legal rules.

---

application of the norm); here, any generalization incurs the risk of being arbitrary. Furthermore, a deeper understanding would require embarking on the specificity of the definition of legal precision.

363 To clarify, marginal product is the variation of the number of prospective cases included in the scope of application of the norm as determined by a marginal variation of the ex ante level of the legal precision.
The previous discussion emphasized that homogeneity is better understood as a matter of degree rather than an all-or-nothing issue. From this perspective, I now attempt to provide some criteria for measuring the degree of homogeneity of the legal demand and evaluating its influence on the production function of the lawmaker. Let us suppose that the lawmaker faces the problem of regulating a given social phenomenon and must choose the most efficient law-making technique. The choice is between regulations through more detailed legal rules or less detailed standards, leaving the specification of the legal command to ex post law-making. I proceed by identifying the sequence of the subsequent steps.

(i) **Identifying the relevant variables for regulating the case.** The lawmaker analyzes the facts of law to identify the elements that trigger the demand for legal protection and the relevant variables for the solution of the coordination problem.

(ii) **Measuring the spread of the range of the values of the relevant variables and the distribution of future legal cases.** Cases that generate the demand for legal protection might be more or less heterogeneous with respect to the factual circumstances surrounding the relevant variables. The lawmaker assesses the distribution of prospective legal cases according to the variation in the values of the relevant variables.

(iii) **Identifying the most frequent scenario.** The lawmaker identifies the expected central tendency in the distribution of the future cases, hence making the first estimate of the most frequent scenario among the possible future cases. The most frequent scenario is defined by the spread of the range in the variation of the value of the relevant variables in the most frequent cases.

(iv) **Formulating the hypothetical efficient norm.** The lawmaker calculates the variance in the cases that define the most frequent scenario. On this basis, the lawmaker identifies the contents of the efficient legal norm for regulating the flow of the future cases. To clarify, the norm is structured as follows. The generalization provided by the factual predicate selects the relevant variables of the case: those triggering the demand

---

304 I am thankful to Professor Edward Iacobucci for his willingness to discuss and provide insightful suggestions regarding the issue discussed in this section. The responsibility for any mistakes is mine alone.
for legal protection and those depending on the efficiency of the legal response. The *consequent* prescribes the efficient *standard* of conduct by taking into account the estimated factual characteristics defining the most frequent scenario.

(v) **Estimating the expected variance of the most infrequent cases.** The lawmaker calculates the variance in the cases not included in the range to define the most frequent scenario. Alternatively, the lawmaker measures the distance that the heterogeneous cases are from the most frequent scenario.

(vi) **Measuring the social costs of having an inefficient rule in the heterogeneous cases.** When the expected variance of the most infrequent cases has been identified, the lawmaker measures the social inefficiency generated by the regulation of these cases through the norm designed for the efficient regulation of the most frequent scenario. I call these costs “maladaptation costs”, which are understood as the costs associated with the mismatch between legal response and coordination needs arising from the facts of law.

(vii) **Measuring the costs of an ex post solution.** The lawmaker repeats steps iii to vi (above), and based on a different scope of the relevant range of frequent cases, modifies accordingly the structure of the generalization provided by the factual predicate.

(viii) **Comparing the costs in terms of social inefficiency of the ex ante and the ex post solution.** At this point, the lawmaker identifies the level of homogeneity of the legal demand through a cost-benefit analysis of the inclusiveness of the generalization provided by the factual predicate. On one hand, *ex ante* regulation generates maladaptation costs associated with the inefficient regulation of the heterogeneous cases. On the other hand, *ex post* law-making reduces the maladaptation costs but (depending upon the degree of frequency of the legal demand) increases the production-information costs.

---

305 The literature on the cost associated with imperfect regulation of heterogeneity (i.e., incompleteness of contracts, inadequacies of governance structures, and so on) uses the expression “maladaptation costs”. I will follow this prevailing terminology and label the costs of *ex ante* regulation of heterogeneity as “maladaptation costs”, or more briefly as “costs of heterogeneity”.

Oliver Williamson first used the term “maladaptation costs” in Williamson, *The Institutions of Capitalism*, supra note 65 at 21. See also, more recently, Eric Brousseau and Emmanuel Raynaud, “The Economics of Multilevel Governance” (2006) ISNIE Conference, University of Colorado at Boulder, at 37 [Brousseau and Raynaud, “The Economics”] identify the maladaptation costs as one of the most relevant disadvantages associated with the centralized provision of a collective order. In particular, they define the concept of maladaptation as “the discrepancy between collective orders and individual coordination needs”. Relatedly, they define maladaptation costs as “the difference between net output obtained by implementing the best (feasible) solution and net output obtained by following a more generic rule” (*ibid.* at 21).
costs in the adjudication phase. The legal demand is *heterogeneous* when the maladaptation costs of the *ex ante* solution exceed the production-information cost differential of the ex ante and *ex post* law-making. Conversely, legal demand is *homogeneous* when the maladaptation costs of the *ex ante* solution are less than the production-information costs differential.

In brief, I suggest that the lawmaker measure the degree of homogeneity of the legal demand by calculating the effect in terms of the social costs of a rule that efficiently regulates in advance the most frequent scenario but that is inefficient with respect to the cases that are located away from the central tendency.

### 3.1.4.4. Conclusions on Legal Heterogeneity

Let me summarize in few propositions some concluding remarks on the three questions about legal homogeneity posed at the beginning of this section (definition, mechanism, and measurement).

(i) **Definition.** Homogeneity that counts focuses on the elements of the facts of law that trigger the demand for a specific type of regulation. The legal generalization provided by the factual predicate hinges on the elements of the facts that trigger the demand for legal intervention. (ii) **Mechanism.** Homogeneity affects the marginal productivity of *ex ante* law-making, holding constant the level of *ex ante* legal precision, when homogeneity increases the number of cases that falls into the optimal scope of the legal norm increases. (v) **Measure.** Identifying the efficient norm for the most frequent scenario, and then measuring the social cost of the inefficient regulation of the *infrequent* cases can measure homogeneity and guide the *ex ante-ex post* choice.

### 3.1.5. Legal Frequency

I define legal frequency as the number of times in which the behaviour falling within the scope of the law takes place, or, more briefly, the number of “legal cases”.

306 As we already know, frequency that counts is the frequency of homogeneous cases, namely, the

---

306 The expression “legal cases” usually refers to the litigation stage, and it is used as a synonym for legal disputes. Here, the legal cases are the factual presupposition of the legal compliance, that is, the compliance with the law presupposes the occurrence of a fact covered by the law that generates the obligation to obey the legal command.
frequency of cases regulated by a given amount of *ex ante* legal specificity. Consequently, I will assume hereinafter that frequency refers to a sequence of homogeneous cases.

In this section, I introduce three complicating factors into the model, all of which are related to legal frequency: i) the effect of a high frequency *per actor* on the individual benefit function; ii) the effect of the degree of *ex ante* precision of law on legal frequency; and iii) the importance of the *per capita interests* at stake for assessing the economies of scale effect. The analysis of these three factors provides insights into the solution of the lawmaker’s optimization problem in case 3 (i.e., low frequency and high homogeneity).

### 3.1.5.1. Aggregate Frequency and Frequency *per actor*

As previously emphasized, the distinction between *aggregate* frequency and frequency *per actor* is important in assessing the behaviour of the aggregate utility function. While *aggregate* frequency is related to the economies of scale effect, the frequency *per actor* is a determinant of the individual utility function. Relatedly, frequency *per actor* enables one to consider how the application of the law is distributed among the individuals subject to the law and how this element affects the shape of aggregate utility functions.

I have defined frequency *per actor*, denoted \( \lambda \), as the number of times in which an individual undertakes the regulated conduct. Aggregate frequency, denoted \( N \), is defined as the sum of each individual’s frequency *per actor*:  

\[
N = \sum_{i=1}^{n} (\lambda)_i \tag{23}
\]

where \( i \) denotes each individual subject to the law, and \( n \) denotes the number of individuals engaging in the course of action covered by the law.  

---

307 For the sake of simplicity, I assume here a homogenous distribution of the application of law. This means that I assume \( \lambda = \frac{N}{n} \) for all the people subject to law.

308 The impact of frequency *per actor* on the individual costs of acquiring information about the law has been emphasized by Ayres, who pointed out that, when the frequency *per actor* is low, a standard might
provided by (22) helps to clarify that \( n \) and \( \lambda \) are the two determinants of aggregate frequency. Three considerations are important.

First, each level of \( N \) corresponds to different possible combinations of the values of \( \lambda \) and \( n \), depending on the distribution of the regulated conduct among the people subject to law. Specifically, holding \( N \) constant, the more the regulated prospective legal cases are concentrated on a small number of people, the more the value of \( \lambda \) increases and become close to \( N \) (\( \lambda \rightarrow 1 \)). On the contrary, the more widely the regulated behaviour is distributed among the people subject to the law, the more \( \lambda \) decreases in magnitude and the difference \( (N - \lambda) \) increases.

Second, the distribution of the regulated conduct indirectly affects the level of compliance through its effect on individual incentives to comply. As we have already seen, frequency per actor affects the intensity of the behavioural effect by influencing individual incentives to comply with the law. Thus, holding the aggregate frequency constant, when the regulated behaviour is concentrated on a few actors, the high frequency per actor reinforces the behavioural effect. On the contrary, when regulated behaviours are widely distributed among the individuals subject to law, the frequency per actor is low, which thereby weakens the incentive to comply. More concentration decreases \( n \), which grows smaller and more distant in value from \( N \). More distribution increases \( n \), which grows bigger and closer to \( N \). Thus, \( \lambda \) and \( n \) oscillate in opposite directions but within the same range from \( 1 \) to \( N \). Namely, they are inversely related as shown in Figure 11. In conclusion, holding aggregate frequency constant, the movement of the two variables along the spectrum of their possible combinations produces effects upon legal compliance and ultimately upon the social welfare function.

---

be more efficient for the individual than more detailed rules. See Ayres, supra note 14 at 9: “Even if the conduct to be regulated is relatively frequent in the aggregate, its frequency per actor may be low. For example, while many people may buy wedding rings, few are frequent purchasers. Thus, the costs of learning complex rules may be especially onerous as opposed to the costs of roughly interpreting a complex standard. Relatively transparent standards may economize on the costs of interpretation – particularly when the frequency per interpreter is low” [my emphasis].
Figure 11 Aggregate Frequency and Frequency per Actor

\[ N \]
\[ \lambda = 1 \]
\[ n = N \]
\[ \lambda \rightarrow N \]
\[ n = 1 \]

The above considerations provide an additional element for the analysis of case 3, in which a low level of aggregate frequency constrains the efficiency of *ex ante* law-making. If the low aggregate frequency is coupled with a high frequency *per actor*, then the behavioural effect should also be included in assessing the efficiency of the *ex ante* law-making. Precisely, the impairing effect on the economies of scale exerted by a low level of aggregate frequency should be counterbalanced by the intensity of the behavioural effect determined by the high frequency *per actor*. This is not enough to conclude in favour of *ex ante* law-making, but it is an important element to bear in mind, for the moment. Let me now introduce two other complicating factors.

### 3.1.5.2. Frequency and Scope of Application of the Law

*Aggregate frequency* and *frequency per actor* are related to the *legal* definition of the *scope of application* of the law. The scope of application is the set of formal requirements that need to be met in order for a concrete case to be considered as covered by a specific legal rule. Legal consequences arise only when the concrete case at issue satisfies the set of requirements formally provided by the law (i.e., a statute, a judicial decision, an administrative act, and so on). Consequently, by looking at the formal definition of the scope of the law, it is possible, in many cases, to predict *ex ante* both the magnitude of the flow of legal cases that will be covered by the norm (i.e., aggregate frequency) and the number of people that will be subject to the application of law (i.e. frequency per actor). Simply put, the way the law *defines* its scope of application affects both the *aggregate* frequency and frequency *per actor*. This is an interesting context where the lawmaker himself or herself can affect the features of the demand for law.
In general, the less specific the scope of application of the law is, the more likely that $N$ is high. If the scope of application of the law contains a small number of qualifying elements or if the elements defining the scope of the law take the form of general standards, then the number of cases to be included in the scope of the legal norms will be higher. However, although a higher level of *ex ante* legal precision tends to reduce aggregate frequency, in many cases, it might increase the *frequency* per actor. Fewer people will engaged in qualified behaviour, but those people will presumably incur the application of the law more frequently.

The reader might be induced to observe that in these cases (i.e., infrequent but more qualified behaviour), the lawmaker faces a trade-off between scale effect and behavioural effect,\(^\text{309}\) that is, the more that the scope of the law is qualified, the less frequent the behaviour and, concurrently, the more intense the behavioural effect. However, while there is no doubt that the behavioural effect intensifies when frequency *per actor* increases, the assessment of the scale effect in the context of the production of law is complicated. As I will explain in the next two subsections, economies of scale should not be evaluated by simply calculating the ratio between the average production-information costs and the number of regulated cases.

\(^{309}\) It is interesting to observe the difference between the casual sequence originating by *ex ante* legal specificity on the frequency per actor, and the effects generated by increasing *ex ante* specificity on the law-making costs (previously discussed in section 3 of this chapter).

First, according to proposition 3 in section 3, increasing *ex ante* specificity impairs the scale effect; in contrast, here a higher level of *ex ante* precision intensifies the scale-effect. The reason is that the model previously discussed did not consider the effect of $\omega$ on $N$; in that model, increasing specificity intensifies the scale effect because of a higher cost-elasticity. In contrast, here we consider the effect of the *ex ante* legal precision on the inclusiveness of the legal norm, so that a higher value of $\omega$ might reduce the magnitude of the scale effect by reducing the aggregate frequency.

The second difference pertains to the behavioural effect. Here, the incentives to comply are reinforced because legal precision increases the frequency per actor, while in the previous discussion the behavioural effect came into play as the consequence of the structure of the individual preferences. In both cases, more legal specificity intensifies the behavioural effect, and, consequently, increases legal compliance.
3.1.5.3. The Magnitude of the Interests at Stake

In this subsection, I elaborate the case in which the frequency of law’s application is concentrated on a small number of actors \((\lambda \to N \text{ and } n \to 1)\).

We already know that for each level of \(N\), the \textit{high} frequency \textit{per actor} \((\lambda \to N)\) militates \textit{towards} a higher level of \textit{ex ante} precision of legal rules because its positive affect on \(b_n\) reinforces the behavioural effect by strengthening individual incentives to comply with the law. However, we also know that when the distribution of frequency is \textit{concentrated}, the number of people that fall into the application of the law becomes \textit{small} \((n \to 1)\), which would appear, at first, to advise \textit{against} the \textit{ex ante} production of significant levels of legal specificity. Because the \textit{ex ante} production of legal rules is generally carried out at high, fixed promulgation costs, a doubt arises as to whether it is \textit{socially} efficient to employ these highly expensive law-making processes in producing legal rules that apply only to a small number of people. In other words, the question is whether it is worthwhile having recourse to the \textit{ex ante} production of detailed legal rules in a situation where the outcome is that the law-making process is not applied to a large number of people. The point here is that frequency \textit{per actor} tells us something important about \textit{individual} aggregate utility functions, but it does \textit{not} tell us enough about the “social” efficiency of \textit{ex ante} legal production.\footnote{This is not to say that frequency per actor is irrelevant from the standpoint of the maximization of social welfare. Instead, frequency per actor affects significantly the behavioural response to the promulgation of the law, and for this reason it is causally related to the maximization of legal compliance. All I am saying here is that individual compliance might have a different impact on social welfare that is not captured by the values of \(\lambda, n\) and \(N\).} In particular, frequency \textit{per actor} does not provide information on the \textit{effect} that the single case exerts on the social welfare function.

When \textit{high} frequency \textit{per actor} is coupled with a \textit{high} level of \textit{aggregate} frequency, behavioural effects and the scale effects \textit{jointly} advise in favour of an \textit{ex ante} law-making process. In contrast, when aggregate frequency is \textit{low}, the choice of the \textit{ex ante} law-making process cannot be justified economically \textit{only} upon a high frequency \textit{per actor}. In this circumstance (i.e., \textit{low} aggregate-frequency; \textit{high} frequency \textit{per actor}), the
assessment of the social efficiency of the *ex ante* law-making should take into account other variables that affect the social welfare function.

If I substitute (4) into the definition of social welfare used in (17) and include $\lambda$ in the equation, we obtain the following definition of social welfare:

$$W = \sum_{i=1}^{n} \lambda \delta [(\pi b_g - c_c) - c_i] - c_p$$  \hspace{1cm} (24)

In this definition, we can easily recognize that the aggregate benefit of the *ex ante* production of law does not depend only on the number of people undertaking the courses of action that fall into the scope of application of the law ($n$), the number of cases per person ($\lambda$), nor the total number of prospective legal cases ($n\lambda$).\footnote{I assume for simplicity that the law is homogeneously applied among the people who fall into the application of the law. Therefore, frequency per actor $\lambda = \frac{N}{n}$ is constant and aggregate frequency $N = n\lambda$.} The aggregate benefit also depends upon the magnitude of the interests at stake, which is expressed by the value of $b_g$.\footnote{Previously (see section 1.1), I assumed the condition of individual incentive compatibility. Thus, the expected individual net-benefit from compliance is assumed positive and large enough to make it efficient for the individual to purchase legal information and decide to comply with the law. I do not have assumed $\pi b_g - c_c$ as constant. Now I want to show how changes in the social welfare depend upon the variation in $\pi b_g - c_c$.} Thus, even if the values of $n$ and $\lambda$ are low,\footnote{No matter whether $n$ or $\lambda$ is the lowest in value, or both variables are small.} the magnitude of $b_g$ may be such that the value of $(\pi b_g - c_i)$ is large enough to ensure the minimum socially efficient scale of *ex ante* production legal specificity. There are two important implications for the purpose of equilibrium analysis. First, in addition to the number of cases to which the law is applied, $N$, and its distribution among the people who are subject to the law, $(N - \lambda)$,\footnote{$N - \lambda$ is not a direct measure of the distribution of the application of the law, but reflects the fact that the more widely the regulated behaviour is distributed among the people who are subject to the law, the more $\lambda$ decreases in magnitude, and the difference $(N - \lambda)$ grows larger. In this sense, for the limited purpose of my model, $N - \lambda$ is taken as an indicator of the degree of distribution.} the assessment of returns to scale effects of the law production should take into account the magnitude of the interests at stake. Second, the higher the value of $b_g$, the more that low levels of $n$ and $\lambda$ might be compatible with a socially efficient production of a high level of *ex ante* legal specificity.

\[311\] I assume for simplicity that the law is homogeneously applied among the people who fall into the application of the law. Therefore, frequency per actor $\lambda = \frac{N}{n}$ is constant and aggregate frequency $N = n\lambda$.

\[312\] Previously (see section 1.1), I assumed the condition of individual incentive compatibility. Thus, the expected individual net-benefit from compliance is assumed positive and large enough to make it efficient for the individual to purchase legal information and decide to comply with the law. I do not have assumed $\pi b_g - c_c$ as constant. Now I want to show how changes in the social welfare depend upon the variation in $\pi b_g - c_c$.

\[313\] No matter whether $n$ or $\lambda$ is the lowest in value, or both variables are small.

\[314\] $N - \lambda$ is not a direct measure of the distribution of the application of the law, but reflects the fact that the more widely the regulated behaviour is distributed among the people who are subject to the law, the more $\lambda$ decreases in magnitude, and the difference $(N - \lambda)$ grows larger. In this sense, for the limited purpose of my model, $N - \lambda$ is taken as an indicator of the degree of distribution.
In conclusion, if the interests at stake are sufficiently large, then the *ex ante* production of law might be socially efficient, even when (1) the total number of the prospective cases is small (i.e., low \( N \)), and (2) the number of cases per actor is small (i.e., low \( \lambda \)). In fact, even under these circumstances, the social value of the law’s application might be significant enough to justify the high fixed cost of *ex ante* legal specificity.

The next section elaborates the importance of the interests at stake in assessing the economy of scale effect.

### 3.1.5.4. The Number of Protected Third Parties

The magnitude of the interests at stake might be difficult to measure, especially from an *ex ante* perspective.\(^{315}\) I do not have an answer for the delicate problem of measuring and assessing the expected social welfare impact of the application of a specific law. However, I suggest that the consideration of the number of people whose interest is “directly affected” by the application of the law provides a good index of the magnitude of the interests at stake. Let me elaborate.

Experience shows that in general, the formal *scope* of application of the law does *not* in itself provide information on the magnitude of interests at stake in a *single* case. The legal definition of the scope of application of a legal command can attribute legal relevance to legal cases in which relatively small and circumscribed interests are involved, as well as to cases involving large numbers of stakeholders and entailing large and diffuse interests. The reason is that the effect on social welfare of legal compliance is not dependent on the number of *recipients* of the legal obligations provided by the law but is related to the number of *individuals holding the interests protected by the rule*. The recipients of the legal command and the holders of the substantial interests protected by the law do not necessarily coincide. Hereinafter, I will refer to the people who are *not* recipients of the obligation arising from the legal command, but whose interests are directly affected by the violation of those obligations, as “protected third parties”.

---

\(^{315}\) The purpose of process efficiency analysis here is that of identifying the optimal level of *ex ante* legal specificity, based upon the characteristics of the demand of law (thus, on the features of the prospective profusion of legal cases). In this view, I recognize that assessing in advance the magnitude of the interests at stake is not an easy task for the lawmaker.
Let me consider, for example, the case of a legal rule that applies to a relatively small number of “large-number externality cases” in which the number of polluters is small, but the number of persons significantly affected by polluting emissions is substantial. In this case, the number of the people who are potentially affected by the negative externality is high, even if the number of the people directly falling into the subjective scope of the legal rule as recipients of the legal obligation might be relatively small. In this case, it would be misleading to assess the return to scale effect based on the number of people who fall within the formal scope of the law as formal recipients of the legal command. Instead, a comprehensive assessment of the interests at stake is required to account for the interests of the protected third parties.

Of course, in many cases, the magnitude of the interests at stake is not predictable nor is the number of individuals whose interests are protected by the application of the law predictable. However, when the regulated activity is qualified, it is possible to make a fairly reliable prediction of the interests at stake in the future application of the law. In externality cases, the more that the objective features of the activity generate externalities and the more that the subjective characteristics of the polluter are qualified, the more that the prospective cases become identifiable and predictable, as well as the objective and subjective features of the external effects. In these cases, despite the fact that the number of people formally subject to the prescriptions of the law might be small, if the per capita interests of the protected third parties are significant enough, from the ex ante perspective, the provision of a high level of ex ante legal specificity might be considered socially efficient.

In conclusion, in the case of low aggregate frequency and high frequency per actor in which a small number of people are addressed by the legal command, the efficiency assessment of ex ante law-making should extend to a substantive inquiry into (i) the range, the magnitude, and the nature of interests protected by the norm, and (ii) the

---

number of people whose interests would be damaged as a direct consequence of the violation of the legal command.

3.1.5.5. Conclusions on Legal Frequency

In this section, I summarize the most important points of legal frequency. I have defined legal frequency as the number of prospective times in which individuals who are subject to the law undertake the regulated conduct. Based on this definition, I have identified four variables related to frequency analysis: (i) aggregate frequency, defined as the total number of prospective legal cases, is a determinant of the economy of scale effect; (ii) frequency per actor, defined as the number of cases per person, is related to the distribution of the application of the law. It presupposes a distinction between the set of people that falls into the scope of application of the law, and, among those, the subset of people that actually engage in the course of action regulated by the law. Frequency per actor affects the intensity of the behavioral effect (i.e., elasticity of the behavioral effect in a change in the level of ex ante legal specificity) through its effect on individual utility function; (iii) the assessment of the economy of scale effect, that is, the evaluation of the average cost function should take into the account the magnitude of the interests at stake; and (iv) one possible reliable index of the interests at stake is given by the number of people whose interests are directly influenced by the regulated conduct.

3.1.6. Conclusions on Production Efficiency

Thus far, I have left unsolved cases 3 and 4, in which low frequency and low homogeneity, respectively, impair the economizing effect of the ex ante production of law. In this section, I verify whether and to what extent the previous discussion about frequency and homogeneity provides any useful insights into these two problematic scenarios. I then conclude by summarizing the most important findings on the ex ante versus ex post creation of law.

Case 3: Low Frequency, High Homogeneity (Low-High)

In this scenario, although the high homogeneity of the demand for law allows for ex ante law-making, the low level of frequency impairs the potential economizing effect of
homogeneity. Specifically, low frequency precludes spreading the high fixed costs of *ex ante* law-making over a sufficient number of legal cases in order to minimize the average costs of production. Nonetheless, the previous discussion identified two characteristics of the demand for law that militate towards *ex ante* law-making despite low frequency. First, when the application of the law is *concentrated* on a restricted portion of the population, then the frequency *per actor* tends to be higher. In these cases, *ex ante* law-making produces a stronger behavioral effect because the individual expected benefit from legal compliance increases. Second, the magnitude of the *interests* at stake should be considered in assessing the economy of scale effect. In assessing the interests at stake, account should be taken not only of the individuals who are formally recipients of the legal command but also of the individuals whose interests are directly protected by the legal command. When the interests at stake thus calculated are sufficiently large, then an *ex ante* production of law might be socially efficient, despite the fact that the *total* number of prospective cases is *small*.

In conclusion, when aggregate frequency is too low to ensure the economies of scale effect, a careful analysis of the demand for law is needed. It should take into account the *distribution* of the application of the law and the *impact* in terms of social welfare of the application of the legal rules. High frequency *per actor* and high impact in terms of social welfare might call for the *ex ante* production of law, despite low aggregate frequency. In these cases, people who are subject to the law are more likely to apply the law, and the application of the law itself has a larger positive impact upon social welfare. In brief, the intensity of the behavioural effect and the magnitude of the interests at stake might counterbalance the lack of scale effect in favour of *ex ante* law-making.

*Case 4: High Frequency, Low Homogeneity (High-Low).*

As previously clarified, legal homogeneity is a condition for the technological *feasibility* of *ex ante* law-making (i.e., *ex ante* law-making is *impossible* in the absence of homogenous legal cases). Before inefficiency, *ex ante* law-making of heterogeneous cases is not feasible. Hence, whereas in case 3, we could identify two conditions of the

---

317 A line of future research might identify criteria for the ex ante–ex post choice in the situation.
legal demand offsetting the lack of low frequency, here no offsetting condition can be identified because of the lack of homogeneity. Instead of looking for any offsetting conditions, the best way to approach case 4 is to “measure” the costs of legal heterogeneity. To follow this analytical direction, it is necessary to abandon all-or-nothing logic and treat homogeneity as a matter of degree. From this perspective, the \textit{ex ante–ex post} choice becomes a problem of calculating the \textit{optimal breadth} of the legal norm as a function of the costs of heterogeneity. Specifically, the social costs generated by the “inefficient” \textit{ex ante} regulation of the infrequent cases should be compared to the higher law-making costs of \textit{ex post} law-making. The choice between \textit{ex ante} and \textit{ex post} law-making hinges upon the relative costs of the \textit{ex ante} regulation of heterogeneity versus the costs of the \textit{ex post} regulation of a high number of legal cases.

\textbf{Figure 12: Ex Ante–Ex Post Choice and the Types of Law Demand}

<table>
<thead>
<tr>
<th></th>
<th>High Frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textbf{High Homogeneity}</td>
<td>\textit{(1) Ex ante,}</td>
<td>\textit{(3) Ex ante if:}</td>
</tr>
<tr>
<td></td>
<td>Economies of scale</td>
<td>-High Frequency \textit{per actor}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>\rightarrow Behavioural Effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-High Interests at Stake\rightarrow Larger Impact on Social Welfare</td>
</tr>
<tr>
<td>\textbf{Low Homogeneity (Heterogeneity)}</td>
<td>\textit{(4) Ex post}</td>
<td>\textit{(2) Ex post}</td>
</tr>
<tr>
<td></td>
<td>Measuring the cost of inefficient \textit{ex ante} regulation of infrequent cases versus the cost of \textit{ex post} law-making of frequent demand for law.</td>
<td>No economies of scale</td>
</tr>
</tbody>
</table>
3.2. Agency Efficiency

3.2.1. Ex-Ante Lawmaking and Commitment Problem

The ex-ante lawmaker provides legal content before any decision is made by the people subject to law, that is, the lawmaker creates legal rules in time $t$, and the individual’s decision to comply with the legal rule is made in time $t+1$. The passage of time between the moment that the law is created and the moment in which the regulated behavior occurs generates the problem of “time-inconsistency”.$^{318}$ That is, it might be the case that in the passage from time $t$ to time $t+1$, the incentives, preferences and constraints of the lawmaker and the people subject to law change. This affects the principal–agent relationship between the lawmaker and the people subject to law. In particular, the problem of time-inconsistency generates two forms of the agency problem, which are related to the capacity of the lawmaker to make a credible commitment. Let me briefly clarify this point.

In the context of a principal–agent relationship, the agent’s lack of the capacity to make a credible commitment leads to suboptimal outcomes that have negative effects on all actors. From an ex-ante perspective, the prospective principals might face incentives to refrain from entering into the agency relationship if the prospective agent does not have the capacity of make credible commitments. Once the individuals are entered into an agency relationship, the agents’ lack of credibility might induce the principal to refrain from continuing the relationship, which it then exits. Similarly, in the context of lawmaking, the lawmaker’s lack of a commitment capacity may result in suboptimal outcomes by generating incentives for the people subject to law either not to abide by the rules, opt for alternative informal systems of regulations, or refrain from engaging in the activity regulated by the insufficiently credible lawmaker. Thus, to achieve its goal of producing legal rules that maximize the social welfare, the ex-ante production of legal

rules presupposes the capacity of the lawmaker to commit credibly to (i) respecting expectations in terms of the legal predictability of the people subject to law, (ii) enforcing legal rules against norm violators.

(i) Time-inconsistency associated with *ex ante* lawmaking creates the scope for opportunistic behavior by the lawmaker. As this study will discuss in length, the lawmaker’s preferences and constraints often diverge from those of the people subject to law. Similar to the preferences of individuals, the lawmaker’s preferences might change over time in a new context; moreover, some new time-inconsistent preferences may conflict with the expectations of individuals who make decisions within the constraints affected by the law. Consider the following scenario: the lawmaker produces a rule in time $t$, and in time $t+1$ an individual decides to make an investment based on the costs and constraints that are affected by the existing rule. In time $t+2$, the lawmaker determines a change in the law that increases the costs and aggravates the constraints on the individual. Now if the individual in time $t+1$ anticipates the possibility of a legal change that affects the returns from its investment, he or she might decide not invest, which would generate a net loss of social welfare.\(^{319}\) This simple example shows how the time-inconsistent preferences of the lawmaker generate a serious problem of credible commitment, which, if not solved, might result in significant losses of social welfare.

The problem of credible commitment, associated with the possibility of opportunistic behavior by the lawmaker, generates both agency costs of the outcome and the costs of reducing the agency costs. First, the economic literature has demonstrated that the lack of credible commitment by governments to respect private property rights tends to generate an undersupply of investments, with adverse consequences for the economy.\(^{320}\) These are outcome agency costs. Second, it has been recognized in the public-choice literature that

---

\(^{319}\) This problem has been largely investigated in the economic literature, see *supra* note 318.

many institutional arrangements are designed to enhance the credibility of the lawmaker (e.g., separation of powers, delegation to independent agencies, and so on); the costs associated with the functioning of these institutional mechanisms are the costs of reducing the agency costs.\textsuperscript{321}

(ii) The behavioral effectiveness of \textit{ex-ante} regulation (i.e., its capability of affecting people’s behavior) presupposes the lawmakers’ capacity to make \textit{credible} threats about the punishment of norm violations. This is truly independent of the degree of centralization of the lawmaking process. For example, in the context of private bargaining, the \textit{ex-ante} dimension creates room for opportunistic behavior, especially in the presence of a high degree of asset specificity.\textsuperscript{322} Here the ability of the individual to make \textit{ex-ante} credible threats is essential to ensure the behavioral effectiveness of contractual regulation. Similarly, in the political context, the lack of the lawmaker’s capacity to make a credible commitment may prevent the lawmaking process from effectively regulating people’s behavior.\textsuperscript{323} In fact, legislative rules that are not supported by a credible sanctioning system are likely to be violated and replaced by informal regulations.\textsuperscript{324} In short, regardless of the degree of centralization of the lawmaking process, the \textit{ex-ante} production of legal rules entails a commitment problem related to the enforcement capacity of the lawmaker. Conversely, \textit{ex-post} lawmaking enjoys a

\textsuperscript{321} The commitment problem associated with the \textit{ex-ante} dimension of lawmaking will become clearer in Chapters 5 and 6, in the discussion of the opportunistic behavior of politicians and bureaucrats.

\textsuperscript{322} On the role of private governance institutions in limiting post-contractual opportunism in cases of asset specificity, see Williamson, \textit{The Institutions of Capitalism}, supra note 65.

\textsuperscript{323} Social science scholars have studied the problem of credible commitment from a broader perspective, including legal and constitutional institutions: see, for example, Jon Elster, \textit{Ulysses Unbound} (Cambridge: Cambridge University Press, 2000).

\textsuperscript{324} It might be noticed that this is an enforcement problem rather than an agency problem. However, it should be considered that the agency problem has peculiar characteristics in the lawmaking context. The lawmaker is an agent to whom principals (i.e., people subject to law) delegate the power to produce a law with which they must comply. Following the logic of the agency model, the faithful agent must be able to enforce legal rules that affect the behavior of the principal. This is because the principal is constituted by large sets of people, including both “law beneficiaries” and “targeted individuals”. To satisfy the interests and preferences of law beneficiaries (the principals), the lawmaker has to be capable of enforcing legal rules against the norm violations committed by the targeted actors.
fundamental advantage of enforcement over regulation because it is triggered by norm violation. 325

The institutional solution to the commitment problem is costly. In fact, it often entails either the use of highly expensive centralized enforcing mechanisms or the individual’s participation in a collective decentralized punishment system. 326 As we will see in Part III, both centralized and decentralized institutional responses to the need for enforceability generate significant rent-seeking costs and collective action problems.

3.2.2. Ex-Ante Lawmaking and Rent-Seeking

Ex-ante lawmaking increases the expected benefits and costs of regulation, thereby generating incentives to rent-seeking behaviour. First, the ex-ante perspective, as repeatedly emphasized, entails the lawmaker to regulate “classes” of cases and identify them through the description contained in the factual predicate. That is, the ex-ante lawmaker proceeds to select the properties of future cases that are legally relevant, as well as the content of the legal consequences associated with these factual elements. In so doing, ex-ante lawmaking create rules that generate streams of future costs and benefits. Therefore, all actors that have future interests at stake face strong incentives to invest resources in influencing the lawmaking outcome in a direction favourable to them. For example, in bilateral bargaining, parties who are negotiating the terms of a contract that regulates their future relationship invest resources (e.g., expenses for legal assistance, opportunity costs of not investing in more productive activity, and so on) to come to an agreement on the definition of the factual predicate of the contract’s provisions. In the judicial process, litigants with future interests at stake invest resources to obtain a

325 In a recent article, Shavell demonstrated that ex-post regulation usually enjoys a fundamental enforcement advantage over ex-ante regulation. Although the author does not discuss the credible commitment problem, his analysis converges toward the idea that ex-ante lawmaking entails higher enforcement costs. See Steven Shavell, “A Fundamental Enforcement Cost Advantage of the Negligence Rule over Regulation” (2012) Paper w18418 National Bureau of Economic Research.

326 To clarify, the problem of credible commitment is generated by the time-consistency associated with the ex-ante dimension of lawmakers, independent of the degree of centralization. The solution to the commitment problem often entails a centralized enforced mechanism. For example, the commitment capacity of individuals in private bargaining (an ex-ante lawmaking mechanism that in general is decentralized) depends upon the enforceability of contractual agreements that is achieved through centralized institutions. However, as I will discuss in Chapter 8, under certain conditions (i.e., close-knittedness), enforcement is also possible in the absence of centralized enforcement authorities. It is important to emphasize here that the ex-ante dimension generates a problem of credible commitment, which in turn generates an increase in lawmaking costs.
precedent that is favourable to them.\textsuperscript{327} In politics, special interest groups invest in rent-seeking activity to influence the legislative process, thereby obtaining monopolistic rents. In summary, the \textit{ex-ante} dimension of lawmaking (whether it is centralized or not) increases the expected return from the external pressure that is aimed at influencing the \textit{ex-ante} provision of the legal content.

\subsection*{3.3. Social Choice Efficiency}

From the perspective of social-welfare maximization, \textit{ex-ante} lawmaking has an advantage over \textit{ex-post} lawmaking regarding social choice, that is, under certain conditions, the \textit{ex-ante} perspective of the decision-making process allows for \textit{imperfect} preferences.\textsuperscript{328} From the perspective of social-welfare maximization, legal rules need not serve a particular set of interests but “should proceed from a disinterested perspective”.\textsuperscript{329} In contrast, the \textit{ex-post} perspective that characterizes fairness evaluations is typically focused on specific situations and specific sets of interests. However, it is crucial to distinguish between the constitutional and institutional stage of choice. At the constitutional stage of choice, it is not unrealistic to assume that the constitution-maker acts under a hypothetical “veil of ignorance”. Conversely, at the institutional level, the \textit{ex-ante} dimension is not sufficient to ensure impartial decision-making. As the subsequent comparative institutional analysis will demonstrate, the strength of the influence by special interests groups is affected (and mitigated) by other institutional features of the lawmaking process (i.e., the reason-based decision-making of judges, the inter-jurisdictional in decentralized lawmaking process, and so on).

\footnotesize
\textsuperscript{327} I will discuss this point in more detail in section 7.1.1.1.
\textsuperscript{328} In the context of social choice, individual’s preferences, as Harsanyi has defined them, are impersonal when “they indicate what social situation [the individual] would chose if he did not know what his personal position would be in the new situation chosen […] but rather would have \textit{equal} chance of obtaining any of the social positions […] forming the highest down to the lowest”. John C Harsanyi “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility” (1955) 63 Journal of Political Economy, 4, 309 at 316.
\textsuperscript{329} Kaplow and Shavell, “Fairness Versus Welfare” supra note 128 at 57; see also the text accompanying the note. Kaplow and Shavell adopt the conventional welfare economic perspective that is based on the idea of the maximization of a given social welfare function. From this perspective, as I emphasize in the text, in symmetrically situated individuals, the ex-ante perspective enables an impartial decision-making process among the various participants in the social choice process (see supra note 160). As already clarified, this is not the perspective adopted in this study.
Finally, it should be recognized that *ex-ante* lawmaking entails significant advantages over *ex-post* lawmaking in dealing with the problem of intransitive cycling in private bargaining (i.e., “empty core”). This problem arises when “unanticipated circumstances put parties in a position of forced market relationship requiring post-contractual negotiations.” In these cases, *ex-ante* lawmaking processes enjoy comparative advantages over private contracting because they provide individuals with the opportunity to devise legal rules *before* the facts leading to empty core bargaining have occurred. I will discuss this issue in greater detail in section 8.2.

### 3.4. Adaptive Efficiency

The *ex-ante* versus the *ex-post* dimension affects the adaptive efficiency of lawmaking for two reasons. First, the general nature of prospective legal rules imposes higher immediate interests, compared to the *ex-post* dimension, which imposes a non-incremental nature on the legal change. This is apparent in the political process, in which the general nature of legislative legal rules increases the visibility of the interests immediately affected by the legal change, which in turn increases both the adaptive transaction costs associated with the decision-making process and the *resistance* costs generated by the losers’ opposition. Similarly, the *ex-ante* perspective in private bargaining generates increasing costs of legal change. For example, the costs associated with the *ex-ante* dimension of rule-making explain why parties involved in a bilateral negotiation leave gaps in contractual provisions or prefer to rely on default rules instead of negotiating the actual terms of contractual provisions. Conversely, the fact-specific nature of *retrospective* legal rules enables the lawmaker to lower the political visibility of the legal change, thereby lowering the *resistance* costs generated by the losers’ opposition.

---

331 *Ibid.* at 1243
Second, the ex-ante perspective can generate a problem in the lawmaker’s time-inconsistent preferences, which results in short time horizons of the decision-making process that underlies the production of legal rules. In particular, if the discount function of the lawmaker is hyperbolic, then a lower benefit from legal rules in the short run is preferred over a higher one in the long run. This is independent of the agency problem analyzed above: the lawmaker’s hyperbolic discount function is compatible with the assumption that the lawmaker is a faithful agent of the principal (i.e., people subject to law) because the lawmaker’s “myopia” might be a consequence of the principals’ myopia. The consequences of hyperbolic time discounting in the lawmaking process are straightforward. The lawmaker might value the expected benefit of a given set of legal rules $x$ more than an alternative set of rules $y$ in the long run, but because of hyperbolic discounting, it produces $y$, which generates higher expected benefits in the short run.

Conclusions

Let me now summarize the findings of the discussion on technical efficiency of the ex ante and the ex post choice.

(i) Frequency and homogeneity generate different effects on law-making costs. On one hand, legal homogeneity is the technical condition of the feasibility of ex ante law-making; on the other hand, legal frequency affects the average cost of ex ante law-making.

(ii) The degree of homogeneity can be explained in two different ways. First, the degree of homogeneity affects the marginal cost of ex post law-making (adjudication phase). Second, the degree of homogeneity affects the marginal productivity of the ex ante law-making (promulgation phase).

---

332 In general, in economic models individuals are assumed to have an exponential discount function. See Paul A Samuelson, “A Note on Measurement of Utility” (1937) 4 Review of Economic Studies 2, 155. The assumption of exponential time discounting ensures that the preferences of the individual are time-consistent. For example, assume that individual prefers $x$ to $y$ in time $t_n$. The assumption of exponential discount function means that in time $t_{n+k}$, for all $n$, he or she will always prefer $x$ to $y$. That is, exponential time discounting entails time-consistent preferences. In contrast, the hyperbolic discount function entails the possibility of time-inconsistent preferences.
(iii) When the effects of homogeneity and frequency converge,\textsuperscript{333} the indication of technical efficiency is straightforward (i.e., cases 1 and 2). In the case of low frequency and high homogeneity (i.e., case 3), two features of the legal demand might justify \textit{ex ante} law-making: the high frequency \textit{per actor} and the large magnitude of the \textit{interests} at stake. In the case of low homogeneity–high frequency (i.e. case 4), the \textit{ex-ante} production of law becomes extremely costly. In such cases, the process efficiency calculus is based on the relative costs of the inefficient \textit{ex ante} regulation of the heterogeneous cases versus the costs of the \textit{ex post} regulation of a frequent demand for law.

\textsuperscript{333} To be precise, here I should use the concept of “degrees” of homogeneity and frequency and identify a threshold under which the combined effect of both the degrees of homogeneity and frequency determine the efficiency of the ex ante or ex post solution, respectively. I refer to the “convergence” of effect as a shorter means of indicating cases that are relatively straightforward for the analysis of technical efficiency.
In this chapter, I discuss the trade-offs to be considered when centralizing the law formation process. In particular, I will examine the costs and benefits associated with different levels of centralization. The goal is the attempt to identify the features of the demand for law that are causally related to the relative performances of alternative levels of centralization. As in the previous chapter, I will focus on core tendencies associated with centralization, while excluding the specific institutional features of alternative centralized law-making mechanisms. Throughout the chapter, I will emphasize the importance of the degree of heterogeneity of the legal demand in affecting the relative advantages and disadvantages of alternative levels of centralization.

The analysis of the choice of the level of centralization, coupled with the discussion (developed in the last chapter) of the ex ante versus ex post law-making, enables the formulation of some general hypothesis concerning the relative advantages and disadvantages in terms of process efficiency of alternative combinations of these two fundamental institutional dimensions of law-making. These hypotheses, which I will synthesize at the end of the chapter, will be verified in Part III through an in-depth analysis of alternative law-making institutions.

---

In the analysis, it is important to separate the advantages and disadvantages generated by the degree of centralization from those related to other specific institutional features of alternative law-making institutions. For example, legislatures, bureaucracy and agencies share the advantages and disadvantages of centralization despite the differences in the institutional features of their decision-making processes.
Introduction

The issue regarding the optimal level of centralization remains the subject of two strands in the economic literature: i) the economic theory of federalism,\(^{335}\) and ii) the new institutional economics.\(^{336}\) The first of these two approaches focuses upon the provision of centralized public goods, and the second is concerned with the general issue of the optimal institutional framework for providing legal order in a society. Both perspectives are relevant here because the law is a public good serving the social order in a given community. Based on the tools provided by these two theoretical frameworks, I investigate the ability of centralized law-making processes to gather and process information (i.e., productive efficiency), to minimize agency problems (i.e., agency efficiency), and to ensure efficient legal change (i.e., adaptive efficiency).

Before proceeding, I will define the concepts of centralization and decentralization used in this chapter. Centralization entails: (i) a centralized lawmaker having a monopoly on the production of law, (ii) a number of local districts vertically integrated under the control of the centralized lawmaker. The choice of centralizing law-making activity determines three changes in the related organizational structure: (i) it increases the number of people that are subject to a given jurisdictional level; (ii) it reduces the number of actors operating on the supply side; and (iii) it increases the distance between the lawmaker and the people subject to law.\(^{337}\) The costs associated with alternative levels of centralization can be traced back to these three structural factors. Additionally, as authoritatively observed,\(^{338}\) the choice between centralized and decentralized law-


\(^{336}\) For a discussion of centralization from an economic-institutional perspective, see Brousseau and Raynaud, “The Economics” supra note 305 and the bibliography suggested there.

\(^{337}\) The distance that is relevant in the context of the process-efficiency analysis is the “functional” distance, which does not necessarily coincide with the geographical distance. Functional distance is related to the ability of the people subject to law to the direct or indirect control (by exerting influence on the individuals operating within the law-making arena) of the outcome of the law-making process.

making resembles the choice between the centralized planner of the legal order and a competitive market for legal rules. From this perspective, centralization entails a “top-down” law-making process, whereas decentralization increases the role of “bottom-up” law-making.

Decentralization is characterized as follows:
(i) it entails a large number of autonomous local jurisdictions, or in its extreme form is based upon private agreements among individuals;
(ii) the degree of reciprocal autonomy among jurisdictions is decided by the centralized lawmaker who aims at maximizing the social welfare of the national community (as composed of the local jurisdictions). \(^{339}\)
(iii) within each jurisdiction, a local lawmaker aims at maximizing the social welfare of its own district, unless the central lawmaker chooses to internalize the spill-over effects across jurisdictions;
(iv) in the production of efficient legal rules, local jurisdictions compete with each other to attract the greatest number of citizens, free to move from one local jurisdiction to another as attracted by legal rules “offered” at the local level.

4.1. *Ex Ante* vs. Centralized Law-making

As in the *ex-ante* production of law, centralized law-making can be usefully conceptualized as a cost-saving technique that allows the lawmaker to regulate the behavior of *more* people at a *lower* cost. Both *ex ante* and centralized law-making increase the *number* of regulated situations included in the scope of the application of the legal norm, and they combine to determine the negative slope of the average cost curve. However, a sound process-efficiency analysis must distinguish the cost-saving effects

---

\(^{339}\) This assumption reflects the methodological approach adopted in this study, based upon the idea of a centralized constitutional convention faced with the choice as to how to organize the law-making process. However, in practice, the choice of centralized versus decentralized law-making can itself be either centralized or decentralized. Indeed, the choice between alternative allocations of the law-making power among different jurisdictional levels can be decided by the central authority (as I am assuming) or by local jurisdictions; this, in turn, has a significant impact on the competitive dynamic within the law-making process. On this point see Wolfgang Kerber “Interjurisdictional Competition Within The European Union” (1999) 23 Fordham Int'l L J 217 at 232-233 [Kerber, “Interjurisdictional Competition”].
produced by centralization from those generated by *ex ante* law-making. Although in practice, the two institutional dimensions are intertwined, the two effects must be kept analytically distinct. Two important differences should be considered.

First, *ex ante* law-making extends the application of legal rules to “classes” of cases defined by the description contained in the factual predicate of the legal norm. In contrast, centralization expands the scope of the law by extending the geographical area covered by the law to wider “territorial jurisdictions”. Thus, while the magnitude of the economies of scale of *ex ante* law-making depends on the degree of specificity of the factual predicate, the economizing effect realized by centralization depends on the extension of the territorial jurisdiction.

Second, both *ex ante* and centralized law-making require the homogeneity of the legal demand to produce economies of scale, but they differ with respect to the lack of homogeneity: whereas *ex ante* law-making is not feasible without the homogeneity of the legal demand, centralized law-making can be used for regulating heterogeneous cases.\(^{340}\) As previously identified, legal homogeneity is the technical condition under which “classes” of cases are generalized and regulated. Consequently, the lack of homogeneity precludes *ex ante* law-making. In contrast, centralization can operate independently of the homogeneous demand for law. The degree of homogeneity influences the magnitude of the economies of scope realized by centralization, but it does not preclude the feasibility of centralization. Thus, by constraining generalization, heterogeneity forces the centralized lawmaker to proceed *ex post* on a cases-by-case basis. Simply put, heterogeneity precludes *ex ante* centralization but does not preclude *ex post* centralization.\(^{341}\)

---

\(^{340}\) Centralization might benefit from economies of scale depending on the degree of homogeneity of the legal demand.

\(^{341}\) Think, for example, of the adjudication activities operated by agencies, or to the many examples of centralized adjudication processes of the constitutional courts.
4.2. Productive Efficiency

In this section, I briefly outline the comparative advantages in terms of the productive efficiency of centralized and decentralized law-making processes.

4.2.1. The Advantages of Centralization

Centralization improves the productive efficiency of the lawmaker and under some circumstances, enhances the quality of regulation. In this section, I identify the following three efficiency advantages of centralized law-making process: (i) economies of scale; (ii) consistency of the legal order; (iii) internalization of externalities.

4.2.1.1. Economies of Scale

I have emphasized that the production of efficient legal rules requires the lawmaker to gather and synthesize a large amount of information. This activity is costly, and bearing part of the information costs only once affords significant advantages in terms of lower average costs. As previously explained, centralization is the institutional arrangement that allows distributing the fixed costs of law-making across a large number of cases when legal demand is homogeneous (i.e., economies of scale).

In addition to increasing the number of regulated cases covered by the law, centralization increases the amount of technical knowledge employed by the lawmaker in producing the law and thus affects the quality of the law-making outcome. Centralization increases the amount of knowledge available to the law-making process in two ways. On one hand, a greater scope of application contributes to the greater accumulation of legal cases.

---

342 Two clarifications are necessary. First, this point is debatable if we consider the long-run perspective: in the long run, case law benefits from the accumulation of knowledge coming from the use of precedents, while customary law benefits from the knowledge effect of the evolutionary process. Second, I am not saying that all centralized systems are characterized by the knowledge-based production of law (legislatures often lack technical knowledge); what I am saying is that centralization allows concentrating greater technical competence in a single law-making body.

343 See, for example, in the context of the debate on a centralized European contract law, Stefan Grundmann and Wolfgang Kerber, “European System of Contract Law – a Map for Combining the Advantages and Disadvantages of Centralized and Decentralized Rule-Making” (2002) 20 at 295 [Grundmann and Kerber, “European System”]: (“If it is true that the quality of law depends on the number of cases which have been decided within this law, and therefore on the cumulative experience, then from the economic point of view dynamic economies of scale exist. If we have only one set of contract law rules in the EU, more experiences can be accumulated within this law and therefore its quality may be higher in the long run than if different contract laws exist simultaneously”).
which then provides greater feedback to the lawmaker regarding the efficiency of the legal rules. On the other hand, centralization allows the pooling of experts in a given field of knowledge, whereby the measurement and decision-making activity can benefit from the greater competence and technical knowledge. This should improve the quality of the legal outcome because greater cognitive resources can be employed in the performance of sound cost-benefit analyses and in the better design of legal rules. Furthermore, as an effect of the increased amount of information and cognitive resources available to the centralized lawmaker, centralization tends to increase the accuracy of the legal outcome. A knowledgeable and competent lawmaker minimizes the errors in making the law and thereby enhances the quality of regulation (i.e., static economies of scale).

4.2.1.2. Uniformity and Legal Consistency

There is an important advantage to concentrating the production of legal rules in a single source of law because it facilitates the consistency of the legal order through the production of a uniform law. In particular, the centralization of law-making favours the reduction of conflicts of norms, which tend to be higher when legal rules are produced

---

344 From a broader perspective, this aspect of centralization is well described by see Brousseau and Raynaud, “The Economics”, supra note 305 at 16 (“Since collective rules are applied to a wide range of different interactions, the actual performance and potential improvement of these rules can be more easily measured and tested than those applied uniquely to bilateral relationships. All things being equal, the designer(s) of a rule benefit(s) from greater feedback on the actual performance of the rule in various contexts”).

345 Recently, Arcuri and Dari Mattiacci, “Centralization versus Decentralization as a Risk-Return Trade-off” (2007) Amsterdam Center for Law and Economics, Working Paper n. 06 [Arcuri and Dari Mattiacci, “Centralization”] pointed out the existence of a risk-return trade-off in the choice of the optimal level of governance. The core of their argument is that not only the amount of available knowledge but also the level of risk depends upon the level of governance. In this perspective, centralization increases the probability that the lawmaker yields the right decision. This point can be can be conceptualized as an increase in the expected returns of the law-making activity. On the other hand, decentralization realizes a spreading of the risk of error, thereby reducing the riskiness of law-making activity. Following this logic, as Arcuri and Mattiacci suggest, centralization becomes preferable with advanced expertise of the decision-maker, while decentralization becomes preferable with poor expertise. Of course, other factors like risk-aversion and magnitude of the interests at stakes should be considered in the evaluation of the risk-return trade-off.

346 On this point see Grundmann and Kerber, “European System” supra note 343 at 295 (“It seems clear that in a centralized, uniform legal system it is easier to attain [...] consistency and to avoid incompatibilities and conflicts between legal rules. Therefore in a more decentralized system of law, in which rules are enacted by different legislators, problems of incompatibility of legal rules or gaps of regulation can be much a greater problem”).
by a plurality of local authorities. Thus, all things being equal, compliance information costs tend to be lower under a regime of centralized law in which economic actors can rely upon a single, coherent set of rules.

As discussed in the previous chapter, lower compliance information costs increase compliance with the law, which then increases the value of legal rules: the greater the number of persons who comply with the law is, the higher the value of legal rules. On one hand, the experience of lawyers, judges, and public officials charged with implementing legal rules increases as the number of persons applying the law increases; this accumulated knowledge thereby increases the expected benefit of law in terms of a reduction of transaction costs for prospective adopters. In turn, this increased benefit from legal compliance incentivizes people to adopt legal rules that are followed by many actors and that are tested and ameliorated by previous practical application. Hence, a more uniform and consistent supply of law centralization generates dynamic economies of scale on the demand side (i.e., the lawmaker achieves greater legal compliance at lower costs). On the other hand, centralization generates network externalities as it facilitates the convergence of individual compliance on some specific legal rules.

Of course, the uniformity of law only partially solves the problem of conflicts of norms: a uniform law might be unclear, ambiguous, or incomplete, which leaves too much room for different and contrasting interpretations of the law.

See Grundmann and Kerber, “European System” supra note 343 at 300: “The advantage of a centralized set of legal rules is that the information costs about the relevant legal situation for the users of the legal rules (firms and consumers) might be considerably lower than in a decentralized system, in which, for example, fifteen different sets of legal rules exists (as within the EU)”. This advantage might be seriously undermined by the inadequate quality of the legal outcome, as it often happens in reality. For example, many scholars have long pointed out the poor quality of legislation, and its detrimental effect on legal certainty. On this point, see the prophetic contribution of Bruno Leoni in “Freedom”, supra note 197. However, it is worth emphasizing that the poor quality of the legislation is not an effect of centralization per se, rather it is an effect of the politicization of the production of law. On this latter point, see Sartori, Democratic Theory, supra note 199 at 306-313.

That is, especially when people have Ayres-type of preferences.

The presence of network externalities increases the value of legal norms. On one hand, it realizes increasing returns on the supply side, that is, it increases the number of people who decide to abide by the rules. On the other hand, it reduces the transaction costs of individuals because it reduces the costs involved in deciding whether to abide by the rules or to seek alternative behaviours (e.g., engaging in activities that change the existing rules; or opting for extra-legal solutions; or deliberately breaching the law). As I will explain subsequently, network-externalities can cause “path dependence”, which might constitute either an advantage (reduction of information costs) or a disadvantage (creates “lock-in” effects impairing legal change). In addition, path dependence is not an exclusive feature of the centralized uniform law-making; it is a phenomenon occurring with respect to all forms of regulation; see Clayton P Gillette, “Lock-in Effects in Law and Norms” (1998) 78 BUL Rev 813 [Gillette, “Lock-in Effects”]. In
Consistent or uniform law provides a “focal point”\textsuperscript{351} around which individuals can converge and coordinate their behaviour.

The second advantage of greater legal consistency is the diminution of the frequency and costs of litigation. Law and economics has long demonstrated that the divergence of expectations among litigants regarding the outcome of litigation is a determinant of the rate of litigation.\textsuperscript{352} Because legal consistency and uniformity facilitate the predictability of the outcome of the adjudication process, they reduce the divergence of parties’ expectations on the outcome of a trial, thereby reducing the frequency of litigation. Furthermore, when parties decide to engage in a dispute before the courts, legal consistency reduces the costs of litigation because it facilitates the resolution of the juridical questions involved in the dispute. In conclusion, centralization favours the uniformity and consistency of the legal order, which in turn produces cost-saving benefits on the demand side: namely (i) reduced costs of determining the applicable law, (ii) dynamics economies of scale, (iii) network externalities, (iv) reduced rate of litigation and related costs.

Before proceeding, I acknowledge that the relationship between centralization and uniformity is very complex. First, centralization might achieve legal uniformity and consistency in many different ways. In North American legal systems, centralization has been achieved mostly through \textit{federal} levels of law production systems or through the activity of \textit{private} (or quasi-public) law-making bodies. In contrast, in civil-law countries, centralization has resulted in \textit{codification}, based on the idea of a complete, coherent, and

\textsuperscript{351}The concept of focal point has been introduced in literature by Thomas, Schelling, \textit{The Strategy of Conflict} (Cambridge, Massachussets: Harvard University Press, 1960). For an introductory discussion on law as an institution providing “focal points” see: Kaushik Basu, \textit{Prelude to Political Economy: A Study of the Social and Political Foundations of Economics} (Oxford: Oxford University Press, 2000). Finally, the idea of centralized uniform law as provision of “focal points” has been proposed in the legal literature by many authors. See e.g. Larry Ribstein, “Choosing Law by Contract” \textit{(1992)} 8 J. Corp. L. at 273, (“Such mechanisms as common law canons or uniform laws can provide “focal points” that facilitate cooperation”); Gillette, “Lock-in Effects”, \textit{supra} note 351 at 838 (“[…] legislators act as repositories of information, providing a focal point to which all interested parties can look to obtain it.”)

\textsuperscript{352}I will discuss in greater detail this point and the related literature in section 7.1.1.2
self-sufficient body of written laws. Second, the degree of uniformity can be increased independently of centralization through bottom-up law-making processes. For example, in the area of contract law, uniformity is pursued through the standardization process mainly promoted by law firms. It is worth noting that each of these alternatives to achieve greater legal uniformity and consistency is characterized by different performances in terms of adaptive efficiency, which I will elaborate in section 4.5.

4.2.1.3. Regulating Externalities

As the literature on fiscal federalism has demonstrated, centralization is economically justified when the effects of economic activities are distributed across the boundaries of local jurisdictions.

To regulate an externality problem efficiently, the lawmaker must take into account the full costs and benefits of the activity generating externalities. For this reason, economic efficiency requires that the territorial scope of the application of the legal rules encompass the geographical scope of the externality. Local jurisdictions will often fail to account for the spillover effects caused by the activities located in their territory. Conversely, centralized processes having jurisdiction over the entire territory, which are affected by an externality, are most suited for solving the externality problem.

4.2.2. The Advantages of Decentralization

Centralized law-making processes are frequently ineffective in accounting for the local preferences of individuals and the local characteristics of legal cases (hereafter referred jointly as “local conditions”). As heterogeneity increases, decentralization of the law-making process can be used as an organizing tool to improve the responsiveness of the law to local conditions. In the discussion that follows, I examine and explain the relative

---

355 See supra note 344.
advantages of decentralized law-making\(^{356}\) in terms of greater responsiveness to heterogeneous local conditions.

As observed at the beginning of this discussion, centralization induces greater heterogeneity as an effect of the increased size of territorial jurisdictions. Thus, the first obvious advantage of decentralized law-making is less heterogeneity. Second, as an effect of greater distance between the lawmaker and people subject to law, centralized processes have greater difficulties in determining local preferences and gathering information about heterogeneous local conditions. Hence, centralization increases the heterogeneity of local conditions and concurrently reduces responsiveness to local conditions. Although widely accepted in the literature, the two previous statements are merely descriptive and do not actually explain the comparative advantage of decentralization (i.e., they do not identify the mechanism that leads decentralized\(^{357}\)

---

\(^{356}\) The decentralization of the law is the subject of an extensive literature. A classic contribution from a philosophical economic perspective is Hayek, *Law Legislation and Liberty*, supra note 201.


In Chapter 8 I will discuss in more detail the decentralized production of legal rules.

\(^{357}\) Decentralization of law-making can take many forms and degrees (e.g., enacting customs, lubricating private bargaining by assigning property rights, and so on). Here, I briefly mention the general principles
lawmakers to greater responsiveness to heterogeneous local conditions). In the following discussion, I identify three factors explaining the comparative advantages of decentralization: (i) the incentive effect; (ii) the knowledge-creating effect; (iii) maladaptation-costs.

4.2.2.1. The Incentive Effect

The first explanation of the reason that decentralized law-making better addresses heterogeneous local conditions assumes that the level of centralization modifies the incentives of actors involved in the law-making process and who gather and process information. More precisely, decentralization increases the accountability of the lawmaker, which therefore increases the incentives for gathering information of both lawmakers and people subject to law.

This point can be first explained through the analytical framework of contract theory, which is based on the principal-agent model.\textsuperscript{358} In smaller jurisdictions, principals have “additional instruments to punish misbehavior”\textsuperscript{359} of agents and to influence final output. This principal’s increased ability to sanction and influence the agent strengthens its incentives to invest resources in observing the agent’s behavior.\textsuperscript{360} In turn, higher control by the principal increases the incentives of the agents to gather the information required to bring about the desired outcomes. From this standpoint, the efficiency advantage of decentralization is not based on the greater ability to gather information about local conditions; instead, it is based on the stronger incentives of the lawmaker to be responsive to the principals’ preferences.

---

\textsuperscript{358} The importance of the insights developed by modern theory of the firm for the analysis of the appropriate degree of decentralization has been discussed by Jacques Cremer, Antonio Estache, and Paul Seabright, “The decentralization of public services: lessons from the theory of the firm” (1994) 1345 World Bank Publications [Cremer et al., The Decentralization”]. Many authors have provided a formal principal-agent model of decentralization of government; see, for example, Mariano Tommasi and Federico Weinschelbaum, "Centralization vs. Decentralization: A Principal-Agent Analysis" (2007) 9.2. Journal of Public Economic Theory 369 [Tommasi and Weinschelbaum, “Centralization vs. Decentralization”].

\textsuperscript{359} See Tommasi and Weinschelbaum, “Centralization vs. Decentralization”, supra note 358 at 380.

\textsuperscript{360} On this point see Cremer et al., “The Decentralization”, supra note 358 at 31.
The economic theories of fiscal federalism and regulatory competition combine the concept of decentralization and the idea of institutional *competition* among jurisdictions. In a system where individuals are free to move from one jurisdiction to another, people choose to migrate to jurisdictions where the supply of legal rules satisfies the individuals’ coordination needs. Consequently, from the standpoint of institutional local actors, the provision of efficient legal rules becomes a way to attract people and investors from other jurisdictions and hence increase intra-jurisdictional social welfare.

The individual choice of regulation triggers what the literature has long defined as “inter-jurisdictional competition”, which is the competitive process between local jurisdictions to produce legal rules that attract economic actors to their territory. Institutional competition among jurisdictions serves a function similar to *market* competition among firms. Just as market competition maximizes consumers’ surplus by eliminating rent-seeking behavior, institutional competition maximizes social surplus by reducing the amount of room for inefficient local regulation. From this perspective, inter-jurisdictional competition is a powerful institutional instrument for inducing the responsiveness of regulation to individuals’ coordination needs. Another important advantage of institutional competition is that it increases the incentives of individuals to monitor local agents. The possibility of choosing between different jurisdictions increases the expected benefit of the investment in information about alternative suppliers of local rules. 361

Despite these potential benefits in terms of greater incentives to satisfy individual local preferences, the process of competition among jurisdictions can suffer from serious deficiencies that hamper the efficient production of legal rules. First, the actual *mobility* of the population among jurisdictions is generally much more limited than assumed by the traditional theory of fiscal federalism. In the context of institutional competition, the exit strategy for residents is more costly than the exit strategy for consumers in market

---

361 This point is well explained in Kerber, “Interjurisdictional Competition” *supra* note 339 at 225 (“An important advantage [of inter-jurisdictional competition] in contrast to collective decisions […] may be the lack of "rational ignorance" because individual choice between different bundles of jurisdictions maintains the incentives to invest in information about different jurisdictions' offers. Individuals, therefore, directly control the performance of jurisdictions regarding the fulfillment of their preferences, a concept known as voting by feet”).
competition is. Consequently, the competitive pressure on local jurisdictions to improve their attractiveness is lower than the competitive pressure on firms in a competitive market to improve the quality (or lower the prices) of their product.\textsuperscript{362}

Second, there is a crucial difference in the institutional nature of jurisdictions and firms. As Vanberg and Kerber point out, jurisdictions are territorial clubs with “no single owner who can claim the residual”.\textsuperscript{363} This means that local agents who are in a position to influence the local supply of legal rules do not fully internalize the positive or negative feedback from competition with other jurisdictions. The economic consequences of the loss of local residents do not fall directly and entirely upon the actors involved in the local law-making process. This is a decisive difference from ordinary market competition, where owners of the firms that have the residual property rights and rights of control bear the consequences of the loss of their competitiveness. Consequently, incentives for local lawmakers to improve the attractiveness of the legal supply and foster the institutional competitiveness of local jurisdictions are less effective than in ordinary markets.\textsuperscript{364} In short, the costs of inter-jurisdictional migration and the lack of owners claiming the residual may considerably weaken the incentives for institutional local actors to improve the attractiveness of their jurisdiction.

Third, the most serious problem related to competition among jurisdictions is the possibility of a “race to the bottom”. The risk is that the competitive pressure triggered by the mobility of individuals leads regulators to seek competitive advantage by holding down compliance costs for individuals instead of adopting surplus-enhancing legal rules. In the effort to attract influxes of people and investors, local lawmakers might adopt legal

\textsuperscript{362} See Viktor Vanberg and Wolfgang Kerber, “Institutional Competition among Jurisdictions: An Evolutionary Approach” (1994) 5.2 Constitutional Political Economy 193 at 205 [Vanberg and Kerber, “Institutional Competition”] (“When residents are dissatisfied with “their” government, they can escape that government only by moving into a different jurisdiction, a transaction that is in general significantly more costly than changing sellers in market exchange. There are potentially considerable exit costs involved, in particular in the form of “sunk capital” that has to be given up or is devalued significantly by the change in residential location. This includes as a major component accumulated knowledge, skills and expertise that are adapted to the particular environment, but have comparatively little value in alternative environments”).

\textsuperscript{363} See Vanberg and Kerber, “Institutional Competition”, supra note 362 at 205.

\textsuperscript{364} Ibid at 206.
standards of behavior below the optimal level required by efficient regulation. Consequently, institutional competition might generate the negative effect of triggering an adverse selection mechanism in which inefficient rules drive out efficient ones. In the latter case, regulatory competition becomes a source of distortions in the production of law, instead of acting as an efficiency-inducing instrument.

An answer to the question of whether institutional competition leads to inefficient regulation or induces fulfillment of individuals’ preferences cannot be reached at a theoretical level. The nature of the problem does not allow a general conclusion; instead, it entails the effort of identifying case-by-case the institutional framework that is the most suitable for avoiding the adverse selection problem and redressing the competitive process toward the efficient production of legal rules. The institutional features leading to efficiency-inducing competition might vary according to different areas of law and different legal systems.

4.2.2.2. The Knowledge-Creating Effect

Decentralization helps the lawmaker deal with Hayek’s knowledge problem. First, in addition to the above-mentioned incentive-effect, institutional competition among jurisdictions generates a “knowledge-creating effect”, resulting in a greater ability of the law-making process to experiment and explore innovative legal solutions. Second, decentralization allows a better use of the information dispersed in the minds of the local agents; consequently, it is better suited than centralization to find legal solutions to widely dispersed coordination problems where relevant information cannot be centralized. Let me briefly clarify these two points.

365 In the context of interstate fiscal competition, the race-to-the-bottom has been described as a “cutthroat competition”, which means that tax competition among states tends to determine lower levels of taxation to such extent that states do not have enough resources to finance socially efficient levels of public services. Another classic example of race-to-the-bottom is the setting of insufficient environmental standards at the state level in an attempt to lower the costs for prospective firms. Both cases speak in favor of a centralized (federal) regulation setting respectively a minimum level of taxation and minimum environmental standards.

366 This point constitutes the analytical focus of the so-called Austrian School of Economics. In particular, the impossibility of the “economic calculus” by a centralized lawmaker is discussed in Hayek, Law Legislation and Liberty, supra note 201. See also the important contribution of James C. Scott, Seeing like a State. How Certain Schemes to Improve the Human Condition Have Failed (New Haven and London: Yale University Press, 1998).

The previously analyzed advantages of centralization are based on the implicit assumption that the information required to create efficient legal rules is available to the lawmaker and that the content of the efficient legal rule is known or relatively easily knowable by the lawmaker. Under these assumptions, centralization works as a cost-saving technique to avoid the costs of parallel rulemaking processes. The only advantage of decentralization is the provision of an *incentive* mechanism toward greater responsiveness to local conditions.

However, once the assumption of perfect knowledge about the optimal legal outcome is dropped, the presence of parallel, decentralized rulemaking processes becomes the source of a comparative informational advantage. The competition among local jurisdictions triggered by decentralization results in the superior ability to find innovative solutions through the better use of dispersed information. This crucial aspect of decentralization is better understood from an evolutionary perspective that is based on *dynamic* competition, which is understood as a discovery process of unknown information and unforeseeable outcomes.

It is necessary to clarify the difference in focus between the static perspective adopted in explaining the *incentive* effect, and the dynamic notion of competition related to the *knowledge* creating effect. Specifically, static competition among local jurisdictions affects the *incentives* to produce efficient legal rules but remains confined within the

---

368 I follow the line of reasoning developed in Martti Vihanto, “Competition between local governments as a discovery procedure” (1992) 411 Journal of Institutional and Theoretical Economics 413-420 [Zeitschrift für die gesamte Staatswissenschaft] [Vihanto, “Competition Between Local Governments”], which draws from the insights of the Austrian School of Economics.


370 The explanation of the incentive effects of decentralization remains within the standard equilibrium framework, based on the common static notion of competition. From this perspective, competition operates on the assumption that the efficient outcome of the law-making process is already known; the main analytical problem is the incentive structure to gather information and to generate efficient outcomes by the actors involved in the competition process.
existing informational parameters. In contrast, dynamic competition is a discovery process of new information, based upon the use of the local knowledge dispersed in the mind of economic actors. Thus, whereas static competition concerns the efficient use of the existing information (i.e., the efficient production of legal rules selected among the available known legal solutions), dynamic competition concerns the discovery of new information (i.e., the production of innovative legal solutions). In summary, static competition operates as an incentive mechanism, whereas dynamic competition acts as an informational mechanism.

These two alternative forms of competition result in comparative advantages of centralized and decentralized law-making processes, depending on the nature of the informational problem generating the need for legal rules. When the production of efficient law-making depends on the amount of available scientific (or technical) knowledge, the centralized lawmaker enjoys the benefits of greater cognitive resources and greater technical expertise. However, when the production of law depends on the knowledge of circumstances of time and place (“factual informational basis”) dispersed in the minds of economic actors, centralization suffers from severe informational constraints. Indeed, the factual informational basis of centralized law-making is limited to the factual knowledge possessed by the agents operating within the central law-making body.371 In contrast, decentralization benefits from an informational advantage. The dynamic competition among local jurisdictions allows for expanding the factual informational basis, thereby overcoming the informational constraint of centralized processes.

Another source of the informational advantage of decentralization derives from the fact that whereas the agents at the central level are selected according to a uniform set of criteria and tend to operate under the same circumstances, the local agents in a decentralized system are selected according to different criteria and act under different

---

371 This point has been the object of Hayek’s classic contribution: Friedrich A. Hayek “The Use of Knowledge in Society” (1945) 35 American Economic Review 519 at 519-30 [Hayek, “Knowledge”].
circumstances.\textsuperscript{372} Thus, even assuming that the combined knowledge of centralized agents equals that of local agents, decentralized decision-making processes can better exploit the variety of factual information dispersed at local levels. Indeed, central agents act under the direction of the same centralized authority and consequently are constrained to a uniform use of the information. On the contrary, local agents are free from the uniformity constraint imposed by centralized structures and variously employ the information in their possession.\textsuperscript{373}

In conclusion, decentralization increases the factual information available to the lawmaker, enabling the exploration and discovery of new legal responses to the demand for law. All things being equal, decentralization should be preferred to centralization, where the factual information is widely dispersed at local levels and cannot be readily gathered and processed by a centralized authority.\textsuperscript{374}

4.2.2.3. Maladaptation Costs
Uniformity is a double-edged sword. On one hand, it allows legal consistency; on the other hand, it prevents legal supply from providing an efficient response to highly heterogeneous local conditions. This subsection is devoted to explaining the latter point.

Even assuming that the centralized lawmaker is able to gather information about heterogeneous local conditions, centralization imposes a severe uniformity constraint that impairs its ability to tailor the supply of legal rules to heterogeneous local conditions. The uniformity constraint operates independently of the incentive and informational constraint. In this respect, decentralized law-making benefits from a further comparative advantage in terms of the flexibility of the law-making outcome that adds to incentive and informational advantages. In this section, I briefly discuss the costs of the uniformity

\textsuperscript{372} See Vihanto, “Competition Between Local Governments” supra note 368 at 417.
\textsuperscript{373} Ibid. at 417 and 419.
\textsuperscript{374} It is worth noticing that, as a careful examination, this argument is not necessarily related to the problem of heterogeneous local conditions. This latter issue concerns the ability of the lawmaker to have access to the information on local preferences or factual conditions; instead, the argument analyzed above concerns the ability of the law-making process to make an innovative use of dispersed information. That is, the argument of dynamic competition is valid even assuming that centralized and decentralized lawmakers have the same information.
constraint imposed by centralization and the role of decentralization as a maladaptation-reducing technique.

First, the uniformity constraint associated with centralization generates maladaptation costs that are similar in nature to those generated by *ex ante* regulation of heterogeneity. Both *ex ante* and *centralized* law-making realize economies of scope by creating legal rules tailored to the most frequent scenario. However, as we have already seen, people subject to law incur maladaptation costs when heterogeneous cases differ from the most frequent scenario; maladaptation costs amount to the difference between the economic surplus generated by the rule designed for the most frequent scenario and the surplus that would be generated by applying the efficient legal response to each single heterogeneous case.

Second, both *ex ante* and *centralized* law-making pose the problem of minimizing maladaptation costs. *Ex ante* law-making requires defining classes of cases, which implies the selection of legally relevant variables and the measurement of the range of variations of their values. The lawmaker optimizes the scope of the law based on this selection and measurement of relevant variables. As we already know, this optimization process can be separated into two parts: the maximization of the number of people to whom law applies and the minimization of maladaptation costs. Similarly, the centralized lawmaker confronts the problem of creating legal rules applicable to all people included in the relevant jurisdiction and concurrently minimizing maladaptation costs associated with heterogeneous cases. Because heterogeneity (and the related level of maladaptation costs) increases with the size of the jurisdiction, the centralized lawmaker confronts the same optimization problem as *ex ante* law-making, which is maximizing the number of people subject to law at the minimum maladaptation cost. In conclusion, by identifying the optimal “scope” of law and optimal “size” of jurisdiction, both *ex ante* and

---

375 I have discussed maladaptation costs coming from *ex ante* regulation in Chapter 4 section 1.4.3.
376 On this point see Brousseau and Raynaud, “The Economics” *supra* note 305 at 22 (“Maladaptation costs should increase as the implementation scope of an order rises. This is a direct consequence of the increasing heterogeneity of coordination needs as the size of the community grows” (assuming that each party has the same bargaining power).
centralized lawmakers face a trade-off between legal inclusiveness and maladaptation costs.

Decentralization can be conceptualized as a maladaptation-reducing technique that helps the lawmaker maximize legal compliance in cases of heterogeneous local conditions. In cases of homogeneity, centralization increases legal compliance by reducing the compliance-information costs resulting from greater legal uniformity and consistency. However, when the demand for law is highly heterogeneous, centralized law-making produces the opposite effect through increased maladaptation costs, which then reduces individuals’ expected benefit from compliance. That is, the benefits of a uniform supply of law are offset by greater maladaptation costs generated by the discrepancy between an undifferentiated centralized legal response and heterogeneous local coordination needs. Thus, in cases of heterogeneity, the reduction of the jurisdiction helps the lawmaker reduce the maladaptation costs, thereby increasing the expected benefit from compliance and maximizing the legal compliance of the people subject to law. In brief, decentralization enhances legal compliance by reducing maladaptation costs.

It is worth noting that in cases of heterogeneity, the relative disadvantage of decentralization is not (as the reader might at first be inclined to think) the increase of production-information costs. Because, in cases of heterogeneity, centralization does not realize economies of scale and scope, the lack of economizing effects cannot be regarded as a comparative disadvantage of decentralization. Instead, the relative disadvantage of decentralization concerns the lack of regulation of inter-jurisdictional spillover effects; hence, in choosing the level of centralization, the trade-off to be considered is between reduced maladaptation costs (through decentralization) and reduced spillover effects (through centralized law-making).

4.2.3. Heterogeneity and Levels of Centralization

After identifying the trade-offs associated with different levels of decentralization, I now combine the ex ante–ex post and centralized vs. decentralized dimensions. In this manner, I construct a two-dimensional model of the supply of law. The model does not
provide a classificatory description of existing sources of law; instead, it aims at identifying the institutional dimensions of the law-making process that most affect the structure of law-making costs.

I assume a hypothetical constitutional convention confronted with a spectrum of progressively increasing levels of legal heterogeneity from perfect homogeneity to perfect heterogeneity. The aim of the constitutional lawmaker is to organize the law-making response to the changing features of the legal demand in order to maximize process efficiency. The starting point is perfect homogeneity, efficiently regulated—as we already know—through an *ex ante* centralized law-making process. As heterogeneity increases, the lawmaker adjusts the structure of the law-making response ("institutional choice") by acting on two levels. First, the lawmaker chooses whether to allocate the greater part of the law-making effort *ex ante* ("high *ex ante*”) or *ex post* ("high *ex post*”). Second, the lawmaker chooses whether to centralize or decentralize either the *ex ante* or the *ex post* law-making. Various combinations of the two institutional dimensions allow the constitutional lawmaker to maximize the relative advantages of the *ex ante*, *ex post*, centralized and decentralized law-making processes.

The table provided in Figure 13 identifies seven law-making techniques ("strategies"). Before analyzing them, I clarify the logical principles underlying the institutional choice:

(i) the lawmaker’s first choice is whether a given (exogenous) amount of legal specificity is efficiently allocated *ex ante* or *ex post*;³⁷⁷

(ii) the constitutional lawmaker can limit its strategy to the *ex ante*-ex *post* choice or, alternatively, can *jointly* make the *ex ante*-ex *post* choice, and the *centralized versus decentralized* choice;

(iii) the two choices are independent of each other; thus, the constitutional lawmaker can choose the level of centralization of either the *ex ante* or *ex post* law-making or,

³⁷⁷ Thus, there is no logical possibility of concurrently allocating the same amount of law-making activity to the *ex ante* or to the *ex post* dimension (e.g., “high”-“high” or “low”-“low”). There are only four possible combinations: "high *ex ante* + low *ex post*”; “low *ex ante* + high *ex post*”; pure *ex ante*; pure *ex post."
symmetrically, can choose to allocate to the *ex ante* or *ex post* dimension either a centralized or decentralized portion of law-making activity;

(iv) alternative law-making techniques can be ordered along a spectrum between the two opposite strategies of *ex ante-centralized* law-making (that is efficient in the case of perfect homogeneity) and *ex post-decentralized* law-making (that is efficient in the case of perfect heterogeneity);\(^{378}\)

(v) regarding the features of the demand-side, I focus on heterogeneity, since this variable is relevant for both the *ex ante-ex post* choice and centralized versus decentralized choice;

(vi) I will specifically mention other features of the legal demand (frequency, preference structure, nature of required information, and so on) when they are decisive in the choice between alternative law-making strategies.

\(^{378}\) Logically, the variables identified in Figure 13 give rise to a greater number of combinations. However, I consider only combinations that reflect the salient characteristics of the sources of law.
Figure 13 Law-making Techniques for Increasing Levels of Legal Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>EX ANTE HIGH</th>
<th>EX ANTE LOW</th>
<th>EX POST LOW</th>
<th>EX POST HIGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pure Centralized</td>
<td>_____</td>
<td>_____</td>
<td>_____</td>
</tr>
<tr>
<td>2</td>
<td>Centralized</td>
<td></td>
<td>Decentralized</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>_____</td>
<td>Centralized</td>
<td>_____</td>
<td>Decentralized</td>
</tr>
<tr>
<td>4</td>
<td>_____</td>
<td>Centralized</td>
<td>_____</td>
<td>Centralized</td>
</tr>
<tr>
<td>5</td>
<td>Decentralized</td>
<td></td>
<td>Centralized</td>
<td>_____</td>
</tr>
<tr>
<td>6</td>
<td>_____</td>
<td>_____</td>
<td>_____</td>
<td>Pure Centralized</td>
</tr>
<tr>
<td>7</td>
<td>_____</td>
<td>_____</td>
<td>_____</td>
<td>Pure Decentralized</td>
</tr>
</tbody>
</table>
Strategy 1 (Pure Ex ante Centralization)\textsuperscript{379}

As noted above, the starting point is a situation in which the demand for law is “perfectly” homogeneous, that is, the prospective cases are identical. In this hypothetical situation, \textit{ex ante} centralized law-making is the most efficient strategy: the \textit{ex ante} perspective yields economies of scope; \textit{centralization} allows economies of scale.

Technical efficiency can be achieved through \textit{ex ante} law-making, even when prospective cases are not exactly identical. When confronted with very low levels of heterogeneity, the lawmaker can regulate prospective cases through a \textit{casuistic} formulation of factual predicates, that is, the lawmaker enhances \textit{ex ante} legal specificity by providing a legal definition of all possible prospective situations. If (via the casuistic method) the lawmaker succeeds in providing an efficient response to each case (i.e., to minimize the cost of legal-heterogeneity), economies of scope are realized, and the law-making cost curve is similar to the cost curve in cases of homogeneity. The extent to which the casuistic method is worth using depends upon the structure of the preferences of people subject to law (i.e., Ayres type vs. Kaplow type). The demand for legal specificity is higher when the prevailing preference structure among people subject to law is of the Kaplow type. In this case, the casuistic method is efficient until a relatively high level of legal specificity is reached and vice-versa.\textsuperscript{380}

\textsuperscript{379}“Pure” forms of law-making, such as pure \textit{ex post}, pure centralized, and so on, are heuristic devices useful for identifying the relevant institutional variables of the law-making process. However, in the reality, “pure” forms do not exist. “Pure \textit{ex post} law-making” is a mere abstraction since law-making authorities always operate according to \textit{ex ante} legal principles that provide guidance for their activity. Analogously, “\textit{ex ante} law-making” usually requires \textit{ex post} implementation legal rules in the concrete case. Additionally, “centralization” usually requires decentralized structures capable of adapting the application of centralized legal rules to local conditions. Conversely, “decentralization” is functionally related to centralized institutions from which it derives legal authority at the local level.

\textsuperscript{380}Once again, other factors affecting the efficiency of \textit{ex ante} law-making are related to other institutional features of the centralized lawmaker. I will consider these aspects in the next chapter.
Strategy 2 (High Ex ante Centralization + Low Ex Post Decentralization)

Strategy 3 (Low Ex ante Centralization + High Ex Post Decentralization)

As heterogeneity increases, compared to the first scenario, casuistic *ex ante* formulation becomes less and less efficient. When *ex ante* specificity of legal norms increases, production-information and compliance-information costs increase until the point at which *ex ante* law-making becomes too costly for the lawmaker, and the effort to appreciate an increasingly complex legal content becomes too costly for the people subject to law. At this point, the centralized lawmaker adopts a different law-making process (by modifying both the institutional dimensions of the supply side).

First, the lawmaker reduces the level of *ex ante* legal specificity and increases the amount of legal content to be specified *ex post*. The economic consequences of this are (i) an increase in variable costs due to the temporal and repetitive effects, (ii) a reduction in compliance-information costs because of the decrease of *ex ante* legal specificity, and (iii) a reduction in the costs of inefficient *ex ante* regulation of infrequent cases. Second, the lawmaker chooses whether to centralize or decentralize the *ex post* law-making process, dependent upon the nature of the informational problem to be regulated, that is, whether the solution to the coordination problem requires technical–scientific knowledge or dispersed factual knowledge. As we have already seen, in cases of dispersed factual knowledge, decentralization is better suited; conversely, in cases where regulation requires greater scientific knowledge, centralization is more efficient.

Strategy 3 is similar to Strategy 2. In both cases, the lawmaker modifies the structure of the two institutional dimensions: part of the law-making effort is allocated to the *ex post* dimension; the *ex post* law-making is decentralized. The difference between the two strategies lies in the magnitude of law-making efforts allocated to the *ex-post* dimension. The lawmaker prefers strategy 3 to strategy 2 at higher levels of heterogeneity.

---

This point requires two specifications. First, the increase in variable costs due to the shift to the ex post dimension becomes lower and lower as the degree of heterogeneity increases; the reason is that the repetitive effect decreases heterogeneity of legal cases increases. Second, I assume that the frequency of the legal demand holds constant in the shift from the *ex ante* to the *ex post* law-making. Consequently, I consider only the increase in variable cost due to the different law-making technique (*ex ante* vs. *ex post*) without considering any frequency variation that may occur because of the behavioral effect.
Finally, it is worth emphasizing that the choice between strategy 1 on the one hand and strategies 2 and 3 on the other hand depends upon two factors: (i) the costs of regulating *ex ante* all possible prospective cases (and the ability to minimize the costs of legal heterogeneity); (ii) the prevailing preferences-structure of people subject to law (i.e., Kaplow type vs. Ayres type).  

*Strategy 4 (Low Ex Ante Centralization + High Ex post Centralization)*

This strategy is efficient under the following conditions: (i) the demand for law is *heterogeneous*; (ii) the frequency of the demand for law is *low*; (iii) the coordination problem produces effects at the inter-jurisdictional level, and/or the production of efficient legal rules requires specialized knowledge.

At this point, the first characteristic does not require explanation. I have previously explained why heterogeneity calls for *ex post* law-making. The second characteristic does require explanation: low frequency allows the economic *sustainability* of centralization in a scenario where both heterogeneity and the *ex post* perspective increase the costs of producing the law. We already know that the *ex post* perspective increases information-production costs and that heterogeneity impairs the economizing benefits of centralization (i.e., economies of scope, economies of scale, and learning effects). The limited frequency of the legal demand helps contain the costs of centralized law-making, which otherwise would be prohibitive. Regarding the third characteristic, both the inter-jurisdictional dimension of the external effects and the need for specialized knowledge call for *centralized* law-making. As previously explained, centralization allows the internalization of external effects and the exploitation of greater cognitive resources.

Note that the limited realm of *ex ante* law-making is economically justified by the limited degree of the homogeneity of prospective cases.

---

382 Holding everything else equal, if the aggregate demand reflects a Kaplow type of preferences, then strategy 1 is efficient; as the aggregate demand gets closer and closer to Ayres-Type strategy 2, then strategy 3 become more efficient.

383 Examples of Strategy 4 are centralized agencies or bureaucracies that provide *ex post* regulation through adjudicatory procedures within legal frameworks provided by legislatures by specific policy guidelines or administrative acts. In addition, judicial bodies might be centralized (e.g., centralized Constitutional Courts) and act within a general legal framework (e.g., Constitutions).
One might ask what would happen if the frequency of the legal demand increased to levels that were unsustainable for the centralized *ex post* lawmaker. Obviously, the answer would require a careful consideration of the particular case. However, it is reasonable to consider strategy 5 as the most suitable strategy for this scenario. I will explain the comparative advantages of Strategy 5 below.

*Decentralizing Ex Ante Law-making*

The salient characteristic of a second set of law-making techniques (corresponding to strategies 4, 5, 6, 7) is the elimination of *ex ante* centralization. Thus far, law-making techniques have been characterized by the allocation of part of the law-making activity to the *ex ante* dimension. However, as heterogeneity increases, the constitutional lawmaker can choose a more radical approach to decentralizing the *ex ante* law-making process.

It is worth emphasizing that the main characteristic of this alternative set of techniques is the absence of *ex ante* centralization. *Ex ante* centralization imposes a strict uniformity constraint on the supply of legal rules. Furthermore, both the *ex ante* perspective and the centralization of law-making are well suited for legal homogeneity, and they are either incompatible (*ex ante*) with or highly inefficient (centralization) in high levels of heterogeneity in the demand for law.

To increase the flexibility of the legal supply and the responsiveness to local conditions, the constitutional lawmaker can either decentralize *ex ante* law-making or radically eliminate the *ex ante* dimension. Obviously, the efficiency of this law-making strategy depends upon the relative effect on the law-making cost function of the degree of heterogeneity and other relevant features of the legal demand. In particular, a delicate problem of *ex ante* decentralization arises when the spillover effects across jurisdictions are significant. In this case, the identification of the efficient law-making process depends upon the relative magnitude of externality costs with respect to the costs of *ex ante* heterogeneity.
Strategy 5 (High Ex Ante Decentralization + Low Ex post Centralization)

Decentralized *ex ante* law-making can be combined with centralized *ex post* adjudication (e.g., a single natural Appellate Court). This law-making technique allows for combining the flexibility of the *ex ante* decentralization with some advantages of centralization.

As previously emphasized, heterogeneity precludes generalizing factual predicates and thus is not compatible with *ex ante* law-making. However, as previously explained, decentralization works as a heterogeneity reducing arrangement by decreasing the size of the jurisdictions. Thus, decentralization can be seen as an institutional arrangement that enables the lawmaker to take advantage, at the local level, of the *ex ante* perspective in cases where high degrees of heterogeneity preclude *ex ante* “centralized” law-making. However, the centralized adjudicatory phase also allows some advantages typically associated with centralization, such as greater legal consistency and learning curves.

Strategy 5 has *ex post* centralization in common with strategy 4, but it differs in that the greater part of the law-making activity is “*ex ante* decentralized” (rather than *ex post* centralized). Thus, the two comparative advantages of strategy 5 are (i) that the allocation of a greater amount of law-making activity to the *ex ante* dimension accommodates increased frequency more efficiently, and (ii) that the decentralization of *ex ante* law-making confers an advantage in addressing the relatively high level of heterogeneity. Obviously, the choice between strategies 4 and 5 depends upon the marginal effects of frequency and heterogeneity on the production cost and demand function.

Strategy 6 (Pure Ex post Centralization)

In very limited circumstances, process efficiency might be achieved through a pure form of *ex post* centralization. The reason is simple: centralization maximizes its advantages (economizing effects) in the presence of *homogeneous* demand. Conversely, *ex post* law-

---

384 Examples of what I have labeled as strategy 4 are cases in which the adjudication authority enforces norms resulting from a decentralized production process (e.g., customary norms or norms emerging from spread contractual practices). See Cooter, “Structural Adjudication” supra note 338.
making maximizes its benefits (responsiveness to local conditions) in the presence of heterogeneity of the legal demand.

I have identified the conditions determining the efficiency of strategy 4 as characterized by “high ex post centralization”. Strategy 6 differs from strategy 4 in that the former lacks the “low ex ante” component of the law-making process. This suggests that strategy 6 is a suitable legal response only in cases characterized as follows: (i) the same efficiency conditions of strategy 4 (i.e., high heterogeneity, low frequency, and inter-jurisdictional externalities); (ii) the absence of the degree of homogeneity justifying ex ante law-making; (iii) a very low level of legal demand frequency. In legal language, strategy 6 might be qualified as an “ad hoc” type of law-making process; namely, a law-making process established to provide a legal response to very exceptional circumstances that require the specialized knowledge of an institutional body fully entrusted with law-making power and perfectly autonomous from pre-existing legal standards. In fact, centralization allows for the exploitation of specialized knowledge, and the “pure” form of ex post law-making allows for the production of law independent of any ex ante legal standards.

_strategy 7 (Pure Ex Post Decentralization)_

At the highest level of heterogeneity, the constitutional lawmaker may choose pure ex post decentralization. In this case, both institutional dimensions of the supply-side facilitate the efficient regulation of heterogeneous demand for law. Decentralization allows for greater responsiveness to local conditions, and the ex post perspective better accounts for the specific features of the facts to be regulated. However, as I will explain below, “decentralization” may be afflicted with many deficiencies in terms of adaptive and agency efficiency. In particular, pure decentralization might prevent the adoption of legal rules to provide a focal point around which individuals can coordinate their transaction needs. For this reason, as the institutional comparative analysis (and legal history) suggests, decentralized ex post law-making is usually accompanied by some forms of centralization.
4.3. Social Choice Efficiency (Reminder)

The centralization of the law-making process intensifies the “collective” nature of law-making decisions in order to exacerbate the social choice problem dramatically. Since centralization increases the number of the people subject to the law, it entails an increasing amount of additional interests that are “brought to the table”, thereby increasing the heterogeneity of individual preferences. In particular, the higher the number of people involved in the law-making process as the effect of centralization, the higher will be the number of outcome alternatives and issue dimensions in the policy space, which tends to increase the decision-making costs associated to preference aggregation.

In general, the centralization of law-making entails (i) a “collective” process and (ii) a “collectivized” political outcome. The process is collective in the sense that decisions result from choices made by a large number of individuals without the possibility of face-to-face interaction (e.g., the election of the members of representative law-making bodies by the electorate). The outcome is collectivized in the sense that it produces binding effects on the entire community to which they are referred. The collective nature of the process, associated with the collectivized nature of the outcome is the source of dramatic increase of decision-making costs associated with the social choice problem.

Alternative law-making institutions deal with the problem of social choice in different ways. Certainly, the centralization of law-making exacerbates the problem of preference aggregation; however, the efficiency in dealing with the social choice problem depends also on other institutional characteristics of law-making institutions (e.g., majority rule versus unanimity rule, preference-based versus reason-based rationality, and so on.).

---

386 Ibid. See also Coleman, “Social Welfare Function”, *supra* note 229 at 1121 (“[…] collective decisions are events in which control is formally distributed over more than one actor […]][emphasis is mine].
387 Ibid (“[Collective decisions tend also to be events in which any outcome has consequences for many actors). (Emphasis is mine)
388 I will discuss this problem in greater detail while analyzing the political process in section 5.1.1.1. (with respect to the informational aspect of the political decision-making process) and section 5.2 (with respect to the social choice problem).
will discuss this aspect of the law-making process in Part III of this thesis. Now it is enough to note that centralization entails social choices and related increases in decision-making costs.

4.4. Agency Efficiency

The centralization of the law-making process increases the severity of the agency problem and multiplies the opportunities for rent-seeking and private capture. In this section, I discuss the centralizing mechanisms that generate agency costs and examine the determinants of relative rent-seeking pressure at different levels of centralization. I organize the discussion around what I identify as the two main sources of agency costs: (i) internal pressures; (ii) external pressures.

4.4.1. Internal Pressures

Internal pressures are those originating from within the law-making body. There are two explanations of why law-making centralization increases internal rent-seeking pressures. The first is related to the formation of the law-making body at the constitutional stage; the second is related to the standard moral hazard problem.

4.4.1.1. The Exit-Downgrading Effect

In the broader context of the economic analysis of multilevel governance, Brousseau and Raynaud provide a convincing explanation of the relationship between levels of governance and the characteristics of rent-seeking activities. In particular, they emphasize the importance of the costs of the “outside options of agents”, which are intended as the costs for the individual who opts out of the collective rules that are adopted at a

---

389 The problem of opportunistic behavior of centralized lawmakers has long been the focus of scholarly attention from public choice theory. In addition, transaction-costs economics has devoted a great deal of attention to the relative performance of centralized versus decentralized organizations. These conceptual frameworks capture two important dimensions of the law-making process. On one hand, there is a principal-agent relationship between rulers and ruled, which create room for rent-seeking activities that operate within the centralized law-making body. On the other hand, the structure of the agency costs changes depending on the degree of centralization of the law-making process.

390 See Brousseau and Raynaud, “The Economics” supra note 305.

391 Ibid at 23-31. Here the term “agents” is synonymous with individual actors.
central level and widely followed by members of a community. This approach offers insight into the law-making process. I borrow some of these intuitions and adapt them to the specific context of centralized law-making.\textsuperscript{392}

The first effect of centralizing the production of legal rules is an increase in the cost of opting out of the legal order. Centralized legal rules reduce the costs of \textit{coordination} among individuals of a given community because in sharing a common set of legal rules, the latter benefit from reduced transaction costs. This (cost-saving) benefit of centralization represents the opportunity cost for individuals who opt out of the legal rules of the community (“outside option”).\textsuperscript{393} The wider the scope of the application of law is, the higher the number of people opting for established legal rules. Similarly, the greater the number of people who follow legal rules, the higher the utility of the rule-following choice is, thereby increasing the cost of the outside option. In summary, centralization triggers \textit{network} externalities that result in greater exit costs for actors whose preferences are more distant from the adopted legal rule. Such “distant” actors have an interest in seeking an alternative institutional legal rule by opting out of the legal system. The second effect of centralization is the increasing \textit{heterogeneity} of preferences among members of the community. Heterogeneity increases in terms of the intensity of preference misalignment and in terms of the number of people whose preferences are distant from centralized law.

The combination of these two effects generates rent-seeking pressure on the centralized law-making body. The greater the number of individuals whose preferences are misaligned and the greater the number of people who are distant from centralized law, the higher the individual incentives to form sub-coalitions based upon the proximity of their preferences to exert pressure \textit{within} the formation process of legal rules. In brief, increased heterogeneity intensifies incentives to influence the collective law-making outcome. Second, higher costs of outside options reinforce the incentives for the


\textsuperscript{393} This applies especially to individuals whose preferences are more distant from the adopted legal rule, and who have an interest to seek an alternative institutional legal rule by opting out of the legal system.
formation of sub-coalitions. In fact, the more difficult the exit solution is, the greater the incentive for the individual to join the sub-coalition that is most capable of influencing the centralized production of legal rules.

The formation of sub-coalitions produces a retroactive feedback effect that boosts the centralizing tendency. In fact, higher exit costs increase the bargaining power of individuals who are already members of the most influential coalition. The increased bargaining power of the most influential coalition increases the internal pressure *towards* centralization because it tends to concentrate the bargaining power in a smaller and smaller number of larger and more powerful coalitions. This increasing push towards centralization further accentuates the downgrading of the outside option, thereby increasing the centralizing tendency, and so on.

The circular process that generates continuous increases in centralization does not continue indefinitely. Indeed, the second effect of centralization counterbalances the exit-downgrading effect. Increasing centralization accentuates heterogeneity and increases the costs for the more distant members of the community. A critical level of heterogeneity exists, at which it is inefficient for alternative sub-coalitions to keep adopting centralized legal rules without the possibility of influencing significantly the outcome of the law-making process. As heterogeneity increases, alternative sub-coalitions exert greater pressure either on the centralized law-making body or on decentralization, depending on the relative costs of exerting pressure and opting out of the centralized law. Therefore, the circular process described above continues until the increasing costs of heterogeneity induce some people to form alternative sub-coalitions, whereby they create an alternative set of legal rules. Only at that point does the circular effect reinforcing the pressure toward centralization stop, and the distant members of the community opt out of the centralized legal order, thereby constructing a local, heterogeneous set of rules.
It is important to emphasize that the triggering element in this circular dynamic is the individuals’ incentive to engage in rent-seeking activity through organized sub-coalitions. As shown in Figure 14, this incentive stems from the heterogeneity of preferences, and it is strengthened by the same centralization process that downgrades outside options. Hence, the degree (and form) of centralization depends upon the degree of homogeneity of members’ preferences: homogeneous preferences will reinforce the pressure toward centralization, whereas heterogeneous preferences will increase the costs of centralization.
(for more distant members of the community) until the breaking point in the circular process.

4.4.1.2. The Moral Hazard of the Centralized Lawmaker

Centralization affects the relation between the lawmaker and the people subject to law: (i) the ruler-ruled distance\(^{394}\) increases the probability that the incentives of rulers and the ruled are misaligned; (ii) the greater the distance, the more difficult it is for the ruled to observe the rulers’ activities and intervene to correct their behavior; (iii) the increased size of the community that is subject to centralized law reduces the individual’s incentive to monitor the lawmaker’s activity, which then increases the moral hazard problem. In particular, centralization exacerbates the problem of coordination among principals because the larger the number of people who share the effort of influencing the lawmaker is, the more severe the collective action problem in controlling political agents. In fact, as the number of people increases, the probability of influencing the outcome of law-making decreases, as does the individual share of the expected return. In summary, high monitoring costs and the misalignment of incentives create room for the severe agency costs associated with centralization.

4.4.2. External Pressures

In this subsection, I examine three mechanisms that explain why centralization increases the magnitude of the influence costs by reinforcing the external pressures on the law-making body.\(^{395}\) (i) centralization increases the expected return from influencing the law-making outcome;\(^ {396}\) (ii) centralization raises the costs of competition among principals of

---

\(^{394}\) Distance is intended as “functional” rather than geographical distance; see supra note 337.

\(^{395}\) See Paul Milgrom and John Roberts, “Bargaining Costs, Influence Costs, and the Organization of Economic Activity” in Perspectives on Positive Political Economy (Cambridge: Cambridge University Press, 1990) 57 at 80 (“[…] any centralization of authority, whether in the public or private sector, creates the potential for intervention and so gives rise to costly influence activities and to excessive intervention by the central authority. The costs need to be weighed against the benefits of centralization to determine the efficient extent and locus of authority”).

influencing the centralized lawmaker; and (iii) centralization triggers competition among agents to increase the surplus of their constituencies. 397

4.4.2.1. The Expected Return from Pressure

The production of uniform law by the centralized lawmaker and the wider scope of application of the law increase the expected return of the rent-seeking activity, 398 exacerbating the external pressure (i.e., pressure exerted by actors outside the law-making body, such as special interest groups) on the centralized lawmaker. In this subsection, I provide three reasons that uniform law and larger jurisdictions maximize the rents extracted from the production of law.

First, uniformity helps coordinate the activities of interest groups. 399 In the case of decentralization, interest groups must engage in rent-seeking activity in each local jurisdiction, whereas in the case of centralization, rent-seeking efforts can be concentrated at the central level. If interest groups succeed in influencing the activity of

---

397 These two dynamics are characteristically related to the mechanism of “political” representation of interests. However, external pressure and distributive conflicts also arise in other centralized decision-making processes, such as those of administrative authorities (e.g., quasi-independent agencies) or centralized jurisdictional institutions (e.g., centralized constitutional courts). I will discuss these aspects in more detail in the following chapter.

398 In the text, I refer to centralized uniformity as to a system in which the law is enacted and enforced by the central lawmaker, without possibility for local jurisdictions to choose not to adopt the centralized law. An interesting stream of scholarship examines (from an economic standpoint) the so called “uniform law process”. This expression refers to the production of law proposals to be adopted at the State level as a uniform law. It is a form of centralized law-making based on the collaboration of private organizations and states. The most successful example is the adoption of the Uniform Commercial Code (UCC) by almost every jurisdiction. The UCC was formulated by the National Conference of Commissioners on Uniform State Laws (NCCUSL). This is a peculiar type of centralized uniformity, which is conditioned to the willingness of each State to adopt the proposed uniform laws.


399 In a federal system, uniformity can be realized through the adoption of uniform Federal laws or the adoption of uniform laws by all States. If interests groups succeed to capture the federal lawmaker, and influence the law-making outcome in their favor, then the resulting inefficiency of the law will spread to all the local jurisdictions. More complicated is the case of the central private lawmaker in charge of the formulation of the uniform law proposal (e.g., NCCUSL); in this case, the capture of the central lawmaker facilitates interests groups, but does not ensure uniformity, which depends upon the adoption of the proposal by the States.
the centralized lawmaker, uniformity ensures that legal rules will be applied and enforced across all local jurisdictions.\textsuperscript{400} In other words, the homogeneity of uniform legal supply produces large economies of scope, which, in cases of private capture of the lawmaker, favours the activity of organized interests groups.

Second, uniformity increases the exit costs of people subject to law.\textsuperscript{401} While decentralized law-making allows the economic actors who are affected by an inefficient rule to move to another jurisdiction,\textsuperscript{402} uniformity does not allow any exit option without dramatically increasing costs. It might be objected that in the absence of exit choices, actors will shift to a voice strategy to engage in some kind of influencing activity. However, in many cases, this option is likely to be more costly and less effective than opting out of inefficient laws (when this latter option is allowed by a decentralized system). Thus, centralization not only increases the number of people who cannot avoid incurring inefficient law but also forces them towards a relatively more costly voice strategy. In contrast, decentralization and inter-jurisdictional competition facilitate the exercise of the exit option, reduce the monopolistic power of the centralized lawmaker and hence limit the expected return from rent-seeking behaviour.

Third, under a regime of the centralized production of uniform law, individuals who are deprived of exit solutions and are left with only the possibility of engaging in a

\textsuperscript{400} The argument that centralization increases the effectiveness of lobbying is widely accepted in literature. See, for example, Redoano, “Does Centralization Affect Lobbies?” \textit{supra} note 396 at 408 (“a decentralized system of policy making allows for competition among jurisdictions. Under a centralized system, […]], policy makers are essentially monopolists and if a special interest manages to capture the regulator, there might be no recourse for those parties that are adversely affected); see Brousseau and Raynaud, “The Economics”, \textit{supra} note 305: (“the larger the institution, the greater the incentives to collude since the “surface” area for collecting rents is wider”). Interestingly, Ribstein (“An Economic” \textit{supra} note 398 at 147), argues that interest groups seek the wider possible uniformity so to prevent the exit-strategy of people negatively impacted by the law; in particular, they are willing to give-up a most favorable treatment in some jurisdictions in exchange of a less favorable but more widely applied law: “[…] an influential group may decide to take less of a benefit from a uniform law than it could win in a few states in order to secure a law that is likely to be widely adopted. Wide adoption may more fully achieve the group's objectives by making it harder for those who lose under the law to escape it by moving or by contracting to apply the law of a non-enacting state”).


\textsuperscript{402} See the discussion developed \textit{supra} in sub-section 4.2.2.1.
competition to influence the legal outcome suffer from an organizational disadvantage with respect to interest groups that are more organized. In this respect, interest group theory has long demonstrated that large, broadly based interests suffer from the problem of collective action, whereas concentrated special interests tend to be advanced more effectively by smaller and more organized interest groups.\textsuperscript{403}

In conclusion, compared to decentralized systems, under a regime of centralized law-making, widely spread interests tend to be in a weaker position. Exit strategies are precluded by the absence of alternative sources of law, and voice strategies are ineffective because of the well-known free-rider problem. In contrast, special interest groups are facilitated in the exertion of external pressures because the uniformity of a centralized supply of law increases the expected return of rent-seeking and facilitates the coordination of rent-seeking efforts through the homogeneity of the legal supply.

4.4.2.2. Competition among Principals

The second mechanism that increases influence costs affects the competitive dynamics among principals. Consider a centralized law-making system with the following characteristics: (i) the scope of the jurisdiction includes two regions A and B with identical populations (i.e., same population size); (ii) the centralized lawmaker creates and enforces a specific legal rule regulating an externality problem (e.g., pollution) that extends homogeneously\textsuperscript{404} to the entire geographical scope of centralized law; (iii) the residents within each region have identical preferences about the best legal rule, but these differ across the region; (iv) the objective function of the lawmaker is that of maximizing the social welfare of the entire community, that is, the sum of the social surplus of the two regions.

Since the coordination problem generates costs and benefits of the same magnitude for both regions, the diminution of externality determined by the law is beneficial in equal measure for all the people of the community. In this context, the promulgation of the

\textsuperscript{403} Subsequently, in section 5.3.1, I provide an in-depth discussion of interest group economic theory, particularly the organizational advantages of interests groups over the unorganized electorate.

\textsuperscript{404} Because my focus is on the influence costs, I assume that spillover effects are evenly distributed across jurisdictions.
legal rule is a *centralized* public good generating positive, non-excludable, and non-rival effects on both regions. Region A and B exert pressure on the centralized lawmaker to influence the outcome and obtain the promulgation of the legal rule that maximizes its surplus. In particular, each region attempts to exert its political influence in order to impose some of the costs of the legal rule on the other region.

Let me assume that the influential weight of each jurisdiction is a function of rent-seeking expenditures (hereinafter, the “influential function”). Assume also that the influence function is an increasing and a concave function, that is, the marginal influence weight is positive but decreases in value as expenditures increase. Under this assumption, each region will sustain rent-seeking expenditures to the point at which the marginal benefit and the marginal cost of the rent-seeking activity are equal. Because I am interested in the effect of centralization upon the aggregate cost function of law-making, I assume for the sake of simplicity that the two regions have the same technological constraints. That is, they have the same ability and the same instruments in exerting pressure on the lawmaker.

Under these circumstances, the two regions share the same influence function because centralization entails a single law-making entity. This implies that an increase in the influence weight of region A determines a symmetrical decrease in the influence weight of region B. If region A seeks to shift part of the costs of the legal rule to region B, the individuals in region B react by increasing the amount of rent-seeking expenditures to increase the influence weight of region B and obtain the promulgation of a regulation to shift the costs of abating the externality to region A. In other words, the situation of the two regions competing to influence the centralized lawmaker is a strategic game in which the relative ability of one region to influence the lawmaker is relative to the other region. This strategic interaction is a “war of attrition” game,\(^405\) in which the two contestants compete for a resource of a given value by persisting in taking an action and constantly accumulating costs over the time. Each region competes for an advantageous position

\(^{405}\) For a textbook introduction to the war of attrition game see Joseph E. Harrington, Jr, *Games, Strategies, and Decision Making* (New York: Worth Publichers, 2009) at 239.
relative to the other in terms of influential weight, and the value of this advantageous position is equal to the amount of costs shifted to the other region.

Assuming no budget constraints, the “attrition war” will persist until regions A and B reach the point at which the amount of rent-seeking expenditures equals the social surplus generated by the legal rule. At that point, the rent-seeking effort is maximized, no region has obtained a relative advantage, and no region is able to influence the outcome of the law-making process by shifting the costs of the legal rule to the other region. The two regions have neutralized each other. In conclusion, the outcome of the strategic game is a negative sum: no one wins, but both regions deploy their social surplus in a costly competition to influence the law-making process.

Under a decentralized regime, where regions A and B are autonomous lawmakers, there will not be competition to influence the centralized lawmaker; however, the external effects across jurisdictions will not be regulated. Thus, once again, the choice would be between two imperfect ways of organizing the law-making activity. Decentralization allows individuals to choose the supply of law, but it precludes the internalization of spillover effects. Centralization allows for internalizing inter-jurisdictional externalities, but it generates an increase in influence costs, which reduces the social surplus of centralized regulation. The trade-off between internalization of externalities and the reduction of influence costs depends on the magnitude of the spillover effects and on the elasticity of the influence function.

4.4.2.3. Competition among Agents

Centralization significantly influences the dynamic relationships among agents. As previously noted, centralized systems typically generate distributive conflicts among community members with contrasting preferences regarding the allocation of costs generated by legal rules. This conflicting dynamic is reflected in the activity of the

\[\text{406}\]

Here, I adopt a principal-agent perspective. The “agents” are the individuals acting within the centralized law-making body on behalf of the people subject to law (i.e., the “principals”) that delegate law-making authority to them.
centralized law-making body, whose members have an incentive to exert their power to maximize the payoff of the most influential principals.\textsuperscript{407}

Let me assume that a centralized lawmaker regulates an inter-jurisdictional externality problem affecting the population of \( n \) regions, with \( n \geq 3 \). The lawmaker is composed of a number of representatives elected locally. The residents in each region support a legal rule that is consistent with their individual preferences. The elected representatives are loyal to their constituencies and support rules in accordance with the electoral mandate. The preferences are homogeneous \textit{within} each region but heterogeneous \textit{across} jurisdictions. Thus, both the number of local jurisdictions and the intensity of differences in preferences among jurisdictions give the measure of heterogeneity. Finally, the law-making process is based on majority rule, so legal rules are promulgated with the consensus of \( 50\% + 1 \) of representatives.

Because the preferences across jurisdictions are heterogeneous, there is no majoritarian jurisdiction within the law-making body; it follows that each legal rule is promulgated with the support of a minimum willing \textit{coalition}. To be re-elected, each representative belonging to the willing coalition wants to minimize the costs imposed on his or her local jurisdiction and maximize the costs imposed on the residents that are not represented by the willing coalition. Thus, the law-making process will generate a legal outcome yielding a distribution of costs favourable to jurisdictions whose representatives \textit{belong} to the winning coalition.

Because I have assumed that representatives are loyal to their constituencies, they will act within the representative body through organized groups (e.g., in political parties)

\textsuperscript{407} As will become apparent in the next chapter, although this effect is exacerbated when the centralized body has a representative nature, it also occurs when the centralized law-making body has a judicial or technocratic nature. In fact, regardless of the institutional configuration of the law-making institution, the centralized lawmaker is not insulated from the community of the people subject to law and, to some extent, has the incentive to consider the preferences of the people subject to the law. As the substantial body of literature demonstrates, both administrative and judicial bodies are subject to external pressure from private interests. I will discuss in more detail the specific dynamics of each law-making institution in the next chapter. Here I merely focus upon the impact of the centralization factor and assume for simplicity’s sake that agents are elected by principals.
expressing unitary votes. The representative body will be divided into \( n \) groups composed of agents with the same preferences, and legal rules are promulgated with the support of a number of groups reaching the requisite 50% + 1 threshold. This means that the higher the number of jurisdictions with different preferences (i.e., heterogeneity) is, the higher the number of representative groups required to form a minimum winning coalition. Consequently, the more that the decentralized system exhibits heterogeneity across jurisdictions, the higher the cost of reaching a majoritarian agreement within the law-making body will be.

I now want to elaborate briefly on the reasons that an increase in inter-jurisdictional heterogeneity (i.e., principals’ heterogeneity) raises the costs of reaching an agreement among agents in the law-making body. The first explanation is straightforward: reaching an agreement between two parties is likely to be less costly in terms of time and effort devoted to the negotiation than reaching an agreement among three or more parties would be. However, another explanation is less obvious: the increased heterogeneity of principals reduces the expected return from centralization because the outcome of the law-making process must compromise additional heterogeneous interests. In particular, increased heterogeneity generates two effects on the cost functions of agents. On one hand, the higher the minimum “number” of representative groups needed to form a willing coalition is, the greater the distance between the promulgated rule and the first-best option for each local district belonging to the winning coalition. On the other hand, the higher the “intensity” of heterogeneity (i.e., differences in preferences) is among the constituencies belonging to the minimum willing coalition, the more likely it is that the distance between the promulgated rule and the rule corresponding to the first-best option of local jurisdictions will be outside the winning coalition. Thus, the intensity of the preference heterogeneity raises the costs imposed on principals whose agents do not manage to enter into the willing coalition.

In summary, the number of heterogeneous jurisdictions affects the expected return to “belonging” to the minimum willing coalition, whereas the intensity of preference heterogeneity increases the costs of “being excluded” from the winning coalition. This is
an important point and applies to all the centralized processes, regardless of the specific institutional features of the law-making body. First, the number of heterogeneous actors and the intensity of heterogeneity of preferences produce opposing incentive effects with respect to external pressures. The higher the number of heterogeneous actors is, the lower the expected return from external pressure and the higher the bargaining costs and hence the lower the incentive to exert external pressure on the law-making process. Second, the greater the intensity of the heterogeneity of preferences is, the higher the cost of being outside the willing coalition and the stronger the incentive to exert pressure on the centralized lawmaker. Consequently, in cases of intense heterogeneity of preferences, the centralization of the law-making process increases external pressures by those excluded by the winning coalition, whereas decentralization reduces external pressures.

The above considerations shed light on the choice between centralization and decentralization in the presence of heterogeneous preferences and spillover effects across jurisdictions. In this case, the lawmaker should consider the trade-off between the costs of the distributive conflicts among agents (i.e., centralization) and the loss of surplus determined by the non-internalization of spillover effects (i.e., decentralization). The calculus of the trade-off should consider the level of externality effects and the degree of heterogeneity of jurisdictions. When the degree of heterogeneity is low, the critical level of externality is such that centralization produces a higher level of surplus, compared to decentralization, if the level of externality across jurisdictions exceeds the critical value. Holding constant the level of externality, greater heterogeneity of principals’ preferences increases the costs of the distributive conflicts among the agents; consequently, the critical value of the externality above which centralization is more efficient becomes higher. This confirms the explanation in previous subsections that a higher level of externality and a higher degree of homogeneity militate toward centralization. However, the underlying economic explanation is different. Here the cost of centralization, which offsets the advantage of internalizing spillovers, derives from a distributive conflict among the agents. The greater the heterogeneity of the principals’ preferences, the higher the drawback of centralization in terms of both the costs of reaching an agreement and the diminution of the return from centralization for local jurisdictions. Furthermore, the
greater the intensity of heterogeneity, the greater are the influence costs for centralized lawmakers.

4.4.3. Heterogeneity and Agency Efficiency

The foregoing discussion leads to the following conclusions about the impact of law-making centralization on the agency problem, depending on the features of the legal demand:

(i) At the constitutional stage, homogeneous preferences reinforce the pressure to centralize; in contrast, heterogeneous preferences increase the costs of centralization for more “distant” members of the community. The critical level of preference heterogeneity is such that it becomes inefficient for alternative sub coalitions to adopt centralized legal rules. Alternative sub-coalitions tend to either generate pressure on the centralized law-making body to influence the collective outcome or exert pressure to decentralize the law-making process, depending on the relative costs and benefits of these two alternative strategies.

(ii) In centralized law-making processes, preference heterogeneity (a) increases the costs of reaching an agreement within the law-making body; (b) decreases the expected return from centralization for the regions not belonging to the winning coalition; (c) increases the distributive conflicts within the law-making body; and (d) increases the external pressure costs for the regions outside the willing coalitions.

(iii) In choosing the level of centralization, the responsiveness to heterogeneous local conditions is a trade-off between the magnitude of the spillover effects and the elasticity of the influence function of local jurisdictions.

4.5. Adaptive Efficiency

As previously explained, adaptive efficiency entails a trade-off between legal certainty (which calls for greater uniformity of legal supply) and legal change (which calls for flexibility of legal rules). In this respect, centralized and decentralized processes differ in

---

408 That is, preference heterogeneity increases the costs of centralization for individuals whose preferences are more heterogeneous with respect to most common preferences.

409 For a definition of “external pressure costs”, see section 2.3.2.1.
their pursuit of legal certainty, as well as in the related mechanisms of legal change. Centralized law-making seeks to maximize legal certainty through top-down production processes that generate the uniform supply of law; conversely, decentralized law-making seeks to minimize legal uncertainty through the coordination of bottom-up processes that generate a highly diversified supply of law. This section will briefly outline the alternative advantages in terms of adaptive efficiency of the “top-down uniformity” versus the “bottom-up heterogeneity” of the legal supply.

4.5.1. The Coordination of Compliance Decisions

When the existing set of legal rules becomes inefficient because of changes in the technological or economic environment, the adoption of a new efficient regulatory regime requires overcoming a number of informational and transactional difficulties. In particular, legal rules are characterized by severe lock-in effects, which are generated by the path dependence hampering the transition to more efficient legal regimes.

The path-dependence effect derives from the dynamic characteristics of the creation of law, independent of the substantive efficiency of legal rules. The production of legal rules is often characterized by learning scales and network externalities that generate increasing returns on the demand side. These increasing returns then make people who

---

410 Here I will not discuss the vast theme of “transition costs” related to the presence of winners and losers caused by the distributive effects generally associated with legal change. On this point, see Quinn and Quinn and Trebilcock, “Compensation”, supra note 36; Kaplow, “Legal Transitions” supra note 36.

411 For an excellent discussion of both the lock-in effect in the law and the relative susceptibility to the lock-in of alternative sources of social regulation, see Gillette, “Lock-in Effects”, supra note 351.


413 In recent years, there has been a growing interest in the legal-economic implications of network benefits. Although this literature is focused on corporate law, much of the argument pointing out the advantages of network benefits from the past use of legal rules is generalizable to the law in general. See Michael Klausner, “Corporations, Corporate Law, and Networks of Contracts” (1995) 81 Virginia Law Review 757; Stan J Liebowitz and Stephen E. Margolis, “Network Externality: An Uncommon Tragedy” (1994) 8 Journal of Economic Perspective 133; Mark A Lemley and David McGowan, “Legal Implications of Network Economic Effects” (1998) 86 Cal L Rev 479.
are subject to the existing law increasingly “change-adverse”. That is, once a legal rule has come into force, and compliance by community members has stabilized, the incentives required to shift to a new legal regime are undermined by the fact that compliance with the previous legal rules has increased their value.  

When a new efficient legal rule emerges, the incentive structure of the individual can be usefully described as a Prisoner Dilemma situation. The increased value of legal rules raises the opportunity-cost of migrating to a new legal regime and makes the solitary transition to a new legal rule too risky an option, despite the fact that under the new circumstances, the consolidated legal path might be a sub-optimal solution. Each individual is unwilling to transit from the old to the new legal regime “without assurance that a critical mass of potential users will do likewise”.

In the absence of a centralized law-making authority, the sub-optimal equilibrium formed under the inefficient, consolidated legal rule could be overcome in two types of situations: (i) a Pareto coordination game; (ii) the presence of “legal entrepreneurs” with substantial incentives to promote the collective migration to a new legal standard. However, these two situations are not likely to occur very frequently in decentralized systems, as the following discussion clarifies.

A Pareto coordination game requires (i) identical preferences among players over the set of possible outcomes; (ii) perfect information among players regarding the payoffs associated with each possible strategy; and (iii) perfect information among players regarding other players’ moves. In the context of legal change, a pure coordination game requires that (i) the new rule is more efficient for every member of the community, (ii) all members of the community know that the new rule is more efficient and that the individual gain from legal change exceeds the transition costs, and (iii) all members of

---

414 It is worth emphasizing that learning scales and network externalities are not necessarily related to the substantive efficiency of the rule. As repeatedly emphasized, compliance with the law increases as the individual’s expected benefit from compliance increases; in turn, the expected benefit might increase as a result of others people following the rule despite the content of the rule do not maximize economic efficiency.

415 Gillette, “Lock-in Effects”, supra note 351 at 818

416 Ibid.
community are confident that a sufficient number of other members will migrate to the new legal rule. In this context, the Pareto-superiority of the new norm is “common knowledge” among members of the community, and the production of law is *perfectly* adaptive. In contrast, in the absence of the above conditions, the fear of losing the network benefits will prevent people from benefiting from the advantages of collective action; no one moves to the new legal regime because there is no common knowledge that a collective migration will support the efficient legal change. This lack of information generates the lock-in effect, and the production of law is characterized by the stickiness of legal rules.

According to Gillette, the law-making process confronts the problem of transforming a Prisoner’s Dilemma (“in which no party will move to the new equilibrium because they are confident that no one else will”) into an Assurance Game (“in which all parties will be willing to move to the new equilibrium because they are confident that a sufficient number of others will”). The production of legal rules by a centralized lawmaker helps overcome the coordination problems that generate the lock-in effect. First, the presence of a central lawmaker facilitates the dissemination of information about the superiority of the new legal rule. Conversely, decentralized mechanisms are less effective in signalling to individuals the Pareto superiority of the new legal rule. Second, the presence of a centralized authority in prescribing a new standard of behaviour through the production of legal rules that are *binding* on all members of the community can generate the

---

417 The concept of “common knowledge” was introduced into the literature of game theory by David K Lewis, *Convention: A Philosophical Study* (Cambridge, Massachusetts: Harvard University Press, 1969) 52-68. Players have common knowledge of a given proposition when every player knows it to be true, knows that the other players know it to be true, knows that other players know that other players know it to be true, and so on.

418 See Gillette, “Lock-in Effects”, *supra* note 351 at 819: (“In short, no single party will incur the costs of deviation from the existing standard without assurances that offsetting benefits can be realized as a result of mass movement to the new equilibrium”).

419 *Ibid* at 819.


421 Chapter 8 will discuss this point at length.
expectation of the transition to the new legal regime by a sufficient number of people.\textsuperscript{422} Hence, centralization can support a mass migration to the new equilibrium by reducing the fear of a solitary transition, which is the basis of the Prisoner’s Dilemma. In fact, by threatening sanctions for violations of new legal standards, centralization allows the “simultaneous movement” to a superior legal regime that constitutes the \textit{condition sine qua non} of the Pareto superiority of the new legal rule.\textsuperscript{423} In brief, the promulgation of legal rules by a centralized lawmaker provides a focal point around which people can coordinate their compliance decisions.\textsuperscript{424}

Despite its informational advantages, decentralized law-making tends to exacerbate the coordination problems posed by the presence of network externalities in the production of legal rules. Whereas efficient adaptation requires information about the positive payoffs of legal change, the dissemination of information in decentralized systems can be prevented by the absence of individual incentives that signal the advantages of collective shifts to a superior norm. In practice, decentralized adaptation requires “norm entrepreneurs”\textsuperscript{425} whose incentives are strong enough to promote collective migration towards more efficient norms. However, the signalling of advantages for new legal standards is a costly activity because it reduces the possibility that the presence of norm entrepreneurs will occur in practice.

The above considerations should not lead to the conclusion that legal change by centralized law-making is \textit{not} subject to efficiency problems. On the contrary, as has

\textsuperscript{422}See Gillette, “Lock-in Effects”, \textit{supra} note 351 at 824 (“Legal rules […] are promulgated by a central authority (a court or legislature, for current purposes), and are thereafter binding on all within the authority's jurisdiction. The presence of the central authority reduces uncertainty, typically present with technological lock-in, about the willingness of others in the network to adopt the superior standard. The fact that judicial rules, once enunciated, bind all within the jurisdiction essentially transforms the Prisoner’s Dilemma into an Assurance Game.”)

\textsuperscript{423}In force of the presence of network externalities, the individual’s expected benefit of the new legal regime depends upon the number of people who will adopt it. Precisely, efficiency requires that the number of people who adopt the new rule positively affects the benefit of belonging to the new legal networks to the point that it exceeds the transition costs.

\textsuperscript{424}At this point in the discussion, I seek to demonstrate that the choice between centralization and decentralization affects the adaptive efficiency of the law-making process. Subsequently, in the context of a comparative institutional analysis, I will discuss the advantages and disadvantages in terms of the adaptive efficiency of alternative forms of centralization (e.g., legislatures, courts, agencies, and bureaucracies, etc.).

\textsuperscript{425}See Gillette, “Lock-in Effects”, \textit{supra} note 351 at at 835.
been demonstrated, centralized legal innovation is often impaired by incentive constraints. In particular, a uniform top-down legal supply entails high levels of information, bargaining, and agency costs such that centralized responses to the demand for legal change are often slow to emerge. In some of these cases, incentive constraints that affect a uniform top-down legal supply can be overcome by decentralization. The crucial point is that incentive problems affecting centralization occur under circumstances that differ from those affecting decentralized law-making. When legal innovation is affected by agency problems, decentralization can favour legal change by triggering inter-jurisdictional competition. On the contrary, when evolutionary traps\textsuperscript{426}, like those described above, preclude legal innovation, centralization might be better suited to favour legal change through the dissemination of information on the efficiency advantage of new legal regimes. In conclusion, both decentralization and centralization benefit from comparative advantages in terms of adaptive efficiency, depending upon whether the legal change is hampered by “agency problems” or by “evolutionary traps”\textsuperscript{427}.

As previously emphasized, a uniform top-down legal supply suffers from informational disadvantages when a) the legal response to transaction and coordination needs depends upon widely dispersed factual knowledge or b) requires the expansion of the actual knowledge available to the lawmaker through exploration of innovative legal solutions. In the latter case, diversified bottom-up law-making processes ensure a more efficient use of the dispersed information and generate innovative legal solutions because of parallel experimentation processes. The features of decentralized law-making facilitate adaptability through legal innovation in cases where adaptation is hampered by the lack of knowledge of new legal solutions because of dispersed factual knowledge or the need for further experimentation.

\textsuperscript{426} I borrow this expression from Cooter, “Structural Adjudication”, supra note 338 at 1687.
\textsuperscript{427} In Part III, I will discuss agency problems and evolutionary traps affecting centralized and decentralized law-making institutions.
A further issue concerning dynamic settings is related to the increasing costs of adaptability as the heterogeneity of the demand for law increases. In fact, increasing heterogeneity affects both the incentive factor and the informational problems confronted by the lawmaker. Regarding the incentive factor, the heterogeneity of preferences exacerbates the collective action problem because it increases the number of different interests to be considered when determining the content of the new legal regime. Hence, centralization, which increases heterogeneity per se, is not well suited for adaptability in cases of heterogeneous preferences for new legal regimes.

Regarding the knowledge aspect, heterogeneity in dynamic contexts poses a further problem, compared to heterogeneity in static environments. When heterogeneous local conditions are evolving, it becomes much more difficult to predict the most frequent scenario to assess the maladaptation costs of centralized law-making. Similarly, when the most efficient legal solution is not known, and its discovery is left to experimentation processes, the maladaptation costs are impossible to assess.

4.5.2. Conclusions on Adaptive Efficiency

The following conclusions can be drawn from the foregoing discussion about the relation between adaptive efficiency and the features of the demand for law.

(i) Centralization favors adaptive efficiency when legal change is hampered by the presence of strong path dependence in the production of law.

---

428 This point is well explained by see Brousseau and Raynaud, “The Economics” supra note 305 (“The impact of heterogeneous preferences is also relevant for evolving situations. As a common rule is a compromise, if changes are needed, reshaping the compromise may be complex because changes in the system of (property) rights are likely to have redistribution effects, either because there are losers and winners or because the benefits of changes are unequally distributed. If negotiations among participants in the collective order are difficult to implement, a collective action problem might emerge. […] These difficulties in modifying an existing order tend to arise with the size and the heterogeneity of a community, and they can be called dynamic maladaptation costs (opportunity costs of poor or bad adaptation to a new context). Consequently, collective orders may fail to evolve in line with changing preferences and needs. This results in discrepancies between ex ante established rules and ex post coordination needs.

429 This point is generally accepted in the literature. See, for example, Gillette, “Lock-in Effects”, supra note 351 at 835, Grundamn and Kerber, “European System” supra note 343 at 302.
(ii) Decentralization favors adaptive efficiency when legal adaptation is hampered by (a) dispersed factual knowledge, (b) the lack of knowledge of the efficient outcome and the related need for experimental or innovative solutions, or (c) when agency problems that are related to the rent-seeking pressure of organized interests groups hamper efficient legal change.

(iii) In dynamic legal settings, when the legal demand is heterogeneous, it becomes much more difficult to assess maladaptation costs.

Integrating the Analysis.

4.6.1. Institutional Dimensions and the Four Component of Process Efficiency
The analysis provided in the last two chapters has demonstrated that the \textit{ex ante} versus \textit{ex post} dimension and the degree of centralization of law-making significantly affect the four components of process efficiency. Because this is a crucial point in the analysis of the law-making process, it is worth emphasizing in the following summary of propositions.

\textit{Productive Efficiency}

(i) The \textit{ex ante–ex post} dimension of law-making affects (a) the magnitude of the compliance-information costs and (b) the structure of the production-information costs.

(ii) The \textit{centralization} of law-making entails (a) the economies of scale effect, (b) legal uniformity and consistency, and (c) efficient regulation of spillover effects across jurisdictions. The \textit{decentralization} of law-making entails (a) the incentive effects for the lawmaker due to inter-jurisdictional competition, (b) the knowledge-creating effect, and (c) the reduction of maladaptation costs.
Social Choice Efficiency
(i) The *ex ante* dimension avoids the cycling problems associated with private bargaining with *empty core*.
(ii) The centralization of law-making entails problems of preferences aggregation that are associated with the *collectivization* of the outcome.

Agency Efficiency
(i) The *ex ante* dimension poses an inter-temporal *inconsistency* problem that is independent from the degree of centralization.
(ii) The degree of centralization affects the intensity of both internal and external pressures.

Adaptive Efficiency
(i) Prospective legal rules (which are general in nature) tend to increase *resistance* costs and *maladaptation* costs.
(ii) The degree of centralization may increase the *resistance* and *maladaptation* costs. However, the degree of centralization increases the ability of the lawmaker to *provide a focal point* around which individuals can coordinate the collective migration toward superior legal rules.

4.6.2. The Interaction between Institutional and Structural Dimensions
The two fundamental *institutional* dimensions of law-making (i.e., *ex ante*–*ex post* and degree of centralization) interact dynamically with the *structural* characteristics of the regulatory environment (i.e., degree of heterogeneity and frequency of the cases to be regulated). The last two chapters have attempted to develop an integrated analysis of the interactions between structural and institutional dimensions. This analysis has enabled us to develop a four dimensional framework that serves to organize the explanatory and prescriptive analysis of alternative law-making institutions, which will be conducted in Part III of this study.
Figure 15 summarizes some of the main insights provided by the proposed four-dimensional framework. In particular, in each type of demand for regulation, the framework identifies alternative combinations of *ex-ante* versus *ex-post* and centralized versus decentralized law-making, thus exhibiting the relative advantages of process efficiency. The main result is that when the regulated environment exhibits a highly homogenous and frequent demand for legal rules, then the *ex ante* centralization of law-making is likely to enjoy significant advantages in terms of productive efficiency. This is not enough to conclude that in these cases *ex ante* centralization is the most efficient institutional law-making structure. Indeed, other institutional features of *ex ante* centralized processes should be investigated and other structural variables should be taken into account. However, given the significant economizing effects associated with *ex ante* centralization in cases of high legal homogeneity and frequency, this analytical result constitutes a useful starting point for the process efficiency analysis. Analogous considerations should be made with respect to the advantages associated with *ex post* decentralization in cases of high legal heterogeneity and low frequency. In comparison, in cases where the degrees of homogeneity and frequency generate counterbalancing effects on law-making costs (i.e., cases 3 and 4), then other structural variables should be considered to affect an efficient institutional law-making structure.

In any case, it must be emphasized that the insights provided by this four-dimensional framework of analysis should be integrated with a careful comparative assessment of the economic effects associated with the additional institutional features and structural variables identified in Part III.
Figure 15 The Four Fundamental Dimensions of Law-making

<table>
<thead>
<tr>
<th>High Homogeneity</th>
<th>Low (Aggregate) Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Homogeneity</td>
<td><strong>Ex Ante Centralized Law-making</strong></td>
</tr>
<tr>
<td>(i.e., Heterogeneity)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ex Ante Centralized Law-making** if: |
| (i) the increase in maladaptation costs associated with inefficient *ex ante* regulation of most infrequent cases does not exceed the increase in production information cost associated with the regulation *ex post* of a frequent heterogeneous demand for law; |

---

430 This table is coherent with the one provided in Figure 13. However, while the table in Figure 13 focuses on the effects associated with seven levels of heterogeneity, the table provided here (i) considers only two variables levels of heterogeneity (i.e., “high” or “low”) and (ii) includes legal frequency as an additional environmental variable condition.
4.7. Hypotheses to be Verified through Comparative Institutional Analysis

The four-dimensional framework of law-making allows us to formulate the following general hypotheses, which will be tested against the results of the comparative institutional analysis developed in Part III.

(i) Under conditions of high homogeneity and frequency of the demand for law, the ex-ante centralization of law-making allows for significant effects of economies of scale, which significantly reduce the average (production and compliance) information costs.

(ii) The ex ante centralization of law-making entails a significant increase in the agency costs of law-making.

---

431 It might be anticipated that centralization can improve adaptive efficiency unless there are conditions of a close-knittedness facilitating spontaneous law-making.
(iii) The various combinations of ex ante versus ex post and centralized versus decentralized law-making affect the structure of the social-choice problem confronted by the lawmaker. However, there are no unambiguous relative advantages of ex ante centralization over ex post decentralization and vice versa, in terms of social-choice efficiency.

(iv) The ex ante centralization of law-making reduces the magnitude of adaptive and information costs associated with legal change; however, it increases the magnitude of resistance and maladaptation costs.

In Part III, I will identify additional institutional and structural variables that affect the process efficiency of alternative law-making institutions. This comparative institutional analysis will enable the identification of core tendencies in the production of law and the assessment of the validity of the four hypotheses identified above.
In Part II, I examined advantages and disadvantages related to process efficiency using alternative combinations of \textit{ex ante} versus \textit{ex post} lawmaking and \textit{centralized} versus \textit{decentralized} law production. In addition, I investigated how these institutional dimensions interact with the degree of homogeneity and the frequency of the demand for legal rules. This analysis enabled me to formulate some general hypotheses on \textit{ex ante} centralization’s effect on the structure of lawmaking costs.

The discussion in Part III proceeds at a lower level of abstraction; it examines four alternative models of the lawmaking process: (i) politics, (ii) bureaucracy, (iii) adjudication, and (iv) spontaneous lawmaking. These simplified models do not provide a full description of real-world lawmaking institutions. However, analysing these models enables me to identify some core tendencies underlying the diversity of real-world lawmaking institutions.

The introduction to each chapter identifies the institutional structure characterizing each lawmaking process. That is, each lawmaking institution is viewed as a combination of the \textit{ex ante} versus the \textit{ex post} dimensions and alternative degrees of centralization with some additional institutional variables. \textit{Politics} is an \textit{ex ante} centralized form of lawmaking combined with the mechanism of political representation. \textit{Bureaucracy} is \textit{ex ante} centralized lawmaking combined with a lawmaker’s technical expertise and a lawmaker’s political insulation. \textit{Adjudication} is an \textit{ex post} form of lawmaking, with varying degrees of centralization, combined with three additional institutional features: (i) settlement-trial decisions by private parties; (ii) adjudication courts; and (iii) the doctrine of precedent. \textit{Spontaneous lawmaking} is characterized as \textit{ex ante} decentralization (i.e. private bargaining) repeated over time in the absence of a centralized public lawmaker.
After identifying the institutional features that are predicted to be most relevant in affecting the efficiency of law production, the line of inquiry revolves around two questions: (i) what are the mechanisms through which the additional institutional variables identified above affect the core tendencies illustrated in Part II; and (ii) what are the environmental conditions (i.e. structural variables) under which each lawmaking institution is predicted to enjoy comparative advantages relative to alternative sources of law. In following this main line of inquiry, I organize the discussion according to the analytical framework provided by the unified taxonomy of lawmaking costs developed in Part I. For each component of process efficiency, I investigate both the demand and supply sides of the lawmaking process.

The conclusion in Chapter 9 integrates the different levels of analysis developed throughout this study, thereby identifying the core thesis and its main supporting arguments.
Chapter 5
POLITICS

Everyone wants a free lunch. It is not possible for everyone to have one. The attempt by everyone to get one through legislation only ensures that everyone over-spends on lunch.

Schwartz, Vote Trading and Pareto Efficiency, 1973

This chapter examines the efficiency advantages and disadvantages associated with the political production of law. It builds on the premise that (i) “ex-ante centralization” and (ii) “political representation” are the two structural features that most influence the efficiency of the political process. The central thesis is that politics is a relatively inefficient and ineffective source of law and that the excessive politicization of the law-making process significantly undermines the substantive efficiency of the law.

A better understanding of how “ex ante centralization” and “political representation” shape law-making costs helps us to identify the structural limits of political law-making. It will become clearer as the discussion proceeds that the advantages of political law-making are associated with the ex ante centralized dimension of the legislative process while the most significant limitations result from the mechanism of political representation and from majoritarian decision-making.

The present discussion is not meant to be comprehensive. No survey of a reasonable length can do justice to the richness and complexity of the literature on the political process that has developed over the past decades. My focused purpose is to frame the issues of major relevance for process-efficiency analysis: I look at democratic politics as

432 In this chapter I refer (i) to the outcome of the political process as “legislation” or “statutes,” (ii) to the actors that operate on the supply side of the political process as “politicians,” (iii) to the actors that operate on the demand side of the political process as “voters,” and (iv) to the institutional arena of the law-making process as “legislatures.”

433 A very useful overview on different approaches to the study of politics is provided by Angelo Panebianco, L’analisi Della Politica. Tradizioni di Ricerca, Modelli, Teorie (Bologna, Il Mulino, 1989).
a “source of law” from an efficiency standpoint, and do not consider other major themes that are naturally associated with analysis of the political process (e.g., political freedom, social justice, and so on). In the pursuit of brevity and coherence many issues will be covered much more briefly than they merit.

Introduction. From the Rule of “Law” to the Rule of “Legislators”

A brief historical excursus helps to identify the essential features of political law-making. The increasing “politicization” of the production of law is a recent phenomenon. For the longest part of legal history, law has been predominantly created through customary processes and judicial decisions; it is only over the last two centuries that the written law, enacted by centralized political bodies, has assumed a predominant role in the organization of the sources of law. This change – and the resultant growing centrality of statutory law at the expenses of customary and judge-made law – has been supported by a renewed trust in political law-making along with a significant change in the conception of “law.” Because this historical passage is crucial for appreciating the structure of political law-making, it is worth devoting a few words on it.

For centuries, since the Roman tradition, the ius (the law) has been inextricably connected with iustum (what is just). Legal rules have been understood as expressing the community’s sense of justice and as reflecting widely accepted values and beliefs among society’s members. Law was qualified by its adherence to some normative values: ius is iustum. From this perspective, the role of the iurisdicatio (jurisdiction) has been conceived as that of finding the law, rather than that of making it; the law has been understood as a set of long-standing principles that provide the context for judicial activity, with the judge’s only task being to formally announce and define these pre-existing natural or fundamental rights. The ius dicere (declaring of law), and not the ius dare (creating of law), characterized the activity of jurists. We will appreciate the importance of this

---

434 Sartori, Democratic Theory, supra note 199 at 310.
conception of the law subsequently in relation to the distinction between “grown” and “made” adjudication orders.

Since the nineteenth century, the emergence of the ideas of both representative democracy and universal suffrage has drastically changed the conceptions of (1) law, (2) law-making activity and (3) law-making institutions. First, under the conception of democratic legislation, law “is no longer a fact qualified by a value (a ius that is iustum),”\(^{435}\) rather, the law is an expression of the will of the political power in forms established by the power itself. The law qualifies as a iussum (a rule enforced by a sovereign), and any sovereign’s iussum qualifies as ius. In this view, “[t]he idea of law is on the one hand restricted to the commands that bear the mark of the will of the sovereign and on the other extended to any order that the sovereign is willing to enforce.”\(^{436}\) Stripped to its essence, “a ‘will conception’ or ‘command theory’ of law gradually replace[d] the […] idea of a free law-making process derived from custom and defined by judicial decisions.”\(^{437}\) Insightfully, Leoni comments on the politicization of law occurring in the nineteenth century as follows:

The most important consequence of the new trend was that people […] accustomed themselves more and more to conceiving of the whole of law as written law, that is, as a single series of enactments on the part of legislative bodies according to majority rule.\(^{438}\)

Second, if the law is no longer regarded as something pre-existing the political power, but as the legitimate expression of the will of the sovereign, the distinction between iurisdiction and gubernaculum (the government)\(^{439}\) dissolves. Law no longer precedes politics; rather, politics precedes and produces the law. Furthermore, law-making by elected politicians has replaced law-finding by the courts as a primary source of law – i.e., the ius dare of the sovereign has come to supplant the ius dicere of the experts. Put shortly, under the new conception, the source of the ius is no longer the technical knowledge of expert jurists but rather the will of the legislator. The principle of political

\(^{435}\) Ibid. at 308  
^{436}\) Ibid. [emphasis mine].  
^{437}\) Ibid. at 310  
^{438}\) See Leoni, Freedom, supra note 197 at 147-149  
^{439}\) Sartori, Democratic Theory, supra note 199 at 311
representation explains why legislative “will” prevails over judicial “knowledge”: the law should be primarily determined by elected legislative bodies because they act as political representatives of the people. The legitimacy of law-making is based on its reflection of the will of the people expressed by faithful elected representatives and not on the discovery of pre-existing natural or fundamental rights by expert jurists. Under the legislative conception of the law, the ius is tied to the iustum by virtue of the principle of representation as opposed to the professional activity of the iurisdiction. As Leoni magisterially summarizes:

[...] the law-making process was no longer regarded as chiefly connected with a theoretical activity on the part of the experts, like judges or lawyers, but rather with the mere will of winning minorities inside legislative bodies.440

Third, this legislative understanding of the law significantly changed the conception of the institutional functions of parliaments. Modern “legislatures” were not originally designed to “make” law but rather to exert “control” over the law created by the sovereign. Parliaments were not considered a suitable institutional arena for the production of legal rules until they were transformed into “legislatures.” The first parliament was formed in England during the late seventeenth century as a collegial body with the power to represent the general interests of the community before the sovereign. In that institutional context, laws were enacted by the king with the approval of both the lords and the commons in parliament (the “King in Parliament”).441 Namely, the role of parliament was to control the power of the sovereign by consenting to the law he or she sought to enact and by voting on the issuance of taxes. Importantly, the laws enacted were a very small portion of the sources of law, which were mainly derived from customs and from the adjudication of disputes.

With the triumph of the democratic ideal and universal suffrage, parliaments acquired sovereign power as legitimate custodians of the will of the people. In this new context, the law took the form of legislation and parliaments become “legislatures.” Parliaments are no longer controllers of the law enacted by the king, but they have become the central

440 Leoni, Freedom, supra note 197 at 147-149 [emphasis mine]
441 See, Sartori, Elementi, supra 385 at 234
political institution of law and its supreme source. As a consequence, the *ius* is no longer the expression of pre-existing natural or fundamental rights, but rather it largely corresponds to the *iussum* of parliament. In essence, (i) the *lawmaker* coincides with the *sovereign*, and (ii) the *law* becomes *legislation*.

To summarize: the political law-making of our times is the result of an historical process dating back to the seventeenth century that culminated in a change “from a system based on the *rule of law* to a system centered on the *rule of legislators*.”\textsuperscript{442} This process has led to an increasing politicization of (1) the sources of law in general, (2) the very idea of law-making, and (3) the function of parliaments. It can be summarized through the following propositions:

(i) the *law* has transformed from “*ius iustum*” to the sovereign’s *iussum*;
(ii) the conception of *law-making* has changed from the expert’s *theoretical* activity to an act of *will* based on the principle of representation;
(iii) the *legislatures* have progressively evolved from *controllers* of the law into *makers* of legislation.

The preceding historical considerations introduce the main argument of this chapter. I challenge, from many different angles, the idea that legislatures are a suitable institutional arena for the production of law, at least to the extent that they function as such today. At a certain point “statutorification” begins to raise the costs of law-making without generating any corresponding improvement in the quality of the law, and – beyond that limit – it produces highly inefficient legal rules. The pages that follow attempt to identify the *economic* reasons for which the politicization of law-making fails once certain efficiency limits are exceeded.

\subsection*{5.1. Productive Efficiency}

Productive efficiency in the legislative process results from the combination of centralized and *ex ante* law-making. As we already know, *centralization* entails an

\textsuperscript{442} Sartori, *Democratic Theory*, supra note 199 at 309.
institutional separation between the lawmaker and the people subject to the law,\textsuperscript{443} which creates informational problems on both sides of the law-making process. On the one hand, the lawmaker faces the challenge of identifying the public’s preferences; on the other hand, the people subject to law make choices based on their “perception” of the lawmakers’ behaviour.\textsuperscript{444} Ex ante law-making exacerbates these informational problems. On the one hand, voters form some sort of expectations regarding the lawmaker’s future behaviour; on the other hand, when creating legal rules, the lawmaker predicts people’s future responses to legal change. The interaction between voter expectations and lawmaker predictions generates a highly complex, and still not entirely understood, informational dynamic, which significantly affects the productive efficiency of the political process. This section investigates this dynamic with the purpose of identifying the strengths and weaknesses of the legislative process as compared to adjudication and to private ordering.

5.1.1. The Demand Side

People play a dual role in the political process. They participate in the election of their political representatives as voters and, at the same time, they are subject to the law produced by legislatures. The two next subsections analyze the informational problems individuals face as both voters and people subject to the law.

5.1.1.1. Voters

Voters are limited in their ability to identify the efficient production of legal rules and to promote it through their voting choices. Economic analysis has long identified the characteristics of the electoral process that undermine (at the aggregate level) the voters’ capacity to promote efficient social choices: voters (i) deal with a Knightian, or immeasurable, uncertainty, (ii) do not have incentives to become informed, and (iii) have incentives to behave irrationally when casting their votes.

\textsuperscript{443} The political outcome is not only determined by the choices of the people subject to law, but it is also largely affected by lawmaker’s choices.

\textsuperscript{444} I discussed these problems in chapters 4, and will not repeat them here.
5.1.1.1.1. Voter Uncertainty

Unlike adjudication,\(^{445}\), both political and market processes are based on acts of individual will. However, there are important differences in the way the individual will finds expression in the market and in the political arena. Unlike consumer choices, political decisions are both “collective” and “collectivized.”\(^{446}\) Political decisions are \textit{collective} in the sense that they constitute the aggregate result of choices made by a large number of individuals without a face-to-face interaction.\(^{447}\) They are \textit{collectivized} in the sense that they produce effects on the whole community subject to them. It is important to distinguish these two contiguous concepts from each other.\(^{448}\) The act of political decision-making is “collective,” but its result is “collectivized,” – that is, political decisions entail a collective decision-making \textit{process} culminating in collectivized outcomes.

The fundamental differences between market decisions and political ones derive from the collective and collectivized nature of political decisions. First, while in private markets individual consumer choices determine \textit{predictable} effects on the consumer’s utility functions, in the political arena the choice of a single voter provides a \textit{marginal} contribution to a \textit{collective} decision, the final effects of which are difficult to predict. As Coleman observes: “because of the divided control [over collective choices], a much more extensive calculation using expected utility, rather than certain utility is necessary in carrying out rational behavior.”\(^{449}\) Second, the collectivized nature of political decisions involves a \textit{mismatch} between those who decide and those impacted by the choice. While in private markets the choosing entity (consumer) coincides with the recipient of the decision (consumer), in the political process the chooser (politicians) differs from the recipient (individuals). That is, the politicians elected by voters are those

\(^{445}\) As the discussion in chapter 7 will clarify, adjudication is based on reason-providing acts.
\(^{446}\) On this distinction and its importance in the analysis of decision-making processes, see Sartori, Elementi, \textit{supra} 385 at 379-380.
\(^{447}\) See James S. Coleman, “The Possibility of a Social Welfare Function” (1966) 56 The American Economic Review, 5, 1105 at 1121 (“[…] collective decisions are events in which control is formally distributed over more than one actor [...]”).
\(^{448}\) \textit{Ibid.} (“[Collective decisions tend also to be events in which any outcome has consequences for many actors] [emphasis mine]
\(^{449}\) \textit{Ibid.}
who decide the policy outcomes that affect individuals’ utility function. This characteristic separation between chooser and recipient makes it more difficult for the individual to predict the effects of his or her choices in the political arena. Third, the collectivized nature of political decisions is associated with a high degree of production \emph{indivisibility}.\footnote{See James M. Buchanan, “Individual Choice in Voting and The Market” (1954) 62 Journal of Political Economy, 334 [Buchanan, “Individual Choice”], now in \textit{The Collected Works James M. Buchanan} (Indianapolis: Liberty Fund, 1999) Vol. 1, 75-88, at 81-82 [Buchanan, “\textit{Collected Works of James Buchanan}”].} In fact, political decisions bundle together a number of heterogeneous issues; this significantly raises the cost of acquiring the necessary information and increases the complexity of the choice.\footnote{See Komesar, \textit{Imperfect Alternatives}, supra note 7 at 71 (“The more complex the social issue the more difficult or expensive it is to recognize one’s position”).} To conclude, the structural characteristics of political decisions lend a high degree of \emph{uncertainty} to the decision-making environment in which voters choose. Buchanan describes this immeasurable uncertainty as “Knightian.”\footnote{Buchanan, “Individual Choice”, supra note 450 at 77.}

It is worth underlining an interesting difference between political and adjudication processes. Both judicial and political decision-making are plagued by limited information, by bounded rationality and by consequent uncertainty. However, while the structure of the adjudication process is designed to help the lawmaker deal with the pervasive uncertainty of the decisional environment, that of the political decision-making process \textit{exacerbates} this uncertainty. \textit{In politics the uncertainty generated by the structure of the process adds to the uncertainty physiologically associated with the production of legal rules.}

\subsection*{5.1.1.1.2. Voter Rational Ignorance}

The choice of representatives in the electoral arena requires a large amount of information; but voters do not have efficient incentives to invest the time and resources required to become informed. First, the \emph{uncertainty} surrounding the outcome of individual voting decisions makes it difficult for voters to estimate the expected return from acquiring political information. Second, as noticed earlier, the production \emph{indivisibility} of the political process generates a \textit{collective action} problem that further
undermines individual incentives to acquire political information. Third, an individual’s expected benefits from participating in the political process are generally low compared to the high costs of acquiring political information. As Down\(^ {453}\) has long demonstrated, the voter incentive structure raises the problem of voter “rational ignorance.” Because of a given voter’s infinitesimal likelihood of affecting the electoral outcome, the voter has no incentive “to invest in acquiring sufficient knowledge to make an informed choice.”\(^ {454}\) It is individually \textit{irrational} to invest in acquiring political information. Therefore the electorate as a whole displays the tendency to under-invest in political information. In economic terms, “an informed electorate is a public good,”\(^ {455}\) the provision of which is plagued by a severe collective action problem.

The efficiency consequences of rational ignorance for political law-making are enormous. When individuals are not well informed (and, even worse, when they are \textit{ill} informed) they \textit{cannot} form and express genuine preferences. This “suggests that democratic election systems always operate at less-than-perfect efficiency.”\(^ {456}\) Namely, the structure of individual incentives to acquire political information “prevents the governed from expressing their \textit{true consent} when electing a government.”\(^ {457}\)

A clarification is needed on the issue of voter rational ignorance. The theory of voters’ rational ignorance does not describe the behaviour of every single individual voter with respect to the acquisition of political information. What the problem of rational ignorance really tells us is that the electorate “as a whole” is likely to be under-informed, and thus \textit{unable to properly represent individuals’ preferences}. To be clear, this does not mean that the aggregate electorate plays no role in the political process. In fact, while the acquisition of political information is a costly and difficult activity, \textit{voting} itself is a much cheaper activity; thus a number of individuals might choose to vote even if they are

\(^{453}\) Downs, “\textit{An Economic Theory}”, supra note 203 chapter 13. The problem of voter rational ignorance is the subject of a great deal of economic literature.


\(^{455}\) Ibid.

\(^{456}\) Downs, “\textit{An Economic Theory}”, supra note 203 at 247.

\(^{457}\) Ibid. at 247 [emphasis mine].
uninformed, and, undoubtedly, an uninformed electorate exerts effects on the political system. I will discuss this point below.458

The problem of “rational ignorance,” which is related to the collective and collectivized nature of political decisions, is a source of comparative disadvantages of politics versus adjudication and private orderings. Unlike political decisions, both market decisions and trial-settlement decisions are individual and individualized decisions, and as such they more effectively influence the decision-making output as compared to a voter’s influence on political outcomes. The consumer’s choice to buy or not to buy generates an immediate, predictable consequence for the consumers’ utility function. Similarly, in adjudication, litigants select the flow of information and judges decide issues brought before them on these grounds.459 In contrast, as already emphasized, the influence of a single individual’s decisions on the final political outcome is uncertain, unpredictable and negligible.

Since voter rational ignorance crucially affects the process efficiency of political law-making, it is worth briefly considering the strengths of the counterarguments that have been raised in the literature. These arguments can be grouped under two types of objections: (i) voters are not as ignorant as predicted by economic theory; (i) the consequences of voter ignorance are not as detrimental to the quality of the electoral process as economic theory predicts because voters can make correct choices despite biased information or their rational ignorance.

458 Second, the distribution of interests at stake, as well as the costs and ability of acquiring political information, vary significantly across the electorate. Uneven distribution of stakes determines different incentives across voters to acquire political information and results in an uneven voter access to the political process. This has important implications in terms of process efficiency, which I will discuss in the next section as part of the agency efficiency problems.

459 To be clear, I am not saying that there is no uncertainty plaguing market and adjudication processes. The consumer’s choice might be burdened by uncertainty in measuring the quality of products, or in assessing the reliability of the seller. Similarly, in adjudication the output of the process is difficult to predict, which generates uncertainty for litigants. What I am saying is that the structure of political decisions exacerbates the uncertainty, which is consequently less pervasive in market and adjudication than in the political-voting process.
(i) **Voters are not as ignorant as predicted by rational ignorance theory.**

Wittman has argued that “[t]he amount of information held by voters has been underestimated.”460 He provides many supporting arguments for this claim. First, political entrepreneurs have incentive to provide voters with political information because in doing so they increase their chances of being elected. Therefore, the costs of political information are not entirely borne by voters, as assumed by the theory of rational ignorance, but rather they are shared with politicians. Second, “[p]arty brand names reduce still further the cost for the voter of acquiring information” and enable voters to “make intelligent inferences.”461 Third, “[v]oters receive a lot of ‘free’ information – in the news, in the mail, and in ordinary conversation.”462

However, the preceding arguments suggest to me that the information most easily accessible to voters is biased – i.e., it is information that politicians have incentive to give to voters. Once the quality of the information provided to citizens by politicians, by political parties and by the media is taken into account, Wittman’s argument can be challenged. Depending on the degree of plurality in the political system and on the independence of the press, these sources of “free” political information might be seen as a source of electorate manipulation.

(ii) **Voters can make correct choices despite rational ignorance or biased information.**

The second series of objections dispute the negative effects of voter ignorance on the electoral process. First, the “law of large numbers” is likely to correct the majority choices.463 In political science literature this principle has been referred to as the “Miracle of Aggregation”: if uninformed, voters make independently random choices; since voters’ randomly-distributed wrong choices mutually cancel out, candidates who win are

461 Ibid.
462 Ibid.
those who convinced the majority of well-informed voters. As Caplan emphasizes, taken to the extreme the Miracle of Aggregation demonstrates that “[d]emocracy with 99% ignorance looks a lot more like democracy with full information than democracy with total ignorance.” If the 99% of ignorant voters make independently random errors, their votes will be equally distributed among candidates. As a consequence, the winning candidate is the one selected by well-informed voters.

The Miracle-of-Aggregation argument is logically sound. However, its major weakness is its reliance on the assumption of independently random errors. On the one hand, voters make “systematic” errors when making political decisions. That is, errors are not random but follow some predictable patterns. On the other hand, it has been long recognized that informational cascades pervasively characterize the formation process of public opinion, and this contradicts both the assumptions that ignorant voters make independently random mistakes and that well-informed voters make “correct” choices. The empirical weakness of these assumptions significantly undermines the predictive force of the Miracle of Aggregation.

The second argument supporting scepticism about the real consequences of voter ignorance is that people are aware of the asymmetry of information in favour of politicians. Being aware of their informational disadvantage, voters react by discounting politicians’ promises. From this perspective, voters are thought to behave as buyers in the market who react to the lack of information about the quality of the product by reducing their willingness to pay. However, this argument fails to capture the very nature of information asymmetry between politicians and voters. In politics, people often do not know how large the imbalance is. The essence of politicians’ informational advantage is

---

466 Caplan illustrates this point through the example of the choice between free-trade and protectionist policies. It is likely that some vote for protection while preferring the effect of free-trade, and some vote for free-trade while preferring the effects of protectionist policies. However, since voters generally “overestimates the benefits of protectionism” (*ibid.* at 11), the former are much greater in number then the latter. Thus, unlike the predictions of the Miracle-of-Aggregation, the median voter chooses in favour of more protection while preferring the effects of free-trade policies.
their ability to *conceal* this asymmetry and to convince voters that they are acting as their faithful representatives. Furthermore, the analogy between the market and the political arena on this point is seriously misleading. In markets, consumers have strong incentives to control information and product-quality because they *directly* internalize the costs of incorrect information and unsatisfactory product quality. By contrast, as I will clarify below, this cost internalization is absent in the political arena. In addition, consumers can easily detect the lack of promised quality and immediately punish the seller. As a result, the seller’s ability to deceive consumers is significantly constrained. By contrast, in the political arena the voter’s ability to detect politicians’ broken promises is much more limited. This point will be demonstrated in section 3 of this chapter.

To conclude on this point, the effectiveness of political manipulation lies in its being *invisible* to the voter’s eyes. Political propaganda works at a much deeper level than market advertising. Political organizations exert influence through mechanisms of group identification and ideological education. Furthermore, even if one were willing to grant that voters are able to correctly recognize the information asymmetry, to react with distrust and to discount information accordingly, this argument would not be enough to demonstrate the efficiency of the political representation mechanism. If anything, this argument demonstrates the existence of a huge *adverse selection* problem plaguing the mechanism of political representation: because of information asymmetry, voter choices do not reward *better* politicians, rather mediocre or unfaithful representatives are frequently selected.467

Third, it has been argued that voters do not actually need much political information to control politicians. Voters observe the “outcome” of politicians’ choices, and, in the case of unsatisfactory results, they punish politicians by “retrospective voting.”468 The

---


assumption underlying this argument is that voters can correctly perceive the actual state of public affairs, even if they are ignorant or biased about the effects of public policies. Put differently, voters are biased (or ignorant) when forming their preferences regarding future public policies, but they can accurately perceive the real consequences of public policies once they are implemented. If politicians implement wrong policies, the electorate will recognize this and punish politicians through “outcome-linked voting.”

Undoubtedly, the theory of retrospective voting identifies an important advantage of the political representation mechanism – i.e., “outcome-linked voting” places limits on the abuses of political power. When a politician’s misbehaviour produces consequences visible to the electorate, individuals can respond with their vote and thus eliminate unfaithful representatives. In particular, voters can employ the “grim trigger” strategy: in response to a single misbehavior by a politician, voters can defect, permanently withdrawing their support from him or her. It is as if voters compensate for their limited information about politicians’ behavior by imposing a “probability multiplier” on their punishing reaction; their “overreaction” compensates for their limited ability to “detect” defections by politicians.

However, retrospective outcome-linked voting is a very limited mechanism. As Caplan emphasizes, outcome-linked voting generates unsolvable dilemmas for politicians. Often voter policy preferences and outcomes are bifurcated such that politicians face a problematic trade-off between implementing people’s preferences in public policies (to gain consensus) and implementing policies that generate outcomes appreciated by the electorate (to keep consensus):

If leaders ignore the public’s policies preferences, they will be thrown out of office no matter how good economic conditions are. If they fully implement those preferences, though, leaders become scapegoats for poor economic performance.

---

470 Ibid.
The consequences of bad policies are rarely reflected in political outcomes “visible” to the electorate. For example, voters tend not to perceive long-term beneficial effects of public policies, and they focus on short-term costs. This creates perverse incentives for politicians to implement short-term public policies thereby generating problems in terms of adaptive-inefficiency for political law-making. A further limitation of retrospective voting lies in the difficulties for the electorate of correctly identifying the causes of political failures. Voters often ignore the complexity of the policy-making process; as a result, they may “punish leaders for problems outside their control.” In addition, because informational asymmetry is high, the probability of politicians being detected is often too low to provide them with strong incentives to pursue efficient policies. Moreover, even if those responsible for misbehaviour or bad policies were identified and punished, retribution would not make politicians fully internalize the negative externalities of their actions. Indeed, while in the market arena the breach of contract entails serious economic consequences, in the political arena the breach of promises might cost re-election or damaged political reputation, but often politicians can still continue their political career (or, in any case, they do not directly bear the costs of their unfaithful behaviour). In short, the incentives provided to politicians by the grim trigger strategy may not be enough to prevent abuses of political power, even if voters impose a multiplier to discipline misbehaviour.

Fourth, both Wittman and Caplan have contended that ignorance does not necessarily entail biased information. However, on this specific point both authors fail to take into account the full complexity of the relationship between political information and political opinion. This latter issue has been the subject of a massive amount of literature, which I cannot analyze here. However, it is worth bearing in mind that the process of individual

---

471 I will discuss this in section 5.4. of the present chapter.
473 As Caplan puts it: “A used-car dealer who gets caught lying to his customers might lose more than their goodwill; he risks a fraud conviction as well. In contrast, after he irreversibly ruins his public reputation, a politician can earn a comfortable living in a law firm. A democratically elected leader can break all his campaign promises without risking a day in jail or a one-dollar lawsuit. Heads he wins, tails he breaks even: a recipe for constant abuse.” (ibid. at 105)
474 Ibid. at 104.
political opinion-formation cannot be understood by maintaining the assumptions of rational choice theory. In a large number of cases an individuals’ political opinions are partially a result of collective social processes: individuals identify themselves with groups, and groups establish the prevailing opinion of their participants. In this respect, it has been observed that in general group identification generates “opinions without information,” in the sense that, shaped by the group context, individuals form opinions prior to receiving information, and then “defend” these ideas against (contrasting) information. In the case of group identification, the individual cognitive process tends to discard inputs contrary to the established opinion in the group. This entails biased information processing rather than rational discounting of a politician’s words.

A second important characteristic of the opinion-formation process is how opinions are linked together. This leads to the identification of another factor responsible for biased flows of opinion: ideology. Ideology is a well-organized, coherent system of ideas and beliefs, which interact with the flow of information an individual receives when forming a political opinion. The ideological mind is likely to be characterized by a strong cognitive closure, which is often associated with an intense emotional state. This generates cognitive biases leading one to select information consistent with the ideology and to reject information that cast doubt on preconceived beliefs. Thus, it is important to distinguish generic voter lack of information from alternative and more specific sources of voter ignorance. When ignorance is generated by social identification, or by an ideological cognitive closure, it usually leads to biased information and opinions rather than to mere lack of information. The mechanism generating cognitive biases is rooted in the activity of collective actors who strategically manipulate biased information. Thus, while the economic theory of voter rational ignorance simply predicts that rational voters lack incentives to acquire political information, the study of the public opinion-formation process suggests the pervasiveness of other sources of voter manipulation, such as group identification, strategically biased information, ideology, and so on. Put bluntly, rational ignorance often accompanies irrational ignorance.
To conclude, analysis of the voter incentive structure suggests that at the aggregate level the electorate is a rather unreliable actor when it comes to identifying and promoting the production of efficient legal rules. Public opinion may certainly be regarded as an effective instrument for raising the costs of inefficient or abusive policies, and this is probably the major advantage associated with this mechanism of political representation. However, when political representation is entrusted with the law-making function it is likely to become a source of pervasive irrationality plaguing the production of law. The following subsection will add supporting arguments for this conclusion.

5.1.1.1.3. Voter Rational Irrationality

The foregoing considerations suggested that the collective and collectivized nature of political decisions diminishes the individual benefits of rational, well-informed, individual choices. In this subsection, I emphasize that the consequences of the collectivization of political choice go far beyond diminishing returns from political information. The most striking effect produced by low voter decisiveness is the pronounced decrease in the material costs of making irrational choices, which dramatically increases the incentives to behave irrationally. The collectivization of the decision-making process not only discourages the acquisition of information, but it also further encourages irrational behaviour. In his book *The Myth of the Rational Voter*, Caplan convincingly demonstrated this point.476

Conventional price theory assumes that individuals make choices based on their rational beliefs. Beliefs are the necessary cognitive resources for selecting courses of action that most satisfy an individual’s preferences. Choices based on false beliefs are likely to not satisfy an individual’s preferences; it follows that the costs of mistaken beliefs amount to the loss of welfare generated by wrong choices. This marks a crucial difference between market and political processes. To explain why Caplan introduces the idea of preferences over beliefs,477 Individuals derive emotional or psychological benefit from believing what they do, independent of whether or not these beliefs hold up to rational scrutiny. In other

words, beliefs are not only a means of ensuring rational choices, as the conventional price theory assumes; they are also a source of emotional or psychological benefits, no matter how rational their content.\textsuperscript{478} It follows that irrational beliefs do not only result in wrong choices and thus loss of welfare, they also produce benefits in terms of emotional and psychological reward.

The crucial point is that while in markets individuals pay for their irrational beliefs, \textit{in politics individual voters do not fully sustain the consequences of their individual irrationality}. The vanishingly small probability that a given individual will cast the pivotal vote dramatically reduces the costs of “departure from the unbiased, rational belief.”\textsuperscript{479} Therefore, in the political voting process there is large room for accommodating the individual demand for irrationality, which is not generally available under market conditions. While in the market irrationality is costly, in the political process voter irrationality is generally costless. \textit{The trade-off between the costs and benefits of irrational belief changes enormously from the market to the political arena.} As Caplan put it: “the institutional structure of democracy makes political irrationality a free good for its ultimate decision-makers, the electorate.”

The behavioural consequences and aggregate implications of free irrationality\textsuperscript{480} are remarkable. First, while in the market arena the individual “is willing and able to live without irrationality,”\textsuperscript{481} in the political arena the individual can afford to be irrational \textit{up to the point of satiation}. An individual will consume irrationality until there is no more emotional or psychological benefit to doing so. As a result, in politics “we should expect voters to be [often] on their worst cognitive behaviour.”\textsuperscript{482} Second, despite the fact that \textit{private} costs of individual irrationality are zero, \textit{social} costs are dramatically high. In fact, because a voter’s probability of decisiveness is vanishingly small, the expected \textit{individual} cost of irrationality tends toward zero. However, at the aggregate level the


\textsuperscript{479} Caplan, \textit{The Myth of the Rational Voter, supra} note 464 at 123.

\textsuperscript{480} Costless in the sense that costs are private. As I explain shortly below in the text, the social costs of voter irrationality are enormous.

\textsuperscript{481} Caplan, \textit{The Myth of the Rational Voter, supra} note 464 at 133.

\textsuperscript{482} \textit{Ibid.} at 132.
expected costs of irrationality are not weighted by the zero probability of decisiveness. Let me assume for simplicity that individuals have identical preferences in the matter of beliefs (i.e., they share the same beliefs). In this case, if everyone chooses to be irrational to the point of satiation, then the final collective outcome will be irrational. The result is that society makes irrational political decisions and pays for it without any corresponding social benefit. While the private costs of irrationality are zero and the private benefits are positive, the social costs of irrationality are positive and the social benefits are zero. In conclusion, the collectivization of individual choices generates privatized benefits and socialized costs of individual irrationality.

Of course, the assumption of identical preferences in matter of beliefs is not realistic. In real-world situations voters have heterogeneous preferences over beliefs and different degrees of irrationality. However, the preceding considerations capture a tendency toward irrational beliefs that is at work in democratic decision-making.

5.1.1.2. People Subject to Law

One of the main arguments in favour of legislation is that it enhances the certainty of law. Once again, to assess this claim it is useful to unpack the ex ante-centralized component of legislating from the mechanism of political representation. As clarified in chapter 4, centralized law-making enhances legal certainty. In addition, the general, forward-looking perspective of statutes, which are based on presumptive, general, “exclusionary rules”, reduces the costs of applying statutory rules to single cases. Thus, if we

483 As Caplan repeatedly emphasizes (ibid. at 145), it is important to recognize that even moderate individual preferences for irrational beliefs generate dramatically high aggregate costs of irrationality. Because of the indivisibility of the political outcome, the aggregate demand for irrationality is not continuous, despite the fact that individual rationality might be continuous. In fact, as long as the voter’s probability of influencing the final electoral outcome is low (and in a large electorate it is close to zero), moderate individual preferences for irrational beliefs are enough to determine irrational voting by all. Since political decisions are often discrete, this might result in highly costly irrational outcomes.

484 Ibid. at 14 (“Sensible public opinion is a public good. When a consumer has mistaken beliefs about what to buy, he foots the bill. When a voter has mistaken beliefs about government policy, the whole population picks up the tab.”)

485 I will clarify the concept of exclusionary rule in chapter 7 when discussing the adjudication process. Here it is sufficient to anticipate that the exclusionary rule is the logical rule (i.e., a rule disciplining the legal reasoning) that excludes some factual circumstances from the set of those excluding the application of a legal rule. Thus, if a factual circumstance is covered by the exclusionary rule it means that it falls into the scope of application of the legal rule.
consider the ex ante centralized dimension of legislation there is certainly some truth into the claim that legislation is a certainty-enhancing mechanism. Legislation enhances legal certainty to the extent that greater legal uniformity reduces the costs of appreciating the content of legal rules. Indeed the need for legal certainty has been one of the main arguments justifying the vast historical process of codification of private law in civil law systems in the nineteenth century.\textsuperscript{486}

However, it must also be recognized that in political law-making ex ante centralization hinges on the mechanism of political representation, which produces incentives that are deleterious to the provision of legal certainty. Due to the ongoing rent-seeking pressures that plague the political process, legislation can provide at most short-term legal certainty, but long-run legal certainty remains illusory or impossible to achieve through the legislative process. This stems from the fact that stability of legislation depends on the will of legislators who are only in power for short periods, and whose incentives to change or to maintain existing laws are highly volatile and dependent on the balance of power among competing interest groups. Leoni strongly emphasized this point:

\begin{quote}
[...] according to our [legislative] system, nobody can tell whether a rule may be only one year or one month or one day old when it will be abrogated by a new rule. All these rules are precisely worded in written formulae that readers or interpreters cannot change at their will. Nevertheless, all of them may go as soon and as abruptly as they came. The result is that, [...] we are never certain that tomorrow we shall still have the rules we have today.\textsuperscript{487}
\end{quote}

In this sense legislation protects people’s expectations less than judge-made laws, which are anchored in a body of legal precedents and whose modification is subject to the constraints of reason-based decision-making.\textsuperscript{488} Centralized law-making might (in the best scenario) increase legal uniformity, but the mechanism of political representation generates incentives detrimental to long-run legal certainty:

The certainty of the law, in the sense of a written formula, refers to a state of affairs inevitably conditioned by the possibility that the present law may be replaced at any moment

\textsuperscript{486} The principle of legal certainty has been celebrated by Max Weber as a precondition of capitalism; see David M. Trubek, “Max Weber On Law And The Rise Of Capitalism” (1972) Wis. L. Rev., 720
\textsuperscript{487} Leoni, Freedom, supra note 197 at 76
\textsuperscript{488} In chapter 7 I will discuss at length the nature and characteristics of the “reason-based” judicial decision-making.
by a subsequent law. The more intense and accelerated is the process of law-making, the more uncertain will it be that present legislation will last for any length of time. Moreover, there is nothing to prevent a law, certain in the above-mentioned sense, from being unpredictably changed by another law no less "certain" than the previous one. Thus, the certainty of the law, in this sense, could be called the short-run certainty of the law.\(^{489}\)

These considerations suggest that the benefits in terms of greater uniformity of legislation are likely to be largely outweighed by the lack of long-run legal certainty. In fact, in the presence of legislative uncertainty, private economic calculations become difficult and reliance on private contracts becomes risky.

5.1.2. The Supply Side

The theory of economic calculation explains that there is no centralized decision-making agency capable of identifying an efficient outcome independently of the existence of an operating competitive market process.\(^{490}\) The reason for this is straightforward: the existence of a market process is an essential requisite for economic calculation. Since it is impossible for an external observer to estimate individual marginal utility functions and to establish relative prices, prices are an essential source of knowledge necessary for identifying the efficient allocation of resources. I have already illustrated these arguments in the previous chapter so I will not rehearse them here. However, as the discussion proceeds, it is important to keep in mind this fundamental criticism affecting centralized legislative processes.

It is not necessary to subscribe to Austrian economic theory to recognize the informational problems facing the legislative process. The following statement by Scholzman and Tierney – who are not Austrian economists but political scientists studying political institutions – provides a useful reference point for describing the dramatic informational problem confronting legislators:

Legislators’ greatest need is information. [Politicians] are confronted with a staggering number of complex issues about which they are expected to make informed judgements. In the face of these complexities, members of Congress need all the help they can get in trying

\(^{489}\) Leoni, Freedom, supra note 197 at 81.

\(^{490}\) The theory of economic calculation problem has been developed by the so called Austrian School of Economics, see supra note 366. I briefly discuss this important economic theory also in sub-section 7.1.2.1.1.
to determine the consequences of their assorted legislative decisions – who will be affected, in what ways, to what extent, and with what reaction. The kind of detailed, up-to-the-minute information that legislators need is often difficult to obtain and expensive to collect.

Adjudication, bureaucracy and politics have in common the information problem described above. What differs across these alternative law-making processes is the institutional mechanism that helps the lawmaker deal with the information problem. Judges can rely on both the doctrine of *precedent* and some varying degree of *decentralization* of the adjudication system. Bureaucracies are technocratic law-making bodies characterized by the presence of *experts* who enable them to draw on substantial cognitive resources. Legislative processes deal with the information problem through a variety of mechanisms, including: (1) *ex ante* centralization of law-making; (2) information provided by interest groups, (3) the committee system and the floor decision-making process, (4) delegation to agencies. Now, with respect to the legislative process, the central question is *whether these information mechanisms are likely to lead political law-making toward the efficient production of legal rules*. I argue that although these mechanisms certainly help legislators to deal with otherwise prohibitive information costs, they are unlikely to lead political law-making to produce efficient legal rules. First, the informational advantage of the more powerful and organized interested groups within the political arena is likely to distort both legislative and electoral processes. Second, the committee system often represents a monopolistic source of information captured by interest groups. Third, delegation to agencies entails various problems of shirking and slippage by bureaucrats and of strategic delegation by politicians. The discussion in the present and the following chapter will examine the last three institutional arrangements mentioned above (committee and floor decision-making, interest groups, delegation) in order to identify obstacles to the efficient political production of legal rules.

Before proceeding let me clarify the relative advantages entailed by the *ex ante* centralization of the legislative process. As largely discussed earlier, *ex ante* centralization allows for significant economies of scale under conditions of legal homogeneity and frequency. However, these cost-saving effects do not necessarily entail

---

relative comparative advantages over the other sources of law. First, the members of legislatures have limited technical and scientific expertise. When the production of efficient law requires scientific or technical knowledge, legislatures are likely to be largely inefficient institutions for making law. In these cases, from the narrow point of view of productive efficiency, technocratic law-making enjoys a comparative advantage over political law-making, as technocrats can gather and process relevant information at lower cost. Technocratic centralization is more than a cost-saving technique; it also operates as an accuracy-enhancing mechanism.\(^{492}\) Second, when the legal demand is characterized by highly dispersed factual information, the centralized structure of legislative decision-making turns into a comparative disadvantage with respect to more decentralized forms of law-making. To summarize, the legislative process is likely to be inefficient in regulatory environments characterized by (i) highly technical or (ii) highly dispersed information.

**Conclusions**

In conclusion, “[p]olitical irrationality is not an *ad hoc* anomaly, but a predictable response to the unusual incentives” characterizing the electoral-voting process.\(^{493}\) These perverse incentives are largely determined by the collective and collectivized nature of political decisions. Voters (i) make choices in conditions of Knightian uncertainty, (ii) have incentives to remain ignorant, and (iii) do not face the costs of indulging in irrational beliefs. Uncertainty, ignorance and irrationality are highly detrimental to the productive efficiency of political law-making.

In this respect, the political process suffers from a comparative disadvantage over litigation and private ordering. In the adjudication process, both inputs and outputs are individualized and have a strong private good component. Therefore, litigants have strong incentives to acquire large amounts of information in an attempt to influence the adjudication outcome. Moreover, they directly internalize the costs of errors in the litigation process. Finally, the reason-giving nature of the adjudication process limits the possibility for both judges and litigants to manipulate the flow of information. In markets,

\(^{492}\) Centralization plays a different role in technocratic law-making as compared to politics.

consumers internalize the costs and benefits of their actions, and have in general a higher control over the seller’s behaviour compared to voter control of politicians’ behaviour. In essence, voting seems to work much better as a limit to the abuse of political power rather than as an incentive toward the efficient production of law.

5.2. Social Choice Efficiency

For the purposes of this study, the social choice problem can be summarized as the following challenge: to produce a legitimate stable outcome that reflects the preferences of the people subject to law while, at the same time, minimizing the bargaining and strategic costs of law-making. On this point, there is a vast and highly sophisticated body of scholarship,\(^{494}\) which I cannot adequately review here. In the interest of brevity and clarity, I focus on one specific point that is most relevant for the purposes of process-efficiency analysis: identification of the factual predicates under which majority rule is more likely to generate efficient, stable and legitimate outcomes. Stated more simply, I investigate the factual conditions under which majoritarian decision-making is conducive to outcome efficiency.

5.2.1. Unanimity versus Majority Rule\(^{495}\)

Unlike decentralized bargaining processes, which are based on the consent of all individuals involved in the interaction, political decision-making is generally based on the majoritarian principle.\(^{496}\) Majority rule enables the lawmaker to reduce the “hold out” problem and other “strategic costs” typically associated with the search for unanimous

---


\(^{495}\) The allocation of the law-making power to a relatively small, centralized, representative political body brings about a reduction in bargaining costs. See Wittman, *The Myth of Democratic Failure*, supra note 460 at 66, and Cooter, *The Strategic Constitution*, supra note 219 at 175. However, the reduced number of political representatives cannot explain the reduction of bargaining costs. The number of voters in modern legislatures is generally too high to reach a unanimous consensus. If we want to identify the efficiency advantages of political law-making we have to look at the structure of the decision-making process.

\(^{496}\) More precisely, unanimous consensus or qualified majorities are generally required at the constitutional stage of choice, where individuals face higher expected benefits due to the greater durability of constitutions. However, at the legislative stage of choice majority rule is the common rule governing the decision-making process.
Compared to unanimity rule, majority rule enables the lawmaker to produce *legitimate* legal rules at *lower* bargaining and strategic cost. This cost-saving effect comes at the price of high outcome inefficiency, as winning majorities can impose involuntary redistributive effects at the expense of losing minorities.

The choice between majority and unanimity rules can be summarized as follows: unanimity rule prevents winning majorities from imposing political externalities on dissenting minorities, but it entails high bargaining costs associated with the hold out problem; on the other hand, majority rule reduces the bargaining and strategic costs, but it enables winning majorities to impose externalization of the costs on losing minorities. In short, the choice between unanimity and majority rules entails a trade-off between outcome efficiency and decision-making costs.

Majority rule is based on the “one-person-one-vote” principle, which has been largely accepted in modern democratic systems as a major source of legitimization of political outcomes. However, from an economic perspective, majoritarian decision-making in representative electoral bodies is the source of many of the efficiency problems associated with political law-making. The section discusses this crucial point at length.

5.2.2. “Intensity,” “Externality” and “Cycling” Problems

I identify three major sources of process inefficiency associated with the majority rule: problems of (i) “externality” (ii) “intensity”, and (iii) “cycling.”

*Intensity Problem*. The intensity problem stems directly from the one-person-one-vote principle that reflects the idea of equality for people subject to law. The principle of equality entails giving *equal* weight to each individual’s vote regardless of how strong he or she feels about the issue. This arithmetical corollary of the equality principle is considered essential by mainstream political thought to the procedural legitimacy of the majoritarian outcome. The majoritarian outcome takes legitimization from the “arithmetical” strength of the selected options, provided that each single vote is counted

---

as equal to the others. Once the value of the votes is *levelled* (“one person one vote”), the arithmetical strength of the selected option appears to be a reasonable and efficient outcome-legitimizing criterion. Levelling the arithmetic value of votes saves strategic and bargaining costs but prevents the decision-making process from capturing the *intensity* of individual preferences. This is a major criticism of majoritarian decision-making. When the minority holds stronger preferences than the majority, the majoritarian outcome may not ensure that the aggregate benefit enjoyed by the community exceeds the costs imposed on dissenting minorities. There is no assurance that the majority outcome will not be substantively inefficient.498

*Political Externality.* Majoritarian decision-making enables majoritarian groups to impose *involuntary redistributions* on dissenting minority groups. Buchanan and Tullock have made this point clear:

[…] majority decision-making (or any decision-making with less-than-unanimity rules for choice) will tend to produce some asymmetry in gain-sharing among the individual members of the group for which the choices are made. The members of the effective coalition will receive differentially larger shares of the benefits expected to result from collective action and/or they will bear differentially smaller shares of the costs of collective actions providing general benefit for the whole group. This amounts to saying that *redistributive* elements must be a part of any collective decision reached by a less-than-unanimity rule.499

Ironically, the inequality among votes that the principle one-person-one-vote would seek to eliminate is not cancelled; it is rather shifted from *ex ante* decision-making to the *ex post* allocative consequences of the vote, which are entirely disconnected from the intensity of preferences. Namely, votes are levelled *ex ante* for the purposes of faster decision-making but, once the votes are cast, those expressed by the majority member generate a higher benefit compared to those expressed by the members belonging to the losing minority. Majoritarian decision-making is a winner-takes-all game; the voter belonging to the losing minority does not obtain any utility from the expressed vote. Thus, the *inequality* of treatment among voters is not eliminated; it is just severed from

---

498 This is largely recognized by economic theory, according to which “the condition for the Pareto optimality of the supply of public goods requires information on the relative intensity of individual preferences” (Mueller, *Public Choice III*, supra note 494 at 104), so that “it is not particularly surprising that the outcome under majority rule may not satisfy the Pareto-optimality condition” (*Ibid.*).

the intensity of individual preference. If we consider the majoritarian voting procedure as a production process transforming inputs (i.e., individual preferences) into outputs (i.e., social preferences), it is easy to see that by adopting the majority rule output inequality persists (winners and losers); it is merely disconnected by input inequality (preference intensity differential between individuals). The one-person-one-vote principle neither removes nor legitimizes inequality; rather, it is a *fictio iuris* (legal fiction) pretending to conceal preference inequality through an artificial levelling of votes.

In economic terms, the involuntary redistribution legitimized by majoritarian decision-making is a negative externality imposed on dissenting minorities. This may also be called a “political externality.” In general, political externalities occur when individuals who are entitled to participate in a political decision-making process are outvoted according to the existing decision-making rules. In the case of majoritarian voting, the political externality is a “structural” feature of the decision-making process rather than one possible outcome.

To be clear, the presence of political externalities does not implies the inefficiency *tout court* of majoritarian decision-making. The unanimity rule allows blocking political externalities for dissenting minorities, but this safeguard against involuntary redistributions comes at the price of much higher bargaining and strategic costs. Thus, as illustrated earlier in chapter 2, when the bargaining-cost saving effect of majority rule plus the outcome aggregate net-benefit effect exceed the magnitude of the external costs imposed on dissenting minorities, majority rule is likely to enjoy a relative comparative advantage with respect to alternative decision-making criteria.

*Intransitivity Cycle.* Majoritarian decision-making entails the risk of outcome instability. As social choice theory has long recognized, there is no guarantee that individually rational (i.e., transitive) preferences yield collectively rational social preferences. When groups of three or more individuals decide among three or more options through a sequence of unlimited direct pairwise comparisons, intransitivity of collective preferences might occur. In such cases, social intransitivity generates *cycling* over options: the group
is not able to reach a first-majority choice, because for any chosen outcome “some majority always prefers another outcome.”\textsuperscript{500} Intransitive social preferences are likely to result in a loss of social welfare.\textsuperscript{501}

As will become clear, intransitive collective preferences constitute a serious threat to the \textit{legitimacy} of a given political outcome. Stability and legitimacy are both conditions of substantive efficiency. This implies that the political lawmaker often confronts a problematic trade-off between two opposing inefficiency effects: inefficiency due to a lack of legitimacy, and inefficiency due to a lack of stability. Great outcome stability might entail less outcome legitimacy, and conversely.

Before proceeding with the discussion, it must be recognized that the three problems associated with majoritarian decision-making are inextricably related to each other. The indivisibility of the majoritarian outcome determines the insensitiveness toward preferences intensities, which results in the imposition of political externalities on dissenting minorities. In turn, as Bernholz has demonstrated, the presence of negative externalities is the precondition for cycling preferences\textsuperscript{502} (i.e., it is the presence of a negative externality that generates incentives for the outvoted minority to attempt to destabilize the existing majority). Therefore, the overall inefficiency of majoritarian decision-making is generated by the following causal sequence:

\textbf{Intensity Problem} \rightarrow \textbf{Political Externality} \rightarrow \textbf{Intransitive Cycle} \rightarrow \textbf{Loss of Social Welfare}

Once the efficiency problems associated with majority rule have been identified, the next crucial step is identifying the \textit{factual} predicates that either exacerbate or dilute the three problems identified above. In this respect, social-choice literature focuses on three

\textsuperscript{500} See Maxwell L. Stearns and Todd Zywicki, \textit{Public Choice Concepts and Applications in Law} (St. Paul: Thompson Reuters, 2009) at 105 [Stearns and Zywicki, “Public choice”].
\textsuperscript{501} Cooter, \textit{The Strategic Constitution}, supra note 219 at 41.
variables: (i) the structure of individual preferences, (ii) the nature of the decision issue, and (iii) the sincere or strategic behaviour of participants in majoritarian decision-making. The remaining part of this section is devoted to the functioning of majoritarian decision-making dependent on these three varying issues.

5.2.3. Majority Equilibrium

I proceed by analyzing four different contexts in which majority rule operates: (i) one-stage voting with majority rule equilibrium; (ii) unlimited pairwise voting with majority rule equilibrium; (iii) no majority rule equilibrium with vote-trading; (iv) no majority rule equilibrium with structurally induced equilibrium. For each situation, the effect of the three factual predicates identified above is considered (individual preferences, nature of the issue, sincere or strategic behaviour of voters).

5.2.3.1. One-Stage Voting

Let me consider an elementary democratic voting process under the following simplifying assumptions:

(a1) One-stage election in which two candidates, “Liberal” and “Conservative,” compete for the votes of three citizens “A,” “B” and “C.”

(a2) Candidate policy-platforms and voter preferences can be ranked on a single normative dimension scale, from the extreme “right” to the extreme “left.” The median platform/preference is at the “centre” of the normative spectrum.

(a3) Voter ideal points are evenly distributed along the single-dimensional normative scale – e.g., A prefers right, B prefers center, C prefers left.

(a4) Voters choose candidates whose political platforms are the closest possible to their ideal point - i.e., platforms yield the maximum utility.

(a5) Candidates choose the political platform that allows them to win the elections - i.e., obtain the majority of votes. Full electoral participation is assumed.

5.2.3.1.1. The Median Voter Equilibrium

Social choice literature has long demonstrated that under the preceding assumptions the choice of candidate converges on (or nearest possibly to) the platform corresponding to
the ideal point of the *median* voter.\textsuperscript{503} In this example, both Liberal and Conservative will converge on the centre platform. This result is called “Black’s Theorem” or the “Median Voter Theorem.” The underlying logic is easily understood. If one of the two non-median candidates choses an “extreme” platform, while the other choses the centrist platform, the latter wins and the former loses. In our example, the centre platform obtains the vote of B plus the vote of A (if the chosen non-centre platform is left) or C (if the chosen non-centre platform is right). The generalizable conclusion is that for any 2n+1 voters the median has (n-1)/2 voters on the left and (n-1)/2 voters on the right of the single dimensional normative scale. That is, the decision-making process gravitates toward the mean platform.

Two analytical points are worth emphasizing. First, Black’s Theorem captures a central tendency of political law-making. *When voter preferences are evenly distributed on a one-dimensional space, the simple majority rule tends to generate an equilibrium corresponding to the median voter.* Second, no first-majority choice among voters is assumed. The median platform results in a majoritarian outcome not because it is the first choice of the majority of voters, but because rational voters whose ideal point lies at the extremes of the political spectrum vote, as a “second best” choice, for the platform closest to (albeit more or less far from) their ideal point.

A closer look at the underlying assumptions reveals the practical relevance of Black’s Theorem. While (a1) (one-stage election) and (a4-a5) (rationality of political actors) are plausible and fairly realistic assumptions, (a2) (the single-dimensional issue) and (a3) (evenly distributed preferences) are more problematic. Certainly, there are cases of single-dimensional choices in which voter preferences are relatively evenly distributed along the spectrum of possible policy positions.\textsuperscript{504} In these cases, social choice predicts that the political process tends to generate a median voter equilibrium despite the absence of a first-majority choice. However, there are many other cases in which the contested political


\textsuperscript{504} Stearns and Zywicki, *Public Choice*, supra note 500 at 133
issues are *multidimensional*, and voter preferences are *unevenly* distributed. In such cases, majority rule might lead to a stable outcome through either a very restrictive set of individual preferences, or through equilibrium-inducing institutional mechanisms. I will discuss these more problematic cases below.

5.2.3.1.2. Efficiency of the Median Voter Platform

Having clarified the assumptions under which the median voter equilibrium holds, I now inquire into its *efficiency* properties according to three substantive criteria identified in chapter 1: Pareto optimality, Kaldor-Hicks efficiency and social welfare. Figure 16 illustrates graphically the voter utility curves for the previous example.

---

505 This part of the discussion is largely based on Cooter, *The Strategic Constitution*, *supra* note 219 at 31-37.
Figure 16: Single Peaked Voter Utility Functions

Utility

Platforms

$U^C(x)$
$U^B(x)$
$U^A(x)$

$X(a)^*$  $x(b)^*$  $x(c)^*$
**Pareto Optimality.** The median voter platform is represented by x(b)*. Recall that an allocation is defined as Pareto optimal when it is not possible to move away from it without making someone worse off. This is the case of x(b)*. Once x(b)* is reached, any movement to the right makes A better off, but B and C worse off; on the other hand, any movement to the left makes C better off, but B and A are worse off. In conclusion, the median voter platform is Pareto optimal.

**Kaldor-Hicks Efficiency.** As a consequence of the intensity problem, there is no guarantee that the median voter platform will maximize the aggregate surplus as required by Kaldor-Hicks efficiency. The one-person-one-vote principle, as explained earlier, does not allow for the cardinal registering of individual preferences. Only the ordinal ranking of preferences counts. All that majority rule captures is the number of individuals who prefer one outcome to its available alternatives. Bluntly put, majority rule is insensitive to how strongly individuals feel about alternative options.506

Turning to our previous example, suppose that Conservative chooses the median platform x(b)* and that Liberal choses the platform x(a)*. In a one-stage majority vote, both B and C prefer x(b)* to x(a)*, and only A prefers x(a)* to x(b)*. Thus, the selected majority outcome is x(b)*. However, if we consider preference intensities, it might be that both B and C moderately prefer x(b)* to x(a)*, while A strongly prefers x(a)* to x(b)*. In this case, majority voting does not register the ratio of the utility loss imposed on A to the utility gain enjoyed by B and C. That is, there is a mismatch between majoritarian decision-making and cost-benefit analysis. As Cooter incisively puts it: “Majority rule counts voters, whereas cost-benefit analysis adds individual values.”507

506 As I have explained earlier, KH efficiency criterion raises many problems. The most serious flaw is that it requires an inter-personal comparison of utility that in turn requires cardinal registering of individual preferences. The discussion on the majority rule confirms the difficulties of measuring cardinal preferences and thus of applying the KH efficiency criterion.

Figure 17. The Intensity Problem
Figure 17 illustrates the intensity problem. A prefers x(a)* to x(b)* much more strongly than B and C prefer x(b)* to x(a)*. In utilitarian terms, the move from x(a)* to x(b)* determines a utility loss for A \[ U^A(x(a)^*) - U^A(x(b)^*) \] that is larger than the sum of the utility gain of B and C \[ (U^B(x(b)^*) - U^B(x(a)^*)) + (U^C(x(b)^*) - U^C(x(a)^*)) \]. Since majoritarian rule cannot register the cardinal preferences of individuals, the selected majority outcome remains the median platform x(b)*, despite the fact that A prefers x(a)* more strongly than B and C prefer x(b)*. As a result, the majoritarian outcome generates a loss of social utility equal to the difference between the sum of B’s and C’s utility gain and A’s utility loss:

\[ (U^B(x(b)^*) - U^B(x(a)^*)) + (U^C(x(b)^*) - U^C(x(a)^*)) - (U^A(x(a)^*) - U^A(x(b)^*)) \]

Notice that the utility loss suffered by A represents the political externality imposed on him or her by majority rule. The externality problem associated with majority rule is a consequence of its inability to capture the differential between individual cardinal preferences.

To summarize, the K-H efficiency assessment of the median rule depends on the relative intensity of individual preferences. When preference intensity is strongly asymmetric among voters, the marginal vote does not reflect the marginal costs and benefits of the incremental shift in political platforms. In this case, majoritarian logic and efficiency logic bifurcate. By contrast, when preference intensity is symmetrical (i.e., it is equal across voters) the additional vote reflects the marginal benefit and costs of the incremental shift in political platforms. In this special case, majoritarian and (Kaldor-Hicks) efficiency logic converge toward the same equilibrium that maximizes the aggregate surplus. In short, in order for the median platform to result in a Kaldor-Hicks efficient allocation of resources an additional assumption to those underlying Black’s Theorem is required: not

\[ ^508 \text{I am assuming here the perfect divisibility of political platforms, and, consequently, the continuity of the related marginal social cost and benefit functions.} \]
only must the ordinal preferences be *evenly* distributed across voters, but individual preference *intensity* must also be *symmetrical* across voters.

*Social Welfare.* A similar line of reasoning can be applied to social welfare. The majoritarian decision-making process maximizes the aggregate social welfare if the costs and benefits registered by the median voter reflect the *mean* value of the costs and benefits to be allocated. Of course, this is quite a restrictive condition, as it presupposes homogenous intensity of preferences across voters, while in real-world situations the set of individuals involved in the political decision-making process is usually characterized by subsets of differentially interested persons. In practice, the social welfare impact of majoritarian decision-making will depend on the relative size of subsets with voters possessing different relative intensities.

### 5.2.3.2. Pairwise Voting

In the previous example, according to (a1) and (a5), three voters A, B and C are engaged in one-stage majority voting to elect candidates by choosing political platforms. Now, I drop these assumptions and consider a situation with *no candidates*, in which A, B and C have to choose *directly* among three alternative political platforms x(a), x(b) and x(c). The remaining assumptions – (a2) (single-dimensional issue), (a3) (evenly distributed preferences), and (a4) (rational voters) – are maintained. Under these conditions it easy to predict that each voter will choose the political platform closest to his or her ideal point, so that no option obtains the majority of votes. That is, *majority rule does not lead to any majoritarian equilibrium*.

In situations like these, majoritarian logic can be applied as an alternative to one-stage majority voting through a sequence of *binary* comparisons between alternative options. Since pairwise voting is frequently used in real-world legislative decision-making processes it is worth examining its social-choice properties.

#### 5.2.3.2.1. The Condorcet Winner

In an *unlimited* sequence of pairwise voting, the option that defeats all the available alternatives is referred to as the “Condorcet Winner” (hereinafter, CW). It is important to
clarify: (i) the normative status of the CW, (ii) the procedural criterion for identifying it when available, and (iii) the conditions for its availability.

**Normative Status.** The Condorcet Winner has an important normative characteristic: despite not being a majority *first* choice, it defeats all available alternatives in direct binary comparisons. Thus, from a normative standpoint, selecting the Condorcet Winner is a method of social choice *more* consistent with majoritarian logic than with other non-majoritarian criteria (e.g., plurality rule). In that sense, the Condorcet criterion enables the decision-maker to apply majoritarian logic outside the one-stage election context, in which only the first choice majority options are relevant. Pairwise voting allows one to identify a majoritarian outcome when no majoritarian first-choice exists. *The Condorcet Winner is consistent with majority rule not because it is the first-choice of the majority of voters, but because it defeats all the available alternatives in direct binary majoritarian comparisons.*

**Procedural Criterion.** Crucially, an available CW can be identified *only* if the number of pairwise comparisons is *at least* equal to the number of the available alternatives for choice. Hereinafter, I will refer to this principle as to the “Condorcet criterion.” A decision-making process that does not satisfy this criterion does not enable the lawmaker to distinguish the CW from other outcomes that can be potentially defeated in suppressed binary comparisons.

**Condorcet Rule ➔ Allows one to distinguish the Condorcet Winner**

**Non-Condorcet Rule ➔ Does not allow one to distinguish the Condorcet Winner**

**Availability Conditions.** Social choice theory has long recognized that a sufficient condition for the existence of a Condorcet Winner is that every voter has *single-peaked* preferences. Preferences are single-peaked when *one* of the alternatives (i) ranked by the individual along a one-dimensional scale\(^509\) (ii) with an intermediate amount of the

\(^{509}\) I.e., there is one single dimension with respect to which alternatives can be decisively ranked.
decisive attribute (iii) is either the most or the least preferred option. When the preferences of group members are single-peaked, the decision-making rules that satisfy the Condorcet criterion can be used to identify a Condorcet Winner. On the contrary, when preferences are multi-peaked the likelihood of intransitive preferences increases dramatically.

- **Condorcet Rule** ➔ **Single-Peaked Preferences** ➔ **CW**

- **Condorcet Rule** ➔ **Multi-Peaked Preferences** ➔ **Unlikely CW**

The next subsection discusses the efficiency of the Condorcet-Winner option. Intransitivity of social preferences and multi-peakedness will be examined subsequently.

### 5.2.3.2.2. The Efficiency of the Condorcet Criterion

It is important to recognize that the Condorcet criterion is consistent with the insight provided by Black’s Theorem. Notice that, in the example illustrating the convergence of majoritarian voting toward the median platform, the preferences of voters A, B and C are single-peaked, as are the preferences generating a CW. That is, the same fundamental social choice insight is expressed in both the Median Voter Theorem and the Condorcet criterion: for groups of three or more individuals considering three or more options, if (i) everyone agrees that there is one single analytical dimension with respect to which all the alternatives for choice can be decisively ranked (i.e., the one-dimensional issue), and (ii) individual preferences are single-peaked, then majoritarian decision-making reaches a stable equilibrium. The equilibrium can be either the median platform in one-stage majority voting, or the CW in two-stage majority voting.

The CW and the median voter platform share the same efficiency characteristics. First, the Condorcet equilibrium is Pareto-optimal. By definition, any option alternative to the CW is defeated in a pairwise comparison with the CW. This is equivalent to saying that any move away from the Condorcet equilibrium is opposed by a majority of voters – i.e., it is

---

510 For a clarification of the notion of intransitivity and its implications for the legislative process, see *infra* sub-section 5.2.3.2.3.
Pareto non-optimal. Second, the CW is likely not Kaldor-Hicks efficient because, as repeatedly emphasized, majority voting is insensitive to preference intensities. Only if individual preferences are symmetrical does the CW coincide with the surplus maximizing outcome. Finally, there is no insurance that the Condorcet equilibrium will correspond to the social welfare maximizing option. Only if the CW reflects the mean cost and utility functions might it be conducive to social welfare maximization. Being based on majority rule (and thus on the one-person-one-vote principle), the Condorcet criterion suffers from intensity, cyclical and redistributive problems.

The next subsections examine the implications of majority rule in situations where conditions that ensure the presence of a CW are absent (i.e., multi-peakedness preferences and the multi-dimensionality issue).

5.2.3.2.3. Intransitive Cycle

Social choice theory has long demonstrated that when a group of at least three people chooses among at least three options through unlimited pairwise voting, the collective decision-making process might result in intransitive preferences for the group as whole. That is, intransitive group preferences occur despite the fact that the individual preferences of group members are transitive. Individual rationality might lead to social irrationality. This fundamental insight of social choice theory is known in literature as the “voting paradox.” It identifies the risk of collective intransitivity generating “cycling” majorities: for each selected outcome there is always a majority that prefers an alternative option. It is easy to recognize that the voting paradox seriously threatens the legitimacy of the decision-making process. In the legislative process, when intransitive cycling over options occurs, the majority outcome is unstable. It is as if legislators were “trapped in a revolving door with no exit.”511 In short, a legislature that cannot generate a stable outcome representing the majority of member preferences cannot produce outcomes that are legitimate according to majority rule.

5.2.4. No Majority Equilibrium

After discussing the majority equilibrium, let me identify the conditions likely to lead to the absence of majority equilibrium. Social choice theory has identified four major circumstances reducing the likelihood of an available CW: (i) the “multidimensionality” of the issue space; (ii) the “multi-peakedness” of individual preferences; (iii) the increasing number of voters; (iv) the increasing number of alternatives. These factual predicates increase the likelihood of an intransitive cycle.

Two theoretical results have been proved that significantly limit the validity of majoritarian logic as a legitimizing decision-making criterion: the Arrow Theorem and the Chaos Theorem. They demonstrate that under plausible circumstances majority voting might lead to either an intransitive cycle or the absence of a majoritarian equilibrium.

5.2.4.1. Arrow’s Theorem

In his pioneering work *Social Choice and Individual Values* Kenneth Arrow formulates his fundamental “General Possibility Theorem”: any method for deriving social welfare functions that satisfies (i) some basic normative requirements of collective fairness cannot concurrently satisfy (ii) the conditions of collective rationality. That is, assuming transitive individual preferences, there is no aggregation method capable of ensuring socially transitive collective preferences without violating, in some fundamental ways, fairness conditions of collective decision-making.

Arrow’s main concern was measuring social welfare rather than analyzing political process. His main focus was investigating the possibility of constructing a social welfare function.
function that allows one to derive social choices from individual preference orderings. However, Arrow’s conclusions on the possibility of a social welfare function result in striking implications for the analysis of political process. Arrow’s Theorem represents “a conceptual barrier to combining individual preferences into some overall measure of social welfare.”\textsuperscript{514} If individual preferences cannot be combined into a “consistent” social ordering, it follows that there is no democratic procedure for collective decision-making capable of producing rational collective outcomes.

To be clear, Arrow’s impossibility theorem should not be understood as a descriptive or predictive analytical tool. Arrow did not expect actual instability or irrationality in the legislative processes to be observed. What he feared was that actual legislative outcomes would reveal themself (upon closer analysis) to be path-dependent and, as such, subject to strategic manipulation by agenda setting controllers. Arrow’s Theorem challenges the legitimacy of the legislative outcome because it demonstrates that individually rational preferences cannot be aggregated into socially transitive preference orderings without violating normative conditions of collective fairness implicit in normal decision-making mechanisms. Thus, one of two consequences follows: either the legislative outcome is unstable – because the principles of legitimate decision-making are fully satisfied and consequently an intransitive cycle results – or the legislative outcome is somewhat stable but not legitimate – because to ensure outcome stability some principles of legitimate decision-making were violated.

5.2.4.2. The Chaos Theorem

Consider a multidimensional choosing space with \( n \) voting legislators. Let \( M \) be the ideal point of one legislator. A line passing through \( M \) is called the median line if, considering the distribution of all other legislator ideal points, no majority is left on either side of the line. Plott has demonstrated that in a multidimensional issue space a majority rule equilibrium exists if each legislator’s ideal point can be paired with its exact opposite on the other side of the median line.\textsuperscript{515} Stated more easily: pairwise symmetry of legislator

\textsuperscript{514} Farber and Frickey, Law and Public Choice, supra note 511 at 39.

ideal points is a sufficient condition for a majority rule equilibrium to exist. This result is known as the “Plott condition” for majority equilibrium, according to which every small movement of an ideal point away from pairwise symmetry precludes the possibility of a majority equilibrium outcome. Thus, as many scholars have recognized, pairwise symmetry is quite a restrictive condition for a stable majority outcome. Its actual occurrence becomes more and more unlikely as the number of issue dimensions and voters increases.

If the Plott condition is not met, majority rule in a multidimensional issue space is globally intransitive (even if one assumes multi-peaked preferences in each dimension). McKelvey and Schofield have demonstrated that if the voting issue is multidimensional, then the social preference ordering resulting from a regime of pairwise majority voting is intransitive over all outcomes in the issue space. That is, for every point in the issue space there is another point that the majority prefers to it. This result, which is known as the “Chaos Theorem,” has three important implications for the legislative process. First, the cycling properties of majority rule might lead to any outcome in the whole set of possible alternatives. Since each point is majority preferred by another point, majoritarian decision-making can wander anywhere in the issue space. Consequently, it is impossible to predict the legislative outcome on the basis of voting actor preferences. The result is chaos and unpredictability in the legislative process. Second, if a monopoly agenda setter exists, he or she can lead the decision-making process to any outcome in the issue space. Namely, “it is theoretically possible to design voting procedures which, starting from any given point, will end up at any other point in the space of alternatives.” I will examine this important implication subsequently. Third, since no majority equilibrium exists (i.e., there is no outcome that defeats all the alternatives in a unlimited series of pairwise votes), every stable outcome cannot be justified on the basis of majority rule. Notice that this latter implication is consistent with Arrow’s theorem.

---

517 McKelvey, “Intransitivities”, supra note 516 at 472.
5.2.5. Observed Institutional Stability

Social choice scholars have long recognized that the empirical observation of real legislatures appears to contradict the insights of social choice theory. In fact, in many jurisdictions, the political process exhibits a substantial level of political stability. This stability is ensured by the legislative processes despite policy space multi-dimensionality, preference multi-peakedness and the high number of voters, which constitute the real conditions under which modern legislatures operate. The observed outcome stability raises the question of “why there is so much stability.”\textsuperscript{518} Social choice scholarship has demonstrated that two major mechanisms explain the stability of the law-making process in modern legislatures. First, many institutional features characterizing legislative processes lead to the suppression of cycling preferences and the production of stable outcomes. That is, the observed equilibrium is “structurally induced” rather than purely “preference induced.”\textsuperscript{519} Second, under certain conditions the practice of vote-trading (or “logrolling”) enables the law-making process to account for preference intensities. This sensitivity toward preference intensities reduces problems of both externality and cycling.

The next two subsections briefly discuss the practice of logrolling and the use of equilibrium-inducing mechanisms as the two major institutional stabilizing devices.

5.2.5.1. Vote-trading

As noted above, majority voting allows for no cardinal expression of preference intensity, but only preference ordering among alternative outcomes; this results in loss of welfare (i.e., the externality problem) and outcome instability (i.e., the cycling problem). Logrolling proponents have long argued that vote-trading “allows voters to express their varying intensity of concern over issues”\textsuperscript{520} thereby reducing both the externality and the cycling problem. However, this point remains highly controversial in social choice literature. In this subsection, I briefly identify the conditions under which logrolling can

\textsuperscript{519} This does not mean that the structure of individual preferences does not affect the outcome. It signifies that the outcome varies depending on the structure of the institutional mechanisms aggregating individual preferences. By contrast, a preference-induced equilibrium does not vary depending on the structure of the aggregating mechanism.
\textsuperscript{520} Nicholas R. Miller, “Graph-Theoretical Approaches To The Theory Of Voting” (1977) American Journal of Political Science, 769 at 52.
improve the efficiency of political law-making, and emphasize the limits of the market for votes.

5.2.5.1.1. Political Externalities

Let me preliminarily identify the structure of logrolling. Let A, B and C be three individuals voting on a series of pairwise comparisons on two issues x and y. Since for each issue the possible outcomes are respectively (X, X) and (Y, Y), the set of possible political outcomes include four possibilities: (X,Y) (X, Y) (X, Y). Assume the following preference orderings:

A: XPX and YPY
B: PX and PY
C: XPX and YPY

Thus, a majority of A and C prefers X to X, and a majority of B and C prefers Y to Y. If voters choose according to their preferences the majority outcome is (X, Y) (hereinafter, the “no trade outcome”). Notice that while C is part of the majority on both issues, A belongs to the majority on x and B belongs to the majority on y.

Now, let us consider A and B, setting C aside for a moment. Let us also assume that A is more concerned with issue y than x, and that, conversely, B is more concerned about issue x than y. Namely, A prefers Y to Y more intensely than he or she prefers X to X, and B prefers X to X more intensely than he or she prefers Y to Y. The intensity of these preferences can be denoted as the utility differential for a legislator between the preferred and the not-preferred outcome. If I denote the utility of a legislator (A, B or C) for a given outcome as $u_i(\text{outcome})$, then:

for A: $|u_A(X) - u_A(X)| < |u_A(Y) - u_A(Y)|$
for B: $|u_B(X) - u_B(X)| > |u_B(Y) - u_B(Y)|$
In such a situation, A belongs to the majority on x, but he or she is more concerned with y; conversely, B belongs to the majority on y, but he or she is more concerned with x. Differently put, A prefers to win on y and to lose on x, but he or she loses on y and wins on x; conversely, B prefers to win on x and to lose on y, while he or she loses on x and wins on y.

It easy to see that, in such a situation, a vote trade between A and B would make both of them better off. Assume that A trades his or her vote for X against B’s vote for Y. Thus, A votes X on issue x, and B votes Y on issue y. The new majority outcome will be (X, Y) (hereinafter, the “traded-outcome”). In this new scenario, A becomes part of the new majority (together with B) on y, and B becomes part of the new majority (together with A) on x. Differently put, both A and B win on the issues that are of greatest concern to them and lose on issues that are less important to them. Hence, the exchange of votes allows A and B to align the majority outcome with their preference intensities thereby realizing a cooperative surplus. The size of the cooperative surplus is equal to the sum of the absolute difference between the preferred and not-preferred outcomes for both issues x and y for both A and B. In more formal terms, the cooperative surplus can be denoted as follows:

$$\left| A_u(X) - A_u(X) \right| + \left| B_u(Y) - B_u(Y) \right|$$

The absolute difference between the utility of the preferred outcome and the utility of the not-preferred outcome can be called the “salience,” so, to the legislator of an issue. Based on this definition, the cooperative surplus can be thought of as the sum of the absolute difference in the salience to A and B of issues x and y:

$$\left| s_A(x) - s_A(y) \right| + \left| s_B(x) - s_B(y) \right|$$

---

521 Ibid. at 1239.
522 So, Salience for A on x = $s_A(x) = \left| A_u(X) - A_u(X) \right|$ Salience for A on y = $s_A(y) = \left| A_u(Y) - A_u(Y) \right|$
Salience for B on x = $s_B(x) = \left| B_u(X) - B_u(X) \right|$ Salience for B on y = $s_B(y) = \left| B_u(Y) - B_u(Y) \right|$
The above example helps identify the constitutive elements of vote-trading together with its structural limitations. First, vote-trading entails a strategic vote switch. In the example, A, whose preference ordering is $XPX$, votes $X$ instead of voting sincerely for $X$; symmetrically, B, whose preference order is $YPY$, votes $Y$ instead of $Y$.

Second, the vote switch must decisively determine a change in the majority outcome. In our example, without A switching his or her vote (from $X$ to $X$) the majority outcome on issue $x$ would have been $X$, and without B switching his or her vote (from $Y$ to $Y$) the majority outcome on issue $y$ would have been $Y$. This means that vote-trading can occur only between "pivot" voters – i.e., voters whose switch is decisive in changing the outcome. At this point, we should observe a first structural limitation of logrolling as an equilibrium inducing mechanism: logrolling is limited only to votes on issues that are close to obtaining a majority and for which the success of majority voting is obtained through traded votes. Logrolling does not occur either on the issues that win by large margins nor on issues that are losing despite the vote-trading. As Mueller succinctly put it: “no trading on losing issues” and “no trading on issues that win by substantial margins.”

Third, vote-trading is “a system of barter,” which requires that A gain from B’s switch and B gain from A’s switch. Differently said, A does not pay B for his or her vote, and vice versa. Rather, the only way the trade occurs is by the exchange of votes. This significantly limits the occurrence of logrolling in real-world legislatures. The exchange of votes is rational only if A and B are on opposite sides on the two issues. A must belong to the majority on the same issue on which B is a minority, and, conversely, B must belong to the majority on the same issue on which A is a minority. Thus, we observe a

---

524 Ibid. at 1237
525 Mueller, Public Choice III, supra note 494 at 110.
526 Riker and Brams, “The Paradox” supra note 523 at 1238.
second structural limitation of logrolling: *logrolling is unlikely to occur when there is strong party discipline and individual vote switches among coalitions are rare.*

Fourth, logrolling must satisfy some obvious utility requirements. Namely, the utility each trader derives from the trade outcome must exceed the utility he derives from the no-trade outcome. That is, the issue on which each voter-trader accepts to lose must be *less* salient than the issue on which he or she seeks a win through vote-trading. In our example, where A trades his or her vote on $x$ against B’s vote on $y$, it must be that:

for A: $s_A(y) > s_A(x)$

for B: $s_B(x) > s_B(y)$

If we combine this obvious utility requirement with the limits imposed by the barter structure of vote-trading, we can easily see that A must belong to the majority on the issue that is most salient to B and on which B belongs to the minority; conversely, B must belong to the majority on the issue that is more salient to A and on which A is a minority. Thus, we observe a third structural limitation of logrolling: *it is effective only in those situations in which the preference of the minority is thwarted on an issue that is not salient for the majority* (i.e., it is *less* salient for the majority than for the minority). A different way to look at the utility requirement is to say that there must be a *positive* cooperative surplus from trading:

$$\left| s_A(x) - s_A(y) \right| + \left| s_B(x) - s_B(y) \right| > 0$$

This signifies that the relatively more salient issue for A must coincide with the relatively less salient issue for B, and, conversely, the more salient issue for B must coincide with the less salient issue for A. To the extent that voters are allowed to win on their relatively more salient issues, vote-trading increases the joint utility gain of traders. When these efficiency conditions are met, it is rational for voters to engage in vote-trading.

---

527 *Ibid.* at 1238

528 Ones that ensure the coherence of voter behaviour and the assumption of rational behaviour.
The preceding considerations illustrate that both the structural and efficiency conditions for rational vote-trading significantly limit its actual occurrence in real legislatures. Furthermore, the practical relevance of logrolling is not limited only by the restrictive conditions under which it can take place. A second order of limitations derives from its social welfare implications. The theoretical literature on vote-trading has overly emphasized its welfare-enhancing potential. Indeed, vote-trading involves political externalities at the expense of non-traders. In particular, vote-trading entails the risk of reversing the majority outcome at the expense of the majority of voters, thereby generating a diminution of social welfare. Thus it is wrong to generalize a priori the welfare enhancing effect of vote-trading. The social welfare impact of vote-trading depends on the structure of individual preferences, which must be carefully examined on a case-by-case basis. If the utility gain of traders exceeds the utility loss of no-trade voters, then vote-trading generates a welfare enhancing effect; if the cost of political externalities imposed on no-traders exceeds the utility gain of traders, then logrolling reduces aggregate social welfare.

Consider the previous example and focus on the impact of the vote exchange between A and B on C’s utility function. In the no-trade outcome, C belongs to the majority on both issues x and y; that is, the no trade outcome maximizes C’s utility function. The trade of votes between A and B completely reverses the majority outcome such that C comes to belong to the minority on both issues x and y; that is, the trade majority outcome minimizes C’s utility function. Despite C remaining passive during the entire trading process, the decision outcome is completely reversed, generating a diminishing impact on his or her utility function. Clearly, the trade of votes between A and B generates an externality impacting C. This simple example illustrates that the social welfare enhancing effect of vote-trading depends on the structure of the preferences involved in the legislative process. The obvious condition for which vote-trading improves social welfare

is when the utility gain from vote-trading exceeds the external costs imposed on the non-traders who belong to the majority under no-trade outcomes.  

Riker and Brams have demonstrated that the possibility of trading votes in the legislative arena generates a Prisoner’s Dilemma-type of situation that is potentially conducive to a highly inefficient social outcome.  

5.2.5.1.2. Cycling

Now that the social welfare implications of vote-trading have been clarified, it is worth briefly emphasizing the relation between the exchange of votes and the cycling problem.

Bernholz has demonstrated that vote-trading situations imply voting cycles and that the absence of voting cycles implies the absence of vote-trading situations. This result is also known as the “Logrolling Theorem.” Based on this theorem, social choice thinkers generally conclude that the market for votes does not solve the cycling problem. Logrolling is just a special case of the general relationship between the conclusion of contracts and the emergence of cyclical social preferences. I understand Bernholz’s results on this point as definitive, and I will not rehearse them here. However, it is worth underlining that – as Bernholz himself recognizes – the institutional framework within which vote-trading occurs plays a crucial role in determining the emergence of cyclical social preferences. Let me briefly explain this point.

As we already know, the cycling of social preferences presupposes the existence of negative (political) externalities. This fact, together with the above principle that vote-

---


532 As Bernholz clarifies, logrolling is just a special case of the more general relationship between the conclusion of contracts and the emergence of cyclical social preferences.

533 See Mueller, Public Choice III, supra note 494 at 104-112

trading situations presuppose cycling, raises an important theoretical issue. Since vote-
trading situations presuppose voting cycles, and voting cycles presuppose externalities,
then “vote-trading situations” must be “negative-externality situations.” Now, is this not in patent contradiction to the idea of Coasean bargaining? Indeed, according to Coasean logic, in the absence of transaction costs, the bargaining process leads to stable Pareto-optimal outcomes. Applied to the legislative context, coasean logic suggests that the market for votes leads to stable Pareto-optimal outcomes. Does this mean that the Logrolling Theorem is wrong or that the Coasean logic is unwarranted? It signifies neither. While the Logrolling Theorem presupposes the absence of binding contracts among voters, Coasean logic presupposes the possibility of concluding binding and enforceable agreements. Once we adopt the same assumptions underlying Coasean logic, the market for votes can be shown to lead to stable and social-welfare-maximizing political outcomes.

This comparison between “social choice” and “Coasean” analytical perspectives is useful as it helps to identify the limits of logrolling as an equilibrium-inducing mechanism. The next subsection thus analyzes the “Political Coase Theorem.”

5.2.5.1.3. The Political Coase Theorem

In a series of papers, Parisi applies the logic of Coasean bargaining to collective decision-making.535 He demonstrates that, under certain assumptions, “political bargaining may provide a solution to the intensity problem and at the same time correct for the cyclical problem.”536 In short, he contends that, when a Coasean exchange of votes is possible through binding political agreements, stability and efficiency are possible in an Arrowian setting. He calls this conclusion the “Political Coase Theorem” (hereinafter, “PCT”).537

537 See also Cooter, Strategic Constitution, op. cit. supra note 219 at 53 (discussing the “political Coase Theorem” and the “Democratic Coase Theorem”). A number of social scientists have proposed some forms of Political Coase Theorem. For example, Becker, pointed out how competition between pressure groups could create a force towards efficient. See Becker “A Theory”, supra note 208; Gary S. Becker, “Public Policies, Pressure Groups, and Dead Weight Costs, (1985) 28 "Journal of Public Economics, 329 [Becker, “Public Policies”].
It is important to analyze the logical relationship between the set of assumptions underlying the PCT and its main analytical results. This enables us to identify the practical relevance of the market for votes as a solution to the problems of political decision-making identified by social choice literature (i.e., intensity, externality, and cycling problems).

**Assumptions.** The Arrowian setting considered by Parisi is a two-dimensional issue space with three voters allowed to exchange votes; voters offer side-payments in the form of a consumption good in exchange for others’ votes on policy issues. The underlying assumptions are the following. (a1) All voter utility functions are continuous and differentiable.\(^538\) (a2) Policy options are continuous so as to allow incremental changes in the policy outcome.\(^539\) (a3) All voters have strictly concave preferences over policy issues. (a4) There are no transaction costs; in particular, there are no free-rider problems nor strategic hold-ups. (a5) There is no *ex ante* uncertainty on the definition of rights (i.e., it is clear which political agreements enjoy legal protection).\(^540\) (a6) Political agreements are enforceable at zero cost. (a7) Marginal utility of wealth is equal across voters (i.e., there is no wealth effect).

**Results.** Parisi demonstrates that, when the above assumptions hold, the majoritarian outcome with vote bargaining is unique, stable and social welfare maximizing. More precisely, under the above assumptions:

(i) Different initial majority coalitions lead to the same final (i.e., stable) policy outcome.

(ii) The choice of alternative decision rules has no effect on the policy outcome.

(iii) Cycling in a multidimensional space is excluded. The majority outcome will be stable, unique and gravitate toward the centre of the policy space.

---

Wittman (*The Myth of Democratic Failure, supra* note 460) pushed this argument further and formulated an informal Political Coase Theorem for democratic societies arguing that, with rational voters, democratic societies generally produce Pareto efficient, even wealth-maximizing, outcomes.

\(^{538}\) Parisi, “Political Coase Theorem”, *supra* note 535 at 3 and 15 n. 23.

\(^{539}\) *Ibid.* at 17, 26, and 29.

\(^{540}\) *Ibid.* at 18.
(iv) The stable and unique bargaining outcome maximizes the aggregate social welfare of the parties.541

The robustness of these results can be appreciated by considering the logical relationship between these propositions and the underlying set of assumptions. First, (a1) (well-behaved preferences) and (a2) (outcome continuity) assume away those indivisibilities that in many real legislative contexts prevent voters from finely expressing the varying intensities of their preferences. Second, (a3) (strict concavity of preferences) is crucial in that it ensures the uniqueness of the bargaining outcome. Third, (a4) (no transaction costs) and (a5) (ex ante assignment of rights) are standard assumptions for the working of the Coase Theorem. They allow the bargaining process to exhaust all the margins of cooperative surplus: in the absence of transaction costs and strategic costs, bargaining will continue until gains from trade – i.e., the exchange of side-payments against votes – are exploited. (a6) (Enforceability of political contracts) is crucial in two respects. On the one hand, it allows one to solve of the cycling problem. Once the unique Pareto-optimal point in the issue space (“global maximum”) is reached, the bargaining outcome crystallizes into an enforceable coalition agreement, which prevents the re-shuffling of coalition majorities thereby ensuring the stability of the policy outcome. On the other hand, the enforceability of coalition agreements ensures that movements in the issue space (resulting from bargaining) converge toward the gravity centre of the policy space.

Let me briefly expand on the stability and efficiency of the majority outcome with bargaining.

Imagine the three ideal points in the policy space. The segments connecting them form a triangle that delimits the Pareto set. The external perimeter of the Pareto set delimits the possible outcomes of majoritarian politics. Parisi has emphasized that without enforceable coalition agreements the majority policy outcome will be located along the triangular perimeter of the Pareto set. By contrast, enforceable coalition agreements

541 More precisely, according to Parisi, if voters’ utility functions have similar curvatures and are centered around equidistant ideal points, then the unique, stable and Pareto-optimal bargaining outcome will satisfy both the Benthamite and the Nash criterion for social-welfare maximization. If voters’ utility functions have different shapes (provided that they reflect strictly concave preferences), and voters’ ideal points are not equidistant, then the bargaining outcome will satisfy only the Benthamite criterion.
allow for movements away from the sides of the triangle toward the *centre* of the Pareto set. To understand why it must be recognized that enforceable political agreements impose limits on the bargaining strategy of non-coalition members. Voters outside of the majority coalition cannot destabilize the majoritarian outcome by offering some members of the coalition the possibility of forming an alternative majority agreement. Enforceability of political agreements excludes the feasibility of this destabilizing strategy and forces non-coalition voters to contract around the policy outcome with the majority coalition “as a unitary group.”\(^{542}\) This blocks *involuntary* redistributions made without the consent of the extant majority so that the only way to improve the outcomes located along the sides of the triangle is to move toward the centre of the Pareto set. Indeed, any movement from the perimeter toward the centre of the Pareto set allows winners to compensate the losers while still remaining better off.\(^{543}\) Of course losers may be compensated through bargaining and side payments. In conclusion, on the one hand, the enforceability of coalition agreements prevents the cyclicality problems; on the other hand, the market for votes between majority and minority allows one to solve the externality problem and to exploit gains from trade, thereby reaching the optimal outcome by moving toward the centre of the triangle.

\(^{542}\) Parisi, “Political Coase Theorem”, supra note 535 at 18

\(^{543}\) *Ibid.* at 29 (“[…] any movement towards the center of the triangle would give rise to utility gains for the non-coalition voters that more than compensate for the utility losses occasioned by the policy change for the coalition members”).
Obstacles to Coasean Bargaining in the Legislative Process. While the Political Coase Theorem illuminates the welfare-enhancing potential of the market mechanism in the legislative process, a closer look at its underlying assumptions enables us to identify the likely obstacles to efficient bargaining for votes in real-world legislative environments. Four major problems plague the market for votes: (i) political indivisibilities; (ii) strategic bargaining and the collective action problem; (iii) incomplete definition of rights and lack of political agreement enforceability; (iv) agency problems.

Indivisibilities. The first fundamental obstacle to political bargaining comes from the nature of the alternatives for voters. As Buchanan has long clarified: “It is in the precise way in which the alternatives mutually conflict that the voting process must be sharply distinguished from the market mechanism.” The market choice, as assumed by

---

544 This picture is taken from Parisi, “Political Coase Theorem”, supra note 535 at 11. It illustrates an example of three voters whose ideal points are equidistant and whose indifference curves have the same circular shape.

545 Buchanan, “Individual Choice”, supra note 450 at 81.
standard consumer theory, entails the allocation of *highly divisible* goods; in this context, the trade-off among alternative options is subject to the principle of diminishing returns. By contrast, voting choices usually entail two or more *discrete* alternatives. In the political context mutual exclusivity derives from the indivisibility of the political outcome (e.g., building a new bridge or not, legalizing abortion or not, engaging in a commercial relationship with country x or with country y, and so on) Thus, instead of confronting the problem of “how much” (as in market choices with divisible goods), the voter often faces the problem of “which one?”

The indivisibility of the political outcome undermines the empirical validity of (a1) and (a2). The presence of indivisibilities in the policy space dramatically alters the structure of the maximization problem legislators face. Economic theory explains that, in the context of discrete outcomes, standard maximization techniques cannot be used to identify the optimum quantity chosen; that is, the optimum is not characterized by standard first-order conditions. By contrast, indivisibility generates discontinuity, thereby impairing the use of marginal optimization techniques. In practical terms, this means that discrete choices might not always allow voters to exploit gains from trade. Discrete choices are of the “all-or-none” type; they do not permit the division of contended resources that is necessary to maximize the surplus from trade.

The indivisibility of the political outcome also takes the form of “issue bundling.” The bundling of different issues is necessary to save the transaction costs associated with a negotiation for every single policy issue. Political agreements tie together many different issues, thereby reducing the otherwise prohibitive negotiation costs. However, for the reason explained above, issue bundling significantly limits the surplus-optimizing potential of vote-trading. Thus, politicians engaging in vote bargaining face a trade-off between minimization of transaction-costs and maximization of the surplus from trade.

---

546 Indeed, many real-world market choices are discrete.
To conclude on this point, the indivisibility of the political outcome accentuates the distributive (conflictual) aspect of politics; consequently, it generates the above-mentioned problems of involuntary redistribution and outcome instability.\textsuperscript{547}

\textit{Strategic Bargaining and The Collective Action Problem}. Another crucial assumption of the PCT is the absence of transaction costs (a4). However, it must be recognized that the legislative process is generally plagued by severe and pervasive collective action problems that are induced by the public good aspect of many political decisions.\textsuperscript{548} Furthermore, the high degree of interdependence among political actors generates strategic behaviour by politicians, which dramatically increases transaction costs. Finally, transaction costs are likely to increase with the number of issues and voters.\textsuperscript{549} All these characteristics of the political environment (publicness, interdependence, multidimensionality) undermine the empirical validity of (a4).

\textit{Incomplete Definition of Rights and Lack of Agreement Enforceability}. The most relevant impediments to the working of the PCT are (i) the lack of certainty (within the political environment) in the \textit{ex ante} assignment and definition of property rights, and (ii) the lack of legal enforceability of political agreements. The above discussion has assumed away these difficulties ((a5) and (a6)); however, these assumptions significantly limit the practical relevance of the PCT. First, the exchange and selling of votes among legislators is either prohibited or not legally recognized in many jurisdictions. Second, there is a general distrust or moral disapproval of concluding contracts concerning decisions in the legislative arena, which prevents the clear assignment and definition of rights.\textsuperscript{550} Third, in general, parliaments have the right to change the assignment of both decisions and

\textsuperscript{547} This latter characteristic is emphasized when the political process decides on the allocation of legal rights. These can be conceptualized as positional goods, defined as good “such that, given the consumption choice of one agent, the second agent must consume a corresponding negative amount of what the first chooses to consume”; see, Ugo Pagano “Legal positions and Institutional Complementarities” in Fabrizio cafaggi, Antonio Nicita, and Ugo Pagano. “Legal orderings and economic institutions” (Abingdon, Oxon: Routledge, 2007) at 63. The competition for positional goods can be harder the competition for “normal” goods. Pagano observes (“Positional competition is much harder, and sometimes more violent, than competition for ‘private’ goods” Ibid. at 64).

\textsuperscript{548} Cooter, \textit{Strategic Constitution}, supra note 219 at 52.

\textsuperscript{549} See McKelvey, “Intransitivities” \textit{supra} note 516.

\textsuperscript{550} Cooter, \textit{Strategic Constitution}, supra note 219 at 63
property rights at their discretion. This creates a problematic – and still poorly understood – institutional overlap between the actors and the regulators of the market for votes (i.e., politicians are concurrently actors in the market for votes and regulators of the same market).

At present, both academic scholarship and historical experience do not provide guidance as to how to identify characteristics of the institutional framework that would better enable legislators to engage in efficient logrolling and ensure the enforceability of political agreements. Thus, the efficiency of the market for votes remains a logically interesting hypothesis, but much work remains to be done to clarify the costs and benefits of its practical application and to solve the many problems affecting the market for votes.

Conclusions on Vote-Trading. The foregoing discussion has identified three orders of limitations that affect vote-trading as a mechanism for inducing efficiency and equilibrium in the legislative process. First, vote-trading is limited by structural and utility requirements, which make it inapplicable to a large number of cases where the problems of intensity and of externality occur. Second, even when it is applicable, social-welfare maximizing conditions must also be met. If these conditions are not satisfied, logrolling might reverse the externality problem without solving it (i.e., the intense minority gains less than the majority loses). Third, the functioning of the market for votes is plagued by exogenous indivisibilities, high transaction costs and agency problems. This undermines the generalizability of the PCT.

5.2.5.2. Structurally Induced Equilibrium

According to many social choice scholars, the ability of non-Condorcet rules to induce outcome stability independent of the existence of a CW explains the observed stability of real legislative processes. In this section, I briefly discuss the most important stabilizing-devices characterizing legislative decision-making processes.

---

551 This sub-section draws on Gerald S. Strom, The Logic of Lawmaking (Baltimore and London: The Johns Hopkins University Press, 1990) at 83 [emphasis mine] [Strom, “The Logic of Lawmaking”]
552 It was seen in previous sections that the Condorcet criterion requires (at least) the same number of pairwise votes as the available choice alternatives in order to distinguish the CW from other potentially defeasible outcomes. When this condition is satisfied, pairwise majority voting produces either (1) a CW
5.2.5.2.1. The Problem of Outcome Legitimacy

Before proceeding, there is an important theoretical point to be clarified. The fact that the observed stability of legislatures depends on institutional stabilizing devices rather than on the preferences of stable majorities poses the problem of the potential arbitrariness of the legislative outcome. The outcomes generated by non-Condorcet rules are inherently path-dependent – i.e., they vary depending on the order of pairwise voting between options. Path-dependence threatens the legitimacy of the law-making outcome for two important reasons. First, it might potentially mask the existence of intransitive cycles or thwarted majority preferences, thereby undermining the majoritarian legitimacy of the actual outcome. Second, path-dependence gives the “agenda setter” the power of strategically influencing the outcome, which obviously undermines the legitimacy of the resulting collective decisions. Both issues merit brief clarification.

In the absence of a CW, the structurally induced equilibria pose legitimacy problems because the outcomes do not result from the defeat of all available alternatives in direct binary comparisons. Consequently, while these outcomes appear to have majority support, this does not mean that an alternative majority could not have selected an alternative option in a suppressed, direct, binary comparison. In short, these outcomes appear legitimate but they might conceal the preferences of an alternative majority. Second, the existence of preference intransitivity can be excluded only after an unlimited pairwise voting has taken place; therefore, the outcomes that do not defeat all available alternatives might conceal the existence of an intransitive cycle. Third, if the outcome depends on the voting order, the one who controls the voting agenda also controls the outcome. Thus, path dependency gives tremendous strategic power to the agenda setter, if one is available, or (2) an intransitive cycle if a CW is not available. This means that the Condorcet rules generates outcome stability only if actual social preferences are transitive. By contrast, when the decision-making process does not meet the Condorcet criterion, then the stability of the political outcome is independent of the existence of a Condorcet Winner: stability is institutionally induced rather than preference-based.

since potentially any result desired by the agenda controller can ultimately be achieved as the final outcome.\textsuperscript{554}

To summarize, non-Condorcet rules suppress a certain number of pairwise comparisons, which prevents the decision-making process from revealing either the preferences of alternative majorities or the existence of intransitive preferences. This stabilizes the legislative process, but at the same time threatens its legitimacy and creates room for serious agency problems associated with the power of the agenda setter to influence the policy outcome. The preceding considerations are not meant to support the conclusion that non-Condorcet rules are inefficient \textit{per se}. Rather, they suggest assessing efficiency in light of the \textit{legitimacy/stability} trade-off. Legitimacy (when grounded in the majoritarian principle) requires legislative outcomes that reflect the preferences of the majority of voters expressed through an unlimited series of binary comparisons. However, issue multidimensionality and the multi-peakedness of voter preferences decrease the likelihood of an available CW. In these cases, non-Condorcet rules ensure that a stable outcome is reached, thereby reducing the costs of outcome instability. Of course, stability comes at a price: the stable outcome is not a majoritarian equilibrium but a “structurally-induced” equilibrium: it reflects the structure of the decision-making process \textit{without} reflecting the preferences of a stable majority of voters. I do not discuss here the more general question of whether a structurally induced equilibrium is a \textit{normatively} defensible outcome. This would require a more sophisticated discussion of the notion of “legitimacy,” which falls outside of the scope of the present study.

In the next subsection, I briefly discuss three stability-inducing mechanism: (i) control of the agenda; (ii) the backward moving agenda process; (iii) division of the question rule, and, (iv) strategic voting. The goal of the analysis is to inquire into the efficiency properties of structurally induced \textit{equilibria}.

\textsuperscript{554} It is sufficient for the agenda setter to identify the problematic amendments that would defeat the favoured one in a pairwise vote, and to organize the voting order so that the problematic amendments are eliminated in early pairwise comparisons. See McKelvey, “Intransitivities” \textit{supra} note 516.
5.2.5.2.2. Control of the Agenda

Control of the agenda can induce stability in legislative decision-making by restricting the set of available alternatives for choice. Consider three legislators – “Conservative,” “Moderate” and “Liberal” – deciding among three options A, B and C. Assume that C is the conservative option: as such, it is the most preferred by Conservative. Let me start with the following assumptions: (a1) individual preferences are single-peaked; (a2) the voting-space is single-dimensional and it can be ordered along a left-right spectrum; (a3) Conservative is the monopoly agenda setter who decides the order of pairwise voting. Under these assumptions, let me consider four scenarios, which are different depending on (i) whether or not the voting process meets the Condorcet criterion (ii) whether or not a Condorcet Winner is available. The discussion of these cases can be usefully applied to parliamentary committees as well as to full parliamentary chambers.

Case 1. Condorcet rule with an available CW

If the voting process satisfies the Condorcet criterion (i.e., at least three pairwise votes, in this example), the voting process will select the Condorcet Winner (if one is available) no matter the voting order set by the agenda controller. For example, suppose that B is the moderate option favoured by Moderate, and A is the progressive option favoured by Liberal. The preference ordering looks as follows:

Conservative: \( C > B > A \)
Moderate: \( B > A > C \) (or \( B > C > A \))
Liberal: \( A > B > C \)

This is a typical case of transitive social preferences. The Condorcet Winner is option B, which is the second majority-choice for both Liberal and Conservative. Given that the voting process is regulated by a Condorcet rule, B will be the stable majority rule equilibrium, no matter the voting order established by the agenda setter.
Case 2. Condorcet rule with no available Condorcet Winner

If a Condorcet Winner is not available, unlimited pairwise voting leads to an intransitive cycle. Assume, for example, the following preference orderings:

Conservative:  \[ C > B > A \]
Moderate: \[ B > A > C \]
Liberal: \[ A > C > B \]

Each selected outcome is majority preferred by an available alternative. There is no majority equilibrium, because the Condorcet rule unveils the intransitivity of social preferences.

Case 3. Non-Condorcet rule with no available CW

Assume an intransitive social preference ordering like the one described in case 2. Assume also that, to ensure outcome stability, a rule is introduced that allows only two pairwise votes. Conservative (the agenda setter who prefers Q), has to decide between three alternative orders: (1) C versus B and the winner versus A, or (2) C versus A and the winner versus B, and, finally, (3), B versus A and the winner versus C. It is easy to see that under order (1) A is the selected outcome, because Q wins against B, and then Q loses against A; under order (2) B is the selected outcome because Q loses against A and A loses against B. By contrast, under order (3) C is the selected outcome, because B wins against A and then B loses against C. Thus, the agenda setter (who prefers C) will choose order (3).

In general, the monopoly agenda setter can influence the outcome (i) by suppressing some binary comparison, and (ii) by organizing the sequence of pairwise voting so that the problematic outcomes (i.e., those that defeat the agenda setter’s favoured outcome) are defeated by the alternatives that will lose against the favoured option.
Case 4. Non-Condorcet rule with available CW

Assume, now, that the two-pairwise voting procedure is adopted despite the fact that a Condorcet Winner is available. Assume, for example, a social preference ordering as in case 1. As already noticed, the Condorcet Winner is A, which defeats all the alternative options in pairwise voting. In this situation there is no way for Conservative to manipulate the agenda in order to obtain C. No matter what the voting order is, A will be the majority outcome.

The table in Figure 19 summarizes the foregoing discussion on these four cases.

Figure 19. One-Dimensional Issue, Single-Peaked Preferences, Three Voters, Three-Alternatives, Pairwise Voting, and Agenda Setter’s Control

<table>
<thead>
<tr>
<th>Available Condorcet Winner</th>
<th>Condorcet Winner (No Agenda Manipulation)</th>
<th>Not Available Condorcet Winner</th>
<th>Intransitive Cycle</th>
<th>Agenda Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condorcet Rule</strong></td>
<td>Condorcet Winner (No Agenda Manipulation)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No-Condorcet Rule</strong></td>
<td>Condorcet Winner (No Agenda Manipulation)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If a Condorcet Winner is available, it will be selected as a majority outcome, whether or not the voting procedure meets the Condorcet criterion. If a Condorcet Winner is not available, a Condorcet rule gives rise to an intransitive cycle. By contrast, the limitation in the number of pairwise comparisons generates stability, but it also creates room for agenda manipulation. It is important to emphasize that the power of the agenda setter comes from the fact that (i) the number of pairwise voting is fewer than that required by the Condorcet criterion, and that (ii) the agenda setter decides the voting order (and, consequently, which binary comparisons are suppressed).
In the four preceding cases, the whole set of available options (A, B and C) enter the decision-making process. Even if some binary comparisons are suppressed, all the three available options are subject to at least one binary comparison. In short, the agenda setter reduces the number of binary comparisons, but not the available options. This significantly limits the power of the agenda setter to influence the selected outcome. However, in real legislative contexts the range of possible outcomes is much larger than in the three-option example illustrated above. Thus, the need arises for limiting the set of available alternatives. In this more realistic context, the agenda setter can influence the outcome by strategically selecting the range of available alternatives. The (i) possibility of limiting the set of available options, combined with (ii) the right to choose the voting order, and (iii) limited pairwise comparison gives the agenda setter tremendous power to strategically influence the outcome. In general, when the agenda setter can limit the set of available alternatives, it is highly unlikely that an available Condorcet Winner will be the outcome of the decision-making process. Let me illustrate this point by examining three scenarios in which the agenda setter has the power to limit the set of alternatives that enter the decision-making process.

**Case 5. Available CW, non-Condorcet rule, agenda manipulation of available alternatives**

Let me assume a forward moving agenda process in which each proposed alternative must be voted against the status quo, Q. The alternative that defeats Q in a direct comparison itself becomes Q and is put against the subsequently proposed alternative. The agenda setter decides which alternative is proposed against the status quo, and has the right to terminate the decision-making process by calling the last question. In this way, the agenda setter can decide which available alternatives enter the decision-making process.

Under these assumptions, consider again the three-legislators and three-alternatives example. Assume that Conservative (the agenda setter) prefers C, which is at the right end of the single-dimensional spectrum. Assume, that B is the Condorcet Winner: it is the first-choice of Moderate and the second choice for both Liberal and Conservative. Option B is located to the left of A and to the left of C at an equidistant point between A and C,
so that both Moderate and Liberal prefer B to C and Conservative prefers C to B. Finally, the *status quo* Q (e.g., the existing legislation) is located to the left of A but to the right of B. In particular, Q is closer to B than to A. Figure 20 illustrates this distribution of preferences.

Figure 20 Available CW, Non-Condorcet Rule, Agenda Manipulation

![Diagram](image)

Conservative is the agenda setter who prefers C as a first-choice. However, if he or she proposes C, it will lose on the first vote against Q, because both Progressive and Moderate would prefer Q to C. By contrast, if Conservative proposes B, it will win on the first vote against Q because both Conservative and Moderate prefer B to Q.

This example shows that in a single-dimensional issue space with single-peaked preferences – i.e., the conditions for the emergence of a Condorcet Winner – when the agenda setter has the power to restrict the set of available alternatives, he or she will manipulate the agenda so as to obtain the outcome that is the closest to his or her preferred outcome. If his or her most preferred outcome is far from the Condorcet Winner – in the sense that there are intermediate outcomes preferred by the agenda setter to the Condorcet Winner – then the legislative decision-making process will not select the Condorcet Winner. *The legislative decision-making process will select the Condorcet Winner only when the agenda setter’s preferred outcome is the Condorcet Winner.* This latter exceptional case is shown in Figure 21.
Put bluntly, the location of the selected outcome in the issue space depends on the distance between the agenda setter’s ideal point and the Condorcet Winner. Finally, notice that this structure-induced equilibrium is not totally insensitive to other voter preferences; however, the preferences of the agenda setter are those that most decisively affect the legislative outcome.

Case 6. Not-Available Condorcet Winner, Non-Condorcet Rule and Agenda Manipulation of Available Alternatives

Consider a sixth case in which the single-peakedness assumption is relaxed. As we already know, if preferences are not single-peaked social intransitivity is likely to occur. In this respect, case 6 and case 3 are similar: a non-Condorcet rule with no available Condorcet Winner creates room for agenda manipulation. However, unlike case 3, in case 6 the agenda setter has the power to choose the alternatives in the decision-making process. This exacerbates the tendency of the decision-making process to depart from the Condorcet Winner. As noticed above, to the extent that the agenda setter’s preferred outcome does not coincide with the Condorcet Winner the selected outcome will be “Condorcet inefficient.”

Case 7. Not-Available Condorcet Winner, Non-Condorcet Rule, Agenda Manipulation of Available Alternatives, in a Multidimensional Issue space

Let me now relax the single-dimensional issue space assumption. McKalvey has demonstrated that in a multidimensional issue space, when a Plott equilibrium does not exist, and when preferences are intransitive, the monopoly agenda setter can lead the pairwise voting toward any desired outcome. That is, the agenda setter can strategically

555 These determine the existence of the Condorcet Winner and its position on the issue space.
manipulate the agenda of pairwise votes and achieve “any outcome in the whole issue space.”

It follows from these preceding considerations that the relaxation of the single-dimensional assumption entails a dramatic increase in the manipulative power of the agenda setter. In a single-dimensional space, with intransitive social preferences, the agenda setter can strategically manipulate the outcome but cannot often achieve his or her ideal point. By contrast, in the case of intransitive multidimensional preferences, the agenda setter can always achieve the outcome he or she chooses.

Finally, let me emphasize that when the agenda setter has the power to restrict the set of available alternatives entering the decision-making process, the majority of legislators are likely to be made worse off (than they would have been without this power of the agenda setter). Indeed, when a CW is available, any alternative outcome generates less utility. But, when a CW is not available, then it is likely that some alternative non-Condorcet rules that do not empower the agenda setter to select the alternatives for consideration might yield better outcomes.

Summary. A short summary of this subsection might be helpful. In a three-person, three-alternatives decision-making process based on pairwise majority votes, the power of the agenda setter depends on (i) the number of pairwise votes, and (ii) the availability of a CW. If a CW is available, there is no room for agenda manipulation, whether or not the number of pairwise votes meets the requirements of the Condorcet criterion. If no CW is available, a non-Condorcet rule gives power to the agenda setter to strategically manipulate the voting order and influence the outcome.

This manipulative power increases when the agenda setter is given the right to restrict the set of alternative outcomes. An example is the rule granting the agenda setter the right to

556 Strom, The Logic of Lawamking, supra note 551 at at 83 [emphasis mine].
557 See Strom, The Logic of Lawamking, supra note 551 at 83.
558 For example, a good strategy for countering the manipulative power of the agenda setter is to adopt succession voting. See Stearns and Zywicki, Public Choice, supra note 500 at 136.
select the option to be put against the *status quo* in a pairwise vote, along with the right to end the decision-making process by calling the last question. In this case, the outcome does not coincide with the Condorcet Winner, unless the agenda setter shares the voter’s median preference peak.

Further, if individual preferences are *multi-peaked*, the manipulative power of the agenda setter increases. Finally, with socially intransitive *multidimensional* preferences the agenda setter can *always* achieve *any* outcome in the issue space he chooses.

5.2.5.2.3. The Backward Moving Agenda Process

Another strategy for ensuring *stability* of the legislative law-making process is the adoption of a “backward moving agenda process.” This procedural rule requires that any alternative resulting from a series of pairwise majority voting must win a final direct comparison against the extant *status quo*, Q. To understand the stabilizing effect of this procedural rule it is useful to introduce the concept of “win set” X. The win set of a given alternative X, denoted by W(X), is the set of alternatives that a majority prefers to X in a direct pairwise comparison. Now, once the backward moving agenda process is adopted, the set of possible outcomes is restricted to the *status quo* and the win set of the *status quo*. Since any selected option must defeat the *status quo* in a majority vote, the issue space is restricted to Q and W(Q). That is, the only changes to the *status quo* that the decision-making process can produce are those preferred by a majority to the *status quo*.

The preceding consideration clarifies that the backward moving agenda process acts as a limit to the issue space. Three clarifications are needed on this point. First, *the constraining effect of the backward moving agenda process depends on the distance in the issue space between the centre of legislator ideal points and the status quo*. Therefore, it would be incorrect to say that the structure of the backward moving agenda process favours the *status quo*. The *status quo* acts as limit to the range of possible outcomes, but there is nothing in this rule that prevents the majority from preferring a radical change to the *status quo*. Whether the change in the *status quo* is radical or incremental in nature depends on the distance between the geometric centre of the ideal points of legislators in the issue space and the *status quo*. When the *status quo* is near the
centre of the ideal points, then the alternatives included in W(Q) are close to the status quo in the issue space. On the contrary, when the status quo is distant from the centre of the ideal points of legislators, then W(Q) will even include alternatives that are distant from the status quo in the issue space. In this latter case, the backward moving agenda does not provide a significant limitation on the range of possible legislative outcomes. Thus, the structure-induced equilibrium is not preference-insensitive; rather, it is significantly influenced by the structure of individual preferences. Inducing equilibrium does not mean rendering the lawmaking process insensitive to individual preferences. Second, by limiting the set of possible outcomes, the backward moving agenda process acts as a limit to the power of the agenda setter. In fact, the agenda setter is bounded by the set of possible outcome including Q and W(Q). Finally, the adoption of a backward moving agenda process is not alone sufficient to establish a unique equilibrium outcome. It reduces the set of available outcomes but is not a sufficient condition to ensure a stable equilibrium.  

5.2.5.2.4. Division of the Question Rule

We know that a major factor of instability in the legislative decision-making process is the multidimensionality of the issue space. In this respect, an institutional constraint that allows one to deal with the problem of issue multidimensionality is the “division of the question rule.” According to this rule the vote on a multidimensional issue must take place one dimension at a time.

Under the assumption that individuals have “separable” preferences, the division of the question rule is sufficient to induce a stable equilibrium in the legislative decision-making process. Voting one dimension at a time allows the decision-making process to identify the median preference peak in each dimension and to eliminate the risk that the median in one dimension is defeated by a majority formed by consideration of other issue dimensions. By imposing the consideration of one dimension at a time, division of the

559 See Strom, The Logic of Lawmaking, supra note 551 at 98.
561 See Strom, The Logic of Lawmaking, supra note 551 at 114.
question rule forces the equilibrium to exist “at the point at which lines (or planes) drawn through the median preference on each dimension intersect.”

As Mueller clarifies, the division of the question rule “adds further constraints to the maximization problem.” It is unlikely that a majority equilibrium point reached through the division of the question rule will satisfy efficiency conditions.

Finally, it is worth underlining that if the assumption of separable preferences is dropped, then the division of the question rule is not sufficient to guarantee a stable equilibrium.

5.2.5.2.4. Strategic Voting

Social choice theory has demonstrated that even assuming the absence of any stabilizing institutional mechanism a stabilizing effect on legislative decision-making can derive from the strategic voting behaviour of individuals in the legislative arena. Namely, limits to the movement in the issue space can result from the sophisticated behaviour of voters seeking to maximize their utility function. Let me succinctly clarify this point, which is important for identifying the conditions under which political law-making is more likely to lead to relatively stable outcomes.

It is necessary to introduce the notion of a “covering relation.” Consider a choice alternative Y that is an element of W(X). Y is said to “cover” X when all the alternatives that are majority preferred to Y are also majority preferred to X. Differently stated, if Y covers X, all the alternatives in W(Y) are also in W(X). The covering relation between alternatives in the issue space has two important implications. First, if Y covers X, it follows that there is no intransitive cycle among X, Y and the alternative Z that belongs to W(Y). If any Z that is in W(Y) is also in W(X), it means that every alternative that defeats Y also defeats X. I call this implication the “intransitivity effect.” Second, if Y covers X, X can be excluded from the set of alternatives for considerations. In fact, since every alternative that defeats Y also defeats X, rational voters would choose either Y or

563 Strom, The Logic of Lawamking, supra note 551 at 101.
564 Mueller, Public Choice III, supra note 494 at 117.
565 By contrast, intransitivity would require Y to be an element of W(X), Z to be an element of W(Y), and X to be an element of W(Z).
any alternative that is in W(Y) rather then choosing X. I call this implication the “eliminating effect.”

The intransitivity effect and the eliminating effect explain why the strategic behaviour of rational voters can bring about a restriction of the issue space. To understand this point it is useful to introduce the concept of the “uncovered set.” Assume: (a1) strategic voting, (a2) a forward moving agenda process, and (a3) no monopoly agenda setter. Consider the alternatives X, Y and Z in the multidimensional issue space. To each alternative corresponds both a set of covered and uncovered outcomes. I denote the set of covered outcomes respectively as C(X), C(Y) and C(Z). These constitute the sets including all the alternatives covered by X, Y and Z, respectively. Then, I call the sets of the uncovered outcomes the “uncovered sets,” which I denote respectively as UC(X), UC(Y) and UC(Z). These represent the alternatives that are not covered by X, Y and Z, respectively. The issue space can be conceptualized as the sum of the covered and uncovered sets of X, Y and Z:

\[ S = [C(X) + C(Y) + C(Z)] + [UC(X) + UC(Y) + UC(Z)]. \]

The left addend on the right member of the equation is the covered set of the whole issue space, which I denote C(S); the right addend on the right member of the equation is the uncovered set of the whole issue space, which I denote UC(S). Thus, the whole issue space is the sum of the uncovered set and the covered set of S:

\[ S = C(S) + UC(S). \]

By virtue of the eliminating effect, C(S) is excluded from the set of alternatives that a majority would choose. This is because each alternative in C(S) is preferred by at least one alternative in UC(S). To be clear, an alternative Z that is in C(S) can be an

---

566 An alternative Z that is in C(S) can be an outcome only if a monopoly agenda setter does not include in the agenda the alternative in UC(S) that covers Z. But, since we assume no monopoly agenda setter, then if Z is in the agenda there will always be a majority that proposes an alternative in UC(S) that defeats Z.
outcome only if a monopoly agenda setter does not include in the agenda the alternative in UC(S) that covers Z. But, since we assume no monopoly agenda setter, then if Z is in the agenda there will always be a majority that proposes an alternative in UC(S) that defeats Z. Thus, with no monopoly agenda setter, any alternative Z that is in C(S) cannot be an outcome. This is a very important conclusion. It means that any possible outcome of legislative decision-making must be in the uncovered set. The stabilizing effect is not induced by any stabilizing procedural rule (e.g., division of the question rule), but it is determined by the inherent attractiveness of the alternatives in the uncovered set.\textsuperscript{567} Therefore, even if a Condorcet Winner does not exist, if legislators strategically eliminate the dominated outcomes, cycling is restricted to the alternatives in the uncovered set.

Now that I have clarified how under certain assumptions the range of possible outcomes is restricted to the uncovered set, let me briefly discuss the efficiency features of these outcomes. This point is the subject of a highly sophisticated literature, which I cannot review in detail here. However, three general (but important) conclusions about the location of the uncovered set in the issue space must be emphasized. First, the location of the uncovered set depends on the distribution of the ideal points in the issue space. Thus, once again, the equilibrium-inducing mechanism does not suppress the decisive role played by individual preferences in determining the legislative outcome. Second, social choice scholars have shown that, under the assumption of the Euclidean preferences of voters,\textsuperscript{568} the uncovered set is located near the geometric centre of the ideal points.\textsuperscript{569} This has an important implication for the quality of the outcome: despite the multidimensionality of the issue space, the selected alternatives are “not too far” from the joint median alternative. Put bluntly, legislative decision-making will continue to adopt centrist policies even in a multidimensional space. Third, Miller has demonstrated that,

\textsuperscript{567} Mueller, Public Choice III, \textit{supra} note 494 at 237 well clarifies this point in respect to a similar three legislators example: “[…] if a candidate chooses a platform from the uncovered set she knows that she is at most ‘once removed’ from defeating any platform her opponent chooses. At worst, her platform will be involved in a cycle of length three with any platform that defeats it. Conversely, if she chooses a platform that is covered, not only can this platform be defeated, but the platforms that defeat it include some that her platform cannot defeat. Thus, her platform can be contained in a transitive triple in which it is the least preferred of the three platforms.”

\textsuperscript{568} Explain what are Euclidean preferences.

\textsuperscript{569} Mueller, Public Choice III, \textit{supra} note 494 at 239.
under certain conditions,\textsuperscript{570} the uncovered set is a subset of the Pareto set. The Pareto set is the set of alternatives that cannot be changed without making at least one legislator worse off.\textsuperscript{571} Based on this definition of the Pareto set, one can appreciate the importance of Miller’s result, which demonstrates that under certain conditions the alternatives in the uncovered set are Pareto-optimal.

The strength of the preceding result depends on the plausibility of the assumptions. First, the absence of a monopolist agenda setter seems plausible. We have seen that the presence of an agenda setter makes the majority of legislators worse off. Thus, it is plausible to assume that legislators do not subject themselves to a monopoly agenda power. Second, the assumption of a forward moving agenda process does not reflect the reality of many legislative decision-making processes, which more often adopt a backward moving agenda process. This, however, does not undermine the essence of the results. Indeed, assuming a backward moving agenda process simply adds a further restrictive condition to the range of possible outcomes. Namely, the outcome of a backward moving agenda process with strategic voting should be located in both the uncovered set and the status quo win set.\textsuperscript{572} Finally, the more problematic assumption is that of strategic voting, which requires legislators to be fully informed about other legislators’ preferences and about the relative salience of multiple issues for other legislators. This is undoubtedly a strong requirement. However, it cannot be ignored that in a significant number of cases the members of the legislative arena are aware of the preferences of other members.

\textbf{5.2.6. Conclusions on Social Choice Efficiency}

The discussion in this section implicitly assumed the absence of the information problems discussed in section 7.1. and of the agency problems to be addressed in section 7.3. Despite these simplifying assumptions, it has been demonstrated that only under very


\textsuperscript{571} Here I am implicitly assuming that legislators act as faithful agents of voters – i.e., I am assuming zero agency costs. However, as I will discuss in detail in the next section of this chapter, in real world situations legislators have utility functions independent of voters.

\textsuperscript{572} Strom, The Logic of Lawamking, supra note 551 at 123.
restrictive conditions is the legislative process capable of reaching a majority outcome that meets both legitimacy and efficiency requirements. This difficulty is embedded in majoritarian logic that, in most circumstances, is incompatible with the logic of economic efficiency. A brief summary of the main points of the discussion may be useful.

(i) **Majority and Outcome Efficiency.**
Majoritarian logic is difficult to reconcile with the logic of outcome efficiency. The majoritarian method of decision-making (a) is insensitive to preference intensity, (b) is a source of political externality, and, (c) in the absence of an available CW, generates unstable and incoherent results. As previously noted, these problems can be reduced through vote-trading and structure-inducing mechanisms; however, the conditions for outcome efficiency remain quite restrictive.

(ii) **Median Equilibrium.**
If the voting issue is single dimensional and voter preferences are single-peaked, then majority voting – both one-stage voting and unlimited pairwise comparisons – produces a majority equilibrium outcome. Under these assumptions, the majority equilibrium is always Pareto-optimal. It is not K-H efficient unless there is symmetry between voter preference intensities. It does not ensure the maximization of social welfare unless the median voter’s cost and utility functions reflect the mean values of the costs and benefits for the people involved in the decision-making process.

(iii) **Intransitivity and Chaos.**
If the single-peakedness assumption is relaxed, intransitive social preferences are likely to occur. In this respect, Arrow’s Theorem demonstrates that adopting the conditions of non-dictatorship and non-imposition of choices, social welfare functions that satisfy the axioms of the irrelevance of independent alternatives and of Pareto optimality cannot be derived. If the single-dimensional assumption is relaxed, according to the Chaos Theorem the legislative process can wander anywhere in the issue space, which results in chaotic decision-making and unpredictable outcomes. Both Arrow’s Theorem and the Chaos Theorem threaten the legitimacy (as defined according to the majoritarian principle) of the
democratic decision-making process. They demonstrate that much of the *equilibria* reached in real legislative processes is structurally-induced, rather than corresponding to the preferences of a stable majority of voters.

(iv) *Structure-Induced Equilibrium.*  
Structure-induced equilibrium is achieved through various stabilizing institutional devices. I have discussed the following: (a) control of the agenda; (b) the backward moving agenda process; (d) division of the question rule, and (d) strategic voting. These institutional mechanisms variously limit the set of available alternatives in the issue space. Significantly, the structure-induced equilibrium is path-dependant and easily manipulated by the agenda controller. Despite the fact that real legislatures are usually *not* subject to a monopoly of agenda control, the path-dependant nature of the structure-induced equilibrium creates room for serious agency problems (in addition to those identified in the next section of this chapter).

Three characteristics of the structure-induced equilibrium should be emphasized. First, it is not “preference-insensitive.” Voter preferences are still decisive in influencing the outcome. However, preferences influence the outcome in a way that is *not* consistent with majority rule. Second, the structure-induced equilibrium is in most cases unlikely to be consistent with outcome-oriented standards of economic efficiency. The efficiency of the structure-induced equilibrium mostly depends on the distance between the preferences of the median voter and the agenda setter’s ideal points. Third, in a multidimensional issue space the manipulative power of the agenda setter increases significantly. McKalvey has demonstrated that in a multidimensional issue space the agenda setter (in the absence of rules limiting his or her power) can influence the decision-making process in the direction of any outcome he or she prefers.

(v) *Market For Votes.*  
Part of social choice scholarship emphasizes the potential of vote-trading as a means of restoring cardinal utilities in majoritarian decision-making. In this way, the vote-trading process would solve, or at least significantly reduce the intensity, externality and majority
problems. However, this discussion has demonstrated that the market for votes is a severely limited mechanism for curing the problems of majoritarian decision-making. I have identified three major problems. First, vote-trading is subject to both *structural* and *utility* requirements, which significantly limits its practical utilization within the legislative arena. Second, if the conditions for social-welfare maximization are absent, vote-trading may enable intense minorities to reverse the majority outcome and to reduce the aggregate social welfare. Third, the assumptions underlying the Political Coase Theorem are in most cases difficult to reconcile with the structural features of legislative law-making. In fact, real-world political-environments are severely plagued by (a) exogenous indivisibilities, (b) strategic bargaining costs, (c) lack of both property rights and agreement enforceability, and (d) agency problems (that I will discuss in the next section).

(vi) *Process Efficiency.*

From the perspective of *process* efficiency analysis, the preceding discussion suggests that political law-making might enjoy comparative advantages (versus private ordering and courts) in those regulatory environments where people subject to law are likely to express *coherent* preferences. This is likely to occur when there is a sufficient degree of homogeneity of opinions across people, which allows for *single-peaked* preferences along one commonly recognized *single-dimensional* issue. Under these conditions political law-making permits significant economizing effects in terms of reduced bargaining and strategic costs, while at the same time generating stable and relatively efficient results. When single-peakedness and single-dimensionality are lacking the social choice problem is exacerbated. In these cases, legislatures can produce stable outcomes only at the price of serious departures from majority rule.

The preceding conclusion needs to be qualified with two methodological observations. First, social choice theory is mostly based on the idea of the political process as a *linear* production process transforming inputs into outputs. Individual preferences are the *inputs* that the political process transforms into political *outcomes*. In other words, social choice theory assumes that the inputs are *exogenous* to the political process. However, it must be
recognized that this very assumption exacerbates the social-choice problem. In real-world politics, the preferences of the individual and the political process interact in a complex circular dynamic. Namely, the mechanisms of the political process filter and shape the preferences of individuals by creating the conditions for strong and coherent ones.\footnote{See Farber and Frickey, “Law and Public Choice”, \textit{supra} note 511 at 57-62.}

Second, in this section I have not discussed the problem of legitimacy in depth. I have assumed – over-simplistically – that the one-person-one-vote principle is the main source of legitimacy along with the requirements of outcome rationality and stability. Based on this assumption, as repeatedly noted, social choice theory identifies a trade-off between legitimacy and efficiency. However, it is crucial to recognize that the political process cannot be reduced to majority rule, and political legitimacy cannot be reduced to the idea that every single outcome must be majority preferred.\footnote{See Buchanan, “Social Choice” \textit{supra} note 573 at 119 (“[…] certainly, majority rule is acceptable in a free society precisely because it allows a sort of jockeying back and forth among alternatives, upon none of which relative unanimity can be obtained. Majority rule encourages such shifting, and it provides the opportunity for any social decision to be altered or reversed at any time by a new and temporary majority grouping. In this way, majority decision-making itself becomes a means through which the whole group ultimately attains consensus, that is, makes a genuine social choice. It serves to insure that competing alternatives may be experimentally and provisionally adopted, tested, and replaced by new compromise alternatives approved by a majority group of ever changing composition. This is democratic choice process, whatever may be the consequences for welfare economics and social welfare functions.”).}

First, many institutional features of the legislative process have an independent normative appeal. In some cases, these features can make legislative decision-making “legitimate,” even if in some instances the single legislative outcome does not correspond to a majority-rule equilibrium. Second, and relatedly, from a constitutional perspective, the inconsistency of majoritarian voting can itself be considered a normatively appealing feature of the political process. Buchanan has clarified this point. Majoritarian decision-making ensures collective action in those situations where the search for unanimity would lead to a deadlock. In these situations, the majority decision is preferred to a stalemate.\footnote{See Farber and Frickey, \textit{Law and Public Choice}, \textit{supra} note 511 at 57-62.} From this perspective, inconsistency of majority decisions creates incentives to search for new and innovative
majority decisions. In other words, the “inconsistency” of majority decisions ensures their “provisional” nature, thereby allowing for further attempts to reach unanimous consensus.576

The problem of political outcome legitimacy requires a much more elaborate discussion than that permitted by the scope of this study.577 Here, it is important to bear in mind that to be sound, process-efficiency analysis must include a careful consideration of the legitimacy of law-making.

5.3. Agency Efficiency

Thus far the present discussion has assumed an undifferentiated electorate: people subject to law have equal “political weight” as dictated by the “one person one vote” principle. In this ideal setting, and in the absence of any agenda-influencing powers, the “number” of votes in favour of various policy alternatives solely determines the movements within the issue space. Although this assumption enables us to capture some important limitations of majoritarian decision-making, it also limits the explanatory power of the social choice approach to the voting process, and it leaves large portions of political reality out of the model. In this section, I relax the assumption of an undifferentiated electorate and attempt to identify the economic features of the political process that cause the dominance of organized groups over the “unaggregated” electorate. The analytical goal of the present discussion is to explain the efficiency impact of organized interest groups on political law-making. I will use the expression “group dominance” to indicate the formation, within the electorate, of organized interest groups and their political superiority over the preferences of unorganized members of the electorate.

This section is organized as follows. First, I briefly summarize the reasons for group dominance, and describe the nature and efficiency implications of interest-group

576 Ibid at 121 (“In a very real sense collective choice cannot be considered as being reached by voting until relatively unanimous agreement is achieved. In so far as the attainment of such consensus is impossible, it is preferable that the actual choice processes display possible inconsistency to guaranteed consistency”).

577 See, generally, Stearns and Zywicki, Public Choice, supra note 500 at 125.
demands. Second, I inquire into how legislatures respond to the demands of organized groups; this issue is closely related to the analysis of imperfections in the principal-agent relationship between voters and legislators. In general, economic analysis suggests that actors in the political process do not have strong incentives to demand and provide for the supply of public goods. Rather, they have incentive to seek divisible private benefits and to externalize costs to the collectivity. The discussion provided in this section will be integrated in the next chapter with an analysis of the principal-agent relationship between legislators and bureaucrats.

5.3.1. The Demand Side

Political thinkers have long recognized the essential role played by interest groups in the political process. Since the birth of modern representative democracy, the elitist theory has described the political process as being dominated by organized minorities. In particular, elite theorists have firmly rejected the idea (of classic democratic theory) that a government is based on the “will of the people.” Rather, they have emphasized the average citizen’s inability to appreciate relevant political issues and the elite’s manipulative skills and power to influence. In stark contrast, the pluralist theory of democracy has offered an optimistic view of the role played by interest groups in the political process as one of the most active forms of political participation. Pluralists expect there to be many diverse and opposing interest groups exerting influence on political institutions; they insist that this multiplicity of interest groups gives rise to a vigorous competition for political influence, which is highly beneficial for democracy. In essence, both elitists and pluralists recognize the predominance of organized minorities over the unorganized electorate; they diverge on the whether this influence is beneficial to the political process. Modern political science literature has also fluctuated between these two opposing attitudes.

578 See, in particular, the classical contributions of Gaetano Mosca, Vilfredo Pareto, Roberto Michels, supra note 208.
580 However, these two perspectives should not be viewed as mutually exclusive.
Since the publication of *The Logic of Collective Action* by Mancur Olson and *The Calculus of Consent* by Buchanan and Tullock, economic literature has considered the major role of interest groups in the political process. Building on these theoretical foundations, in the 1980s economic literature on the role of organized interest groups exploded, and a consolidated analytical framework was developed based on the idea that legislation is a product of the disproportionate influence of special interest groups on political actors. For brevity of expression, I refer to this entire body of scholarship as “Interest Group Theory” (hereinafter, “IGT”). Within the IGT, a realistic view of interest group politics has been opposed to a more optimistic pluralistic view; both perspectives have significantly influenced legal scholarship. To be clear, the IGT does not provide a comprehensive explanation of politics, but it does identify some core tendencies of the political process that prove useful for process-efficiency analysis.

In this section, I draw from the extended bodies of economic and political science literature to briefly identify: (i) the reasons for the emergence of group dominance, (ii) the implications of group dominance in terms of process efficiency, and (iii) the factual predicates that are likely to exacerbate rather than attenuate interest-group influence.


582 For a useful overview of the different approaches to the legislative process within legal scholarship influenced by public-choice insights, see Stearns and Zywicki, *Public Choice*, supra note 500 at 264-295. See also Jerry Mashaw, “The Economics of Politics and the Understanding of Public Law” (1989) 65 Chicago-Kent Law Review, 123.

5.3.1.1. The Emergence of Group Dominance

Understanding the emergence of organized groups and the reasons for their political superiority over the unorganized electorate is the first step toward understanding the nature of interest-group political demands. From this perspective, it is worth briefly summarizing the major arguments identified by the IGT to explain the emergence of group dominance.

First, the IGT assumes that interest group members are narrowly self-interested.\textsuperscript{584} It follows that interest groups participate in the political process to pursue regulatory choices that maximize the economic interests of their members. That is, interest groups are not concerned with the social-welfare impact of the regulatory choices they advance. Second, the IGT emphasizes that their political activity (i.e., seeking political influence by exerting pressure on political institutions) has the character of a public good. It produces indivisible benefits that can be enjoyed by an indeterminate number of individuals; those who cannot afford to make a personal contribution can hardly be excluded from enjoying the advantages of political influence. For this reason, political efforts by members of a community are seriously burdened by collective action problems.\textsuperscript{585} In particular, the theorem of free riding predicts that since each rational individual has an incentive to free ride on the efforts of others, a collective political action is difficult to organize when it comes to promoting broadly dispersed public interests. The collective action dilemma is crucial in two respects. On the one hand, it determines the interests that are better promoted and those that are likely to remain unrepresented through the political process; on the other hand, it determines the nature of the groups that succeed in forming and achieving political influence. In essence, the free riding dilemma largely shapes the character of the demand side of the political process. The more the public interest is dispersed (i.e., the per capita stakes are low) the more collective action is unlikely and political inaction is likely. As a result, high per capita stakes are more likely to be represented in the political process than low per capita

\textsuperscript{584} I.e., they are exclusively concerned with the maximization of their selfish personal advantage.

\textsuperscript{585} Mancur Olson first explained this point; see Olson, Collective Action, supra note 208.
stake. Interest group politics “is skewed dramatically toward narrow economic
interests.”

Third, the self-interest assumption combined with the logic of collective action explains
the emergence of group dominance in the political process. Aranson and Ordeshook have
identified four activities in which the existence of organized interest groups generate
political superiority over the unorganized, undifferentiated electorate: (i) communications, (ii) bargaining, (iii) monitoring, and (iv) sanctioning. First, the
existence of an organized group enables political leaders to communicate more efficiently
and more effectively the preferences of group members to officeholders. Conversely,
group members – who have high per capita stakes, and consequently, high incentive to
gain political information compared to rationally ignorant individual citizens – can better become informed about officeholder behaviour on the public-policy issues relevant
to them. As Aranson and Ordeshook observe: “interest group members benefit from real
economies of scale, integration and specialization in carrying out political tasks. These
advantages become especially pronounced in the collection, processing and evaluation of
political information at a level that would be impossible for unorganized voters to
achieve.”

Second, politicians prefer to bargain with groups than with an
undifferentiated multitude of single individuals. The organization of interests in pressure
groups dramatically simplifies the bargaining process between politicians and voters,
which confers a bargaining advantage on groups over individual voters. Finally, groups
can more efficiently and effectively monitor and sanction the behaviour of group
members to ensure their compliance with the terms of the transactions concluded with
politicians. The same advantage concerns the monitoring of officeholders’ compliance
and their sanctioning in the event that the terms of the agreement with voters are not

587 Becker “A Theory”, supra note 208 at 392.
588 Peter H. Aranson and Peter C. Ordeshook, “The Political Bases of Public Sector Growth in a
Representative Democracy” (1978) Paper Law and Economics Center, University of Miami School of
Law, at 9 [Aranson and Ordeshook, “Political Bases”]. See also Peter H. Aronson and Peter C.
Ordeshook, “Public Interest, Private Interest, and the Democratic Polity”, at 118-121 [Aranson and
Ordeshook, “Public Interest”] in Roger Benjamin and Stephen L. Elfin, The Democratic State (Lawrence,
The University Press of Kansas, 1985)
fulfilled. The result is that “at the margin […] the preferences of organized citizens surely outweigh the preferences of those who are not organized”.\textsuperscript{589}

To conclude, according to the IGT regulatory outcomes are the responses of the political process to the preferences of various interest groups weighted by their political influence. The benefits associated with successful collective action have a “public good” nature, which generates a pervasive free-rider problem plaguing the political process. Crucially, the free-rider problem affects the various interest groups \textit{differentially}. Groups that are more effective in solving the free-rider problem gain competitive advantages over other groups in the struggle for political influence.\textsuperscript{590} This asymmetric impact determines an unequal, asymmetric \textit{access} of interest groups to the political process.

\textbf{5.3.1.2. Efficiency Implications of Group Dominance}

Once the emergence of group dominance has been explained, the next step is to identify the nature of the political demands expressed by interest groups. In particular, what is crucial to process efficiency analysis is the mechanism whereby the political process provides opportunities for the people to demand and obtain “rents” through legislation.\textsuperscript{591} This mechanism is related to the characteristics of the political process discussed earlier, including: (i) the rational \textit{ignorance} of the electorate as a whole, (ii) the \textit{majoritarian} character of political decision-making, and (iii) the \textit{collectivized} nature of the legislative outcome. These characteristics generate, as previously clarified, the externality problem \textit{intrinsic} to political law-making. Group dominance \textit{exacerbates} the externality problem as a result of the interest-group tendency toward rent-seeking. As the following discussion clarifies, the \textit{intensity} and the \textit{scope} of rent-seeking and rent-extortion mark a crucial difference between the market and the political process. As Peter Aranson (summarizing Leoni’s view on the legislative process) observes:

\begin{flushleft}
\textsuperscript{589} Aranson and Ordeshook, “Political Bases”, \textit{supra} note 588 at 11
\textsuperscript{590} Olson, \textit{Collective Action, supra} note 208; Becker “A Theory”, \textit{supra} note 208.
\textsuperscript{591} This point represents one of the most relevant distinguishing features (in terms of process efficiency) of the political process as compared to markets.
\end{flushleft}
[...] people with special interests in particular issues can pursue benefits from the political process that they cannot secure under conditions of unanimity and perfect information, which is to say, the market.592

5.3.1.2.1. The Non-Distortive Functions of Interest Groups

Before discussing the negative political externalities generated by group dominance, let me briefly emphasize the non-distortive effect generated by the presence of organized interest groups. First, joining an organization and investing resources in political influence is a way for voters to express the intensity of their preferences. In this way, interest groups mitigate the intensity problem affecting majoritarian decision-making. By facilitating the law-making process to register preference intensities, lobbying activity increases the efficiency of the political process.593 Second, interest groups increase the amount of information available to both voters and legislators. Interest groups are often better informed than lawmakers on issues that are relevant to them. This informational advantage concerns not only the preferences of group members, but also the policy issues to be regulated and the consequences of prospective regulations. This further augments the efficiency of the political process by increasing the probability of more accurate decisions. Interest groups also increase the amount of information available to voters.

5.3.1.2.2. Competition Among Interest Groups

A useful way to identify the efficiency implications of group dominance is to conceptualize the political process as a “regulatory marketplace.” We have already encountered the politics-like-market metaphor with respect to vote-trading in the legislative arena. A similar approach can be used to analyze the market dynamic underlying the relationship among interest groups and between the latter and legislators.594 The IGT proceeds along these lines and analogizes political competition to market competition.

593 Wittman, The Myth of Democratic Failure, supra note 460.
594 Traditionally, the IGT has focused on the influence exerted by pressure groups on policy makers. The post-electoral dimension is the one considered to determine the political outcome, whereas voters have been absent from the models. However, more recently, economic literature has devoted more attention to the effect of lobbying activity on electoral competition. In the pursuit of brevity, I will not discuss at length this latter aspect.
Gary Becker has developed a model in which the competition among pressure groups for political favours produces collective outcomes that converge toward Pareto-optimality. Becker applies the Coaseian logic to the political arena, which leads him to formulate an overly optimistic model of the political “invisible hand” according to which the collective influence generated by selfish pressure groups maximizes their aggregate interests. Becker’s model is constructed on the following assumptions. (a1) Interest groups are assumed to maximize their members’ income by spending the optimal amount of resources or political pressure activity. (a2) Interest groups compete with each other for political influence to obtain favourable treatment in terms of taxes and subsidies. (a3) Rent-seeking is conducted under zero-sum conditions – i.e., the amount of resources raised in taxes by the government must be equal to the amount of resources spent in subsidies. (a4) The only distortion in resources induced by taxes and subsidies are the deadweight losses associated with monopoly rents.

Under the above assumptions, competition among interest groups is conceptualized as a Cournot-Nash game in which each group takes as given the political expenditures of other groups and chooses its own optimal level of political effort. The political equilibrium depends on (i) the relative efficiency of each group in producing pressure, (ii) the effect of additional pressure of one group on other groups’ influence function, (iii) group size, (iv) the deadweight costs of tax and subsidies. The ability to control the free-rider problem among group members affects the efficiency of each group in producing political pressure. Groups that control the free-rider problem better are more efficient in political pressure and, consequently, they obtain increased subsidies or more favourable tax regimes. In Becker’s model, the deadweight losses generated by inefficient policy proposals are the major incentive for interest groups to exert political pressure. The larger the deadweight loss generated by a policy the more the losers lose relative to what the winners gain. It follows that groups that are harmed by a policy have a competitive advantage in terms of political influence. Becker suggests that the groups harmed by inefficient policies have an intrinsic advantage in the competition for influence, and groups benefiting from efficient policies have an intrinsic advantage over groups harmed.

592 Becker “A Theory”, supra note 208; Becker, “Public Policies” supra note 537.
by these policies. On the whole, analysis suggests that only efficient policies will prevail in the political market because policies that reduce deadweight losses will be politically popular.

Although it provides a useful benchmark for the analysis of competition among interest groups, Becker’s model fails to capture some essential features of the political process that are likely to lead the competition among pressure groups to the diminution of social welfare. First, Becker seems to assume equal access to the political arena. In his model, the free-rider problem is regarded as a factor that raises the costs of lobbying and, in this way, affects the political efficacy of group lobbying. However, the free rider problem not only affects the influence functions of interest groups, but more significantly, it prevents numbers of groups from accessing the political process. 596 Second, Becker assumes away the possibility of inefficiencies internal to group organizations that prevent groups from maximizing their members’ income. Once the transaction costs associated with internal group organization are factored into the model the competition among interest groups is likely to be less virtuous than that hypothesized by Becker. 597 Third, Tullock has emphasized that for Becker the only distortion in resource allocation induced by different taxes and subsidies is reflected in the deadweight losses, but he does not consider the costs of acquiring and maintaining rents (i.e., rent-seeking costs). Fourth, and relatedly, Becker does not recognize that the amount of resources invested in rent-seeking depends on “the magnitude of rents available from the political market.” 598 This correlation between rent-seeking and the structure of the political market is likely to alter the efficiency of the Cournot-Nash equilibrium of Becker’s model. This important point will become clearer as the discussion proceeds.

Beyond the criticisms mentioned above, the major limitation of Becker’s model is how it treats rent-seeking as a pure demand phenomenon. His model does not include the supply

596 Olson repeatedly emphasizes this point; see Olson, Collective Action, supra note 208
side of the political process.\textsuperscript{599} As a result, Becker fails to account for characteristics of the political process that hinder the competition in the market for regulation. For example, Becker does not consider that political outcomes are of the “all-or-nothing” type.\textsuperscript{600} This is a long-recognized, important difference in nature between individual decisions in the market and in the political arena.\textsuperscript{601} In addition, as we know, the political outcome is collectivized, which means that preferences can be expressed only on “bundles” of issues – i.e., decisions between alternative political goods must be made simultaneously.\textsuperscript{602} This empirical observation contradicts the maximizing assumption underlying Becker’s model, because discrete choices are not subject to the logic of maximization. Furthermore, political decisions are less frequently re-examined once they are made than individual decisions in private markets. For all these reasons, \textit{it cannot be assumed that an invisible hand mechanism is at work in the political process as it is in private markets.}

\subsection*{5.3.1.2.3. The Search for Private Benefits at Collective Costs}

This subsection discusses the interdependence dynamic between various interest groups and its social welfare implications.

The incentive structure embedded in the political process generates a Prisoner’s Dilemma situation in which groups seek their own private divisible benefits by giving up the potential, non-excludable, indivisible benefits of collective action. In other words, organized political minorities compete against each other to obtain favourable regulatory treatments, each of which is individually advantageous for the single group, but the sum of which entails a substantial diminution of total social welfare. This incentive structure

\textsuperscript{599} This criticism of Becker’s model is made clear by Mitchell and Munger “Economic” \textit{supra} note 597 at 536 (“[…] Becker […] does not really include a supplier or government (i.e., elected officials and bureaucrats). Government is not even a ‘black box’; it does not exist. The explanation ultimately rests upon the size of the deadweight costs and the assumptions that (1) politicians are perfectly informed and (2) have some incentive to care about the relevant magnitudes of costs and benefits as they affect relevant groups.”)

\textsuperscript{600} See, Becker “Competition” at 372-373 (“[…] groups do not entirely win or lose the competition for political influence because even heavily taxed groups can raise their influence and cut their taxes by additional expenditures on political activities. This contrasts with the all-or-nothing outcomes implied by many other formal models of political behavior, where the ‘majority’ clearly wins and the ‘minority’ clearly loses.”).

\textsuperscript{601} Buchanan, “Individual Choice”, \textit{supra} note 450.

\textsuperscript{602} Stigler, “The Theory” \textit{supra} note 581 at 10.
seriously affects the quality of the political demand expressed by interest groups, and it constitutes a major source of social welfare elimination associated with the political process. Thus it is worth illustrating this point more analytically.

Aranson and Ordeshook developed a simple model demonstrating how “each group will pursue politically a program that provides its members with a private, divisible good, and not [...] a program that provides all citizens with a public good.” Assume the following. (a1) Two interest groups – 1 and 2 – have a preference for the production, by the public sector, of a given public good from which they both benefit. (a2) Each group has a different preference for the public production of private goods. (a3) Each group has a given amount of resources to allocate in pursuing production by the public sector of either the public or the private good. (a4) All members of the electorate belong to either one of the groups such that each group’s choice is critical for the production decisions of the public sector. (a5) The public and private goods are equal in the magnitude of the benefits they provide and the costs they impose; the only difference lies in their public (indivisible)/private (divisible) nature.

Under these assumptions, each group’s decision about whether to lobby for the public or private good depends on the expectation of the other group’s decision. Consider the payoff structure for group 1. If both groups 1 and 2 lobby for the public good, group 1’s expected payoff is $B - \frac{C}{2}$. If group 1 lobbies for the public good but group 2 lobbies for its private good, group 1 still enjoys B but it pays $\frac{C}{2}$ for its public good and $\frac{C}{2}$ for group 2’s private good, such that its expected payoff is $B - \frac{C}{2} - \frac{C}{2}$, or $B - C$. Thus, from group 1’s perspective, there is a differential payoff equal to $\frac{C}{2}$ depending on group 2’s choice. It follows that the decision of group 1 will depend on its expectation of group 2’s behaviour. Now, to figure out what group 1 will expect from group 2, consider group 2’s

---

603 Aranson and Ordeshook, “Public Interest”, supra note 588 at 121-128.
payoff structure. If group 1 lobbies for the public good, group 2 receives $B - \frac{C}{2}$ in the event that it lobbies for the public good, and $B + B - \frac{C}{2} - \frac{C}{2}$ - or $2B - C$ - if it lobbies for the private good. In short, holding that group 1 lobbies for the public good, groups 2 is better off lobbying for the private good. Since the payoff structure is symmetric, group 2 also anticipates that if it decides to lobby for the public good, group 1 is better off lobbying for the private good. Thus, both groups know that if they each lobby for the public good they are both individually better off by $\frac{C}{2}$. However, they do not have any insurance that the other group will cooperate; on the contrary, they anticipate that the other group has an incentive not to cooperate. As a result, they will both decide not to lobby for the public good. Each group will receive a net benefit of $B - \frac{C}{2}$, or $B - C$. The social cost of the non-cooperative game is equal to the difference in payoff between the cooperative and the non-cooperative solution: $\left( B - \frac{C}{2} + B - \frac{C}{2} \right) - (B - C + B - C) = C$

This simple model suggests two important conclusions. First, the incentive structure is such that interest groups lobby for the public supply of private divisible benefits, rather than for the public supply of public indivisible goods. Consequently, rather then sharing the costs of the joint consumption of an indivisible public good, they end up sharing the costs of the sum of private divisible benefits pursued individually by each group. As Bruno Leoni vividly observes, the mechanism of political representation gives rise to a “potential legal war of all against all, carried on by way of legislation and representation.” Second, the same economic characteristics (indivisibility and non-rival consumption) that call for the public supply of public goods concurrently determine the incentive structure for interest groups not to mobilize in support of their public provision.

---

605 More precisely, group 2 is better off lobbying for the private good whenever $2B - C > B - \frac{C}{2}$, or $B > \frac{C}{2}$. However, if $B < \frac{C}{2}$ then the public supply of goods has a negative net benefit so that both groups will not lobby for any public provision of goods. Thus, we can assume that $B > \frac{C}{2}$. It follows that if group 1 lobbies for the public good, then group 2 is better off lobbying for the private good.

This is an extraordinarily important point. It signifies that the idea of solving collective action problems by means of the political supply of public goods needs to be reconciled with the perverse incentive structure that plagues the political demands of interest groups. As in private markets, public goods tend to be undersupplied by the political process. For this reason, interest groups will support legislation to the extent that benefits are narrowly conferred on particular interest groups. Symmetrically, legislation will provoke opposition by interest groups to the extent that costs are narrowly imposed.

5.3.1.2.4. Monopolistic Rent Costs

As repeatedly emphasized, group dominance exacerbates the externality problem embedded in majoritarian politics. To elucidate this point, let me discuss two major sources of political externalities generated by interest groups politics: (i) monopolistic rent costs, and (ii) rent-seeking costs. This subsection and the following subsection identify the source and structure of these costs.

The unequal political influence of various pressure groups entails distortions in the political allocation of resources. Political institutions use legislation as a mean of redistributing wealth among relevant interest groups according to their relative political influence. Monopolistic rent costs correspond to losses of social welfare associated with the inefficient allocation of resources. In the language of process efficiency analysis, monopolistic rent costs are a type of outcome inefficiency cost. The legislative outcome might take several forms – e.g., narrow tax exemptions, legal barriers to markets entry, protectionist regulatory policies, protective tariffs, industry subsidies,

---

607 The term “rent-seeking” was coined in Anne O. Krueger, “The Political Economy Of The Rent-Seeking Society” (1974) The American Economic Review, 291. However, it was Gordon Tullock in his paper “The Welfare Costs Of Tariffs, Monopolies, And Theft” (1967) 5 Economic Inquiry, 3, 224 who introduced the idea that the search for rents (of the kind associated with monopolies or tariffs) increases social waste, possibly by an amount equal to the value of these rents. I will discuss this point in more detail further on in the text.

608 Incidentally, the word “distortion” is imprecise unless a normative baseline is established to assess the efficient allocation of political influence. The discussion of this problem would carry us too far away. I will briefly address the importance of this issue at the end of this sub-section.

609 In this context, the relevant interest groups are those that “enjoy political influence upon regulatory outcomes”; see Dean, Interest Groups” supra note 597 at 9.

610 IGT conceptualizes government as a wealth-transfer process. ITG postulates that government does not produce any public good; rather, it is a purely redistributive process. See, generally, McCormick and Tollison, Politicians, supra note 208
and so on – each of which produces distinct economic effects. However, a common feature of these legislative outcomes is the creation of conditions for favoured interest groups to exert a monopolistic pricing strategy. Namely, political actors use legislation as a means to transfer resources from one group to another by creating the conditions, for the members of benefitting groups, to raise the price above the levels in competitive markets. The expression “monopoly rent” precisely emphasizes the difference between profits under competitive and non-competitive pricing strategies.  ⑥1

In short, from the IGT perspective, the political process distributes wealth among interest groups through the assignment of monopoly rents. The assignment of monopoly rents not only has a redistributive effect, it also generates a net loss of social welfare, usually referred to as a monopoly “deadweight loss.” Let me be more precise on this point.

Let \( p^* \) denote the market price, and \( p^m \) the monopoly price. By definition, under a monopolistic market, \( p^m > p^* \). Let \( q^* \) denote the quantity sold at \( p^* \) and \( q^m \) the quantity sold at \( p^m \). ⑥12 Monopoly theory tells us that when the monopolist raises prices above market level – i.e., above the point where the price equals the marginal production cost – two effects are produced. First, those consumers willing to pay an amount not lower than the monopoly price still buy, but they pay \( p^m \) instead of \( p^* \). This has a redistributive implication. In fact, these consumers lose the surplus that they would have realized at \( p^* \), which is equal to \( q^m(p^m - p^*) \). This foregone consumer surplus is not lost, but it is transferred from the consumers to the monopolist. In short, the price increase (i.e., the monopoly rent) results in a direct wealth transfer from consumers to the monopolist. Second, a number of transactions that would have occurred at \( p^* \) do not occur at \( p^m \). This is because those consumers would have bought at \( p^* \) but who will not at \( p^m \) do not buy. Thus, part of the quantity that would have been sold in the absence of the monopolistic

  ⑥1 More generally, a monopoly rent can be defined as the difference between the actual return of an economic activity and the opportunity cost of the wealth-producing asset. Differently stated, monopoly rents arise each time an economic activity generates wealth in excess of available alternative market opportunities. For a helpful explanation of the relationship between the concepts of “opportunity cost” and “monopoly rent,” see Stearns and Zywicki, Public Choice, supra note 500 at 35-37 and 47.

  ⑥2 For simplicity, I implicitly assume a constant marginal cost of production (i.e., I assume a flat supply curve).
pricing strategy remains unsold; consequently, the consumer surplus \((q^* - q^m)p^*\) that would have been enjoyed by consumers through purchasing the unsold quantity is lost.

It is easy to conclude that the assignment of a monopoly rent generates a net loss of social welfare. The monopolist pricing strategy generates a loss of consumer surplus that equals the sum of the monopoly deadweight loss and the monopoly rent. This means that the monopolist receives less than the consumers lose. In the history of the US, there are many historical examples of regulatory policies aimed at eliminating monopolistic rents by encouraging multiple suppliers based on the idea that competition among multiple providers would enhance consumer surplus. Just to mention two classical historical cases in which the promotion of competition among providers has dramatically increased the efficiency of the industrial sector in question, consider the Airline Deregulation Act of 1978 and the Telecommunication Act of 1996.\(^{613}\)

5.3.1.2.5. Rent-Seeking Costs
The ITG assumes that interest groups behave rationally in pursuing maximization of their members’ income. This implies that interest groups are willing to invest resources to secure and maintain monopoly rents up to the expected values of the rents available in the political market. This obvious consideration has important social welfare implications. The inefficiency of interest-group politics (as depicted by the IGT) is not limited to outcome inefficiency associated with monopoly pricing strategies. In addition to monopolistic rent costs, there are considerable costs associated with the process of rent-seeking. In the effort to secure and maintain politically conferred rents, it is rational for selfish maximizing interest groups to invest resources in lobbying, political campaigns, and other rent-seeking activities up to expected value of available rents. In this way, interest groups face an incentive to “dissipate” the value of the monopoly rent. In equilibrium, the possibility of securing monopoly rents will continue to attract rent-

\(^{613}\) A legal economic analysis of these two statutes is provided in Joseph D. Kearney and Thomas Merrill, “The Great Transformation of Regulated Industries” (1998) Columb. L. Rev., 1323 [Kearney and Merrill, “Great Transformation”].
seeking investments until the value of the monopoly rent is \textit{completely} dissipated.\footnote{The “exact dissipation hypothesis” – in which the total amount of rent-seeking expenses equal the magnitude of the available rent – has been formulated by Richard Posner in his paper “Social Costs Of Monopoly And Regulation” (1974) 83 4 The Journal of Political Economy, 807 [Posner, “Social Costs”]. This point is discussed in Tullock, “The Efficient Rent-Seeking” [Tullock, “Efficient rent-seeking”] in James M. Buchanan, Robert D. Tollison, and Gordon Tullock, \textit{Toward a Theory of Rent-Seeking Society} (College Station : Texas A and M University, 1980) [“Toward a Theory of Rent-Seeking”] 97-112; Tullock, Rent-Seeking \textit{supra} note 598 at 63-66. See also, Charles K. Rowley, Robert D. Tollison, and Gordon Tullock, eds., \textit{The Political Economy Of Rent-Seeking} (Boston: Kluver Academic Publishers, 1988) [Rowley et al., “Political Economy”].} According to the so-called Virginia School of Political Economy,\footnote{See, generally, Gordon Tullock, \textit{Virginia Political Economy}, in The Selected Works of Gordon Tullock (Indianapolis: Liberty Fund, 2005), vol. 1} the process of \textit{rent-dissipation} represents the more substantial social cost generated by the monopoly power. Legislation is inefficient not just for the outcome inefficiencies associated with captious legal rules, but also for the huge amount of resources deployed in rent-seeking activities.

The foregoing analysis raises the central question of how much of the available rent is dissipated by interest groups – i.e., what is the ratio of rent-seeking expenses to the size of available rent in political markets. There is no clear, generalizable answer to this question.\footnote{This problem is addressed in William J. Corcoran, “Long-Run Equilibrium and Total Expenditures” in Rowley et al., \textit{Political Economy}, \textit{supra} note 614 at 95-101; William J. Corcoran and Grodon V. Karels, “Rent-Seeking Behavior in the Long Run” in Rowley et al., \textit{Political Economy}, supra note 614 at 107-126; William R. Dougan and James M. Snyder “Are Rents Fully Fissipated?” (1993) 77 Public Choice, 4, 793.} However, public choice literature has identified many factors that are likely to affect the amount of rent-seeking costs, including the following: (i) the number of interest groups competing to secure and maintain rents, (ii) the degree of competition between them, (iii) the structure of the rent-seeking production function (e.g., economies or diseconomies of scale in lobbying activities), (iv) the presence of asymmetric information between interest groups, (v) the attitude to risk of interest groups, (vi) the presence of rent-extraction by politicians, and so on. In addition, it can be reasonably predicted that both the zero rent-seeking costs hypothesis and complete-dissipation outcomes are unlikely to occur in reality. Complete-dissipation is unlikely to the extent that rent-seekers recognize that “they are engaged in a game \textit{ex ante} which lowers their expected net wealth.”\footnote{Tullock, “Rent-Seeking”, \textit{supra} note 598 at 64.} The zero rent-seeking hypotheses\footnote{In this case the only effect of rent-seeking are those related to monopoly rent.} can hold only under the
unrealistic assumptions of either no rent-seeking at all or perfectly competitive rent-seeking activity.\textsuperscript{619}

The economic impact of the rent-seeking process is not limited to the dissipation of the rent. The \textit{opportunity cost} of resources dedicated to rent-seeking represents social waste. To the extent that time, effort and resources invested in rent-seeking are diverted from more productive activities, rent-seeking imposes a \textit{deadweight} cost on the economy as a whole.\textsuperscript{620} There are various sources of deadweight losses that rent-seeking imposes on the economy. Let me mention two of the most relevant. First, the resources invested by \textit{unsuccessful} rent-seekers represent a cost that is not offset by anyone’s gain. Second, the resources invested by rent-seekers that successfully capture rents are also in part a deadweight loss to society to the extent that they do not improve productivity in the economic activity generating monopoly rent. From the individual standpoint, investing in rent-seeking might be a negative, zero, or positive sum game; but from the social standpoint rent-seeking is a \textit{negative} sum game.\textsuperscript{621} Third, it has been demonstrated that in countries where political authorities are more responsive to rent-seeking efforts the costs of rents increase dramatically. In these cases, rent-seeking works as an economic barrier to entry for new firms, because the large costs of obtaining rents makes the investment in rent-seeking a much less attractive option.\textsuperscript{622}

To conclude, rent-seeking theory helps to demonstrate that the welfare losses imposed on society by an excessive recourse to political law-making can be enormous. Olson, for example, has argued that the influence of interest groups on the political process can


\textsuperscript{620} See Buchanan, “Rent-seeking and Profit Seeking” in \textit{Collected Works of James Buchanan, supra} note 450 Vol. 1 at 108 (“Resources devoted to efforts to curry the queen’s favour might be used to produce valued goods and services elsewhere in the economy, whereas nothing of net value is produced by rent-seeking.”). See also Posner, “Social Costs”, \textit{supra} note 614.


generate external costs for the economy as a whole to the point of generating unemployment, dampening industries’ productive potential, and inhibiting the economic growth of nations.\textsuperscript{623} From the same perspective, Tullock formulates an explanatory hypothesis of the Industrial Revolution as the consequence of reduced room for rent-seeking consequent to the English Revolution.\textsuperscript{624} Tullock emphasizes that one important institutional development of the Revolution was the central role acquired, within the English judicial system, by the common law courts in which juries largely made decisions. For well-identified historical reasons, juries were reluctant to convict in rent cases; as Tullock observes: “The jury would tend to feel that the person who violated this rent was either a person to be congratulated or in any event somebody who had done nothing particularly awful.” The use of the juries, along with other institutional developments, made rent-seeking activity much less attractive because the rents could not be enforced easily and so they lost much of their economic value. According to Tullock, the reduction of rent-seeking, as an unanticipated and unintended consequence of the English Revolution, was one major factor in the development of the Industrial Revolution in England.

More recently, many empirical studies have attempted to assess the social costs of rent-seeking. These attempts face a serious and difficult methodological problem: providing a definition and an objective measure of rent-seeking costs. In fact, the identification of rent-seeking costs requires an underlying definition of the best, most productive use of available resources. Claiming that the amount of resource $x$ would have been better invested in activity B rather than its actual investment in activity A is a counterfactual hypothetical assessment based on strong normative assumptions. It must be demonstrated that the investment of $x$ in A reduces a society’s production possibility frontier compared to what might have been experienced if $x$ had been invested in B. Thus, it must be recognized that, as Tollison lucidly observes, “as with all empirical work, these various

\textsuperscript{623} Mancur Olson, \textit{The Rise And Decline Of Nations: Economic Growth, Stagflation, and Social Rigidities} (Yale University Press, 1984) [Olson, “Rise and Decline”].

approaches are only as good as the theories and models upon which they are based.’’

Furthermore, in many cases investment in rent-seeking activities is very difficult to trace. However, once one is aware of these methodological problems, these empirical analyses are useful in that they provide a more precise idea of how large the inefficiency losses associated with rent-seeking and rent-dissipation can be in real-world economies.

Laband and Sophocleus attempted to quantify, sector-by-sector, the amount of rent-seeking costs in the U.S. economy in 1985. They tried to measure the magnitude of the aggregate resource investment by entrepreneurs in transfers of wealth that do not involve mutually beneficial exchange. Such non-exchange (and non-charitable) transfers of wealth include, but are not limited to, “theft and theft prevention, fraud and government-sponsored transfer programs.” So, for example, they counted expenditures on such items as locks, safety deposit boxes, burglar alarms, computer security, airport safety equipment, contested divorces, strikes, insurance, police, military expenditures, and expenditures on lobbyists as activities driven by rent-seeking. The conclusion of their work is that a significant fraction of all economic activity conducted in the U.S. is concerned with the transfer of resources driven by rent-seeking or rent-protecting incentives, as opposed to the creation of wealth. Crucially, they also find that a relatively large amount of transfer activity is subject to the political process. In short, the transfer of activity through the political process absorbs a significant fraction of all economic resources in the U.S..

5.3.1.2.6. Rent-Seeking Environments

Once the economic nature of interest-group demand has been identified along with its efficiency implications, the next important step is to identify the conditions of the regulatory environments affecting the degree of potential influence by interest groups. As Komesar observes: “[…] there are significant sources of variation or gradation in the dominance of the few and the dormancy of the many. […] [T]he degrees of majority

626 A survey of this research is provided by Stephen P. Magee, William A. Brock and Leslie Young, eds. Blackhole Tariffs and Endogenous Policy Theory. (New York: Cambridge University Press, 1989) at chapter 15.
dormancy vary across political issues and political jurisdictions.\textsuperscript{627} This is an important point for process efficiency analysis. The circumstances under which the activity of organized groups is especially likely to be effective are potential sources of the comparative disadvantages of political law-making versus other sources of law.

It is possible to identify the following factors.\textsuperscript{628} First, the discussion of the emergence of group dominance suggests that the structure of per capita stakes is a decisive determinant of the effectiveness of collective action. Holding constant the ratio between the magnitudes of the stakes of opposing interest groups, the distribution of the per capita stakes affects the relative strength of the groups. The groups with higher average per capita stakes are likely to be more effective than those with lower ones, even if the absolute aggregate stakes of the latter are higher. High per capita minorities tend to dominate low per capita minorities. Second, the structure of the political conflicts is also important. An organized group is likely to be more effective when acting as a veto player rather than as a promoter of a new piece of legislation: groups that support the status quo have an advantage over groups that advocate legal change.\textsuperscript{629} Third, the characteristics of the market for regulation affect an interest group’s likelihood of success. For example, a group is more likely to succeed in rent-seeking where there are no competing interest groups on an issue. Or, if the different groups on one side of a controversial legislative issue are unified and coordinated they are more likely to succeed. Fourth, the visibility of an issue is also important. “Interest organizations’ influence is likely to be less substantial on highly visible issues that engage public passions or media coverage and on issues in which there are strong competing ideological, partisan, or constituency pressures.”\textsuperscript{630} In addition, the degree of complexity of the regulatory issue is important. Complexity raises information costs, which increases the costs of mobilizing group members, thereby heightening the cost threshold for access to the political process. Fifth, interest groups are more likely to exert influence on single amendments than on entire pieces of legislation.

\begin{footnotesize}
\textsuperscript{627} Komesar, Imperfect Alternatives, supra note 7 at 73.
\textsuperscript{628} Schlozman and Tierney, Organized Interests, supra note 491.
\textsuperscript{630} Schlozman and Tierney, Organized Interests, supra note 491 at 314. On this point see also Trebilcock, Dealing with Losers, supra note 37 at Chapter 2.
\end{footnotesize}
There are many other factors influencing the ability of groups to influence law-making through politics. What is important to emphasize here is that process efficiency analysis should be based on a careful investigation on a case-by-case basis of factors such as the nature of the demand, the nature of the regulatory issue, the structure of the political conflict and of the market for regulation, and the size and distribution of opposing interests.

5.3.1.2.7. Normative Implications
The IGT provides powerful tools for the positive analysis of the political process. In particular, it helps explain the emergence of interest groups, and it allows one to predict which organized interests are more likely to exert influence on political institutions.\textsuperscript{631} Drawing normative prescriptions from the IGT is more problematic. The major criticism lies in the lack of a normative baseline for assessing the degree of influence of interest groups. That is, the IGT does not address the problem of defining a normative standard of political legitimacy.\textsuperscript{632} I do not have useful suggestions for resolving this delicate problem, which has been the subject of significant contributions in the literature.\textsuperscript{633} However, once one is aware of this methodological baseline problem, the framework provided by the IGT offers useful insights for assessing the advantages and disadvantages of interest group politics in terms of process efficiency.

The politics of interest group is problematic in terms of agency efficiency for two reasons. First, as previously emphasized, since the free-rider problem affects interest groups differentially, some interests are more likely to be represented in the political process than others. In economic terms, this implies that some groups internalize the benefits of legislation and externalize the costs on others. This is problematic to the extent that an unequal representation of interests is determined by the different impact of

\textsuperscript{631} More recently, IGT has been criticized by Croley, Regulation, supra note 583.
\textsuperscript{632} To qualify the political influence of an interest group as “disproportionate” one has to first define an “ideal claimant to political influence”; see, see Dean, Interest Groups” supra note 597 at 198.
the free-rider problem rather than being related to some “normative” criteria of political legitimacy. Interest groups exacerbate the externality problem embedded in majoritarian decision-making because they tend to use the political process to externalize the costs of private benefits. Furthermore, as the preceding discussion on rent-seeking has demonstrated, interest group politics entails the dissipation of economic resources, imposing net welfare losses on society.

5.3.2. The Supply Side

The political process can be usefully conceptualized as a chain of agency relations. Along this chain, each principal appoints an agent to act on his behalf; in turn, the agent becomes a principal and so forth, up to the final agent. The principal-agent paradigm enables us to examine the incentive structure inducing political actors to be responsive to the preferences of the people subject to law. As discussed in chapter 4, the centralization of law-making generates a systemic misalignment of the marginal benefits and costs the lawmaker and the people subject to law face. This gives rise to the conditions for a pervasive agency problem affecting centralized law-making processes. In this subsection, I briefly analyze some additional features of the agency problem in political law-making processes. The line of inquiry is two-fold: investigating how politicians respond to the political demand of voters and interest groups, and determining the efficiency implications on the supply-side of interest-group politics. This subsection focuses on the relationship between voters and legislators. The next chapter will discuss at length the relationship between legislators and bureaucrats.

634 “The central, essential perspective or outlook of a democratic society is surely one in which the relationship between citizens and public officials is one of principals and agents, at least in an economically meaningful sense even if not in a legal recognized sense. A government, after all, is owned by its individual citizens, for they are residual claimants with respect to the conduct of government, while its operation resides with a relatively few public officials”, Richard E. Wagner, “Agency, Economic Calculation, and Constitutional Construction” in Rowley et al., Political Economy, supra note 614 at 428.

635 In real-world processes this chain includes a number of actors: “voters may have difficulty monitoring and controlling a legislature, a legislature may have similar difficulties monitoring and controlling its committees, and so on through the relationships between oversight committees and regulatory agencies and between agency heads or commissioners and their bureaucracies”, Michael E. Levine, and Jennifer L. Forrence, “Regulatory Capture, Public Interest, And The Public Agenda: Toward A Synthesis” (1990) JL Econ and Org, 167 at 171 [Levine and Florrence, “Regulatory Capture”].
5.3.2.1. Agency Slack

Agency theory has long recognized\(^{636}\) that there is a discrepancy between what the principal wants and what the agent actually delivers (i.e., there is often inefficiency in the principal-agent relationship). When the principal is prevented from observing the agent’s behaviour, then the agent has no incentive to conform his or her behaviour to the principal’s preferences and goals. The “nonmonitored situations in the principal-agency relationship of voter and representative”\(^{637}\) are referred to in the literature as “agency slack.”\(^{638}\) Generally, agency slack originates from three structural conditions hampering the principal-agent “cooperative game”:

(i) the difficulty of specifying \textit{ex ante} the provisions of the “agency contract” establishing the terms and conditions of the principal-agent relationship (i.e., the “contractual incompleteness” problem);

(ii) the presence of \textit{asymmetries} of information between the principal and the agent, which entails a lack of information by both parties as to the other’s behaviour and needs (i.e., the asymmetric information problem);

(iii) the presence of high costs for the principal of \textit{monitoring} and \textit{controlling} the agent’s behaviour, which prevents the principal from rewarding good performance and sanctioning bad performance (i.e., the monitoring-cost problem).

Based on the analysis of these problems, agency theory identifies the inefficiency associated with the principal-agency relation between voters and legislators. It will become clearer, as the discussion proceeds, that beyond the “natural” agency costs associated with the structure of the agency relationship between politicians and voters, the political process involves four other categories of agency costs: monopolistic-rent costs, strategically augmented costs, rent-seeking costs, and costs related to the extractive attitude of politicians. In the following sub-sections I discuss all these categories; in each subsection the analysis is organized around three theoretical points: (i) the \textit{mechanisms},


\(^{637}\) Levine and Florrence, “Regulatory Capture” \textit{supra} note 635 n.11

\(^{638}\) \textit{Ibid.}
(ii) the *determinants*, and (iii) the *efficiency implications* of the agency problem affecting the political process.

5.3.2.1.1. Mechanisms

The causal mechanisms generating agency-slack in the voter-politician relationship can be explained by examining the characteristics of both the demand-side and the output of the political process. It is the institutional structure of political law-making that determines the scope for opportunistic behaviour by legislators.

(i) “*Contractual*” Incompleteness.

First, the agency relation between voters and politicians does not rest on a formal, enforceable mandate. Once elected, the successful candidate is not subject to any formal obligation concerning any specific parliamentary choice. On the other hand, the elected politician’s loyalty to the political mandate of voters, or to acting in accordance with voter interests, does not create any obligation for voters to support the loyal candidate in the next round of elections. The absence of a formal, enforceable agreement between voters and politicians reduces significantly both the principal’s incentives to monitor agents and the agent’s incentives not to “shirk.” Furthermore, it must be emphasized that even in the presence of an ideally enforceable mandate, a complete (or at least “well specified”) contract would be impossible to achieve at low cost given the “open-ended mandate of legislatures.”

(ii) *Rational Ignorance of Principals*.

Agency slack is to a great extent the result of voter “rational ignorance.” Agency theory suggests that when the principal does not have access to accurate information on an agent’s activity, the agent has an incentive to behave *opportunistically*. Since in many cases – as discussed earlier – the rational voter has an incentive to remain ignorant or indulge in irrationality, legislators face strong incentives to “shirk” and to act in accordance with their own private interests. By contrast, when slack-reducing

---

639 See Lewis A. Kornhauser, “Aggregate Rationality in Adjudication and Legislation” (2008) 7 1 Politics, Philosophy and Economics, 5 at 12-18 [Kornhauser “Aggregate Rationality”] at 18. Kornhauser emphasizes that the open-ended nature of the mandate for legislators is related to the broad scope of legislative power, and that this is an important difference between courts and legislatures.
mechanisms are at work because voter interest on an issue is high, the room for shirking by politicians diminishes dramatically. As Caplan puts it:

If [voters] care deeply about an issue [...] politicians have almost no slack. One word costs them the election. In contrast, if voters find a subject boring [...] if emotions and ideology provide little guidance, their so-called representatives have “wiggle room” to manoeuvre.

The agency relation between voters and politicians is collective in nature; the number of principals is large, while the number of agents is relatively small. This entails joint-consumption by principals of the benefits of monitoring the agents. Monitoring is a public good; as such, it is plagued by collective action problems that are exacerbated by the increasing number of principals. As Komesar states:

Smaller jurisdictions or polities and, therefore, smaller numbers of voters and fewer issues for majoritarian response may decrease the associated connection between majority positions and free riding and increase the probability of majoritarian response.

(iii) Interest Group Dominance.

It should also be considered that, due to group dominance, principals are not undifferentiated. Voter rational ignorance and monitoring difficulties are often exacerbated by group dominance. First, interest groups exert influence on political outcomes by influencing voter information (e.g., through investment in policy advertising, investment of political influence on media coverage of issues that interest them, and so on). Since voters are either ignorant or irrational, they can be easily manipulated by the influence of organized special-interest groups. Second, the presence

---

640 This point is made clear by Caplan, *The Myth of the Rational Voter*, supra note 464 at 20 (“How strictly do elections constrain politicians? My view is that it depends on voters themselves. If they care deeply about an issue […] politicians have almost no slack. One word costs them the election. In contrast, if voters find a subject boring […] if emotions and ideology provide little guidance, their so-called representatives have “wiggle room” to manoeuvre.”)

641 Komesar, *Imperfect Alternatives*, supra note 7 at 73.

of group dominance suggests that principals are not undifferentiated. It might be objected that, should the generality of voters overcome the information and monitoring costs and arrange some form of accountability, the agent runs the risk of a public backlash by favouring interest groups. Although this is certainly a possible scenario, it is not the most likely one. Unorganized individual voters do not have incentives to become informed enough to identify and to reward agents who do not shirk or to punish those who do. By contrast, special interest groups have greater incentives to become informed enough to monitor agents and to reward those who provide them with favourable regulatory treatments. Normally, the lawmaker has slack and pursues political outcomes that benefit special interests who are willing to offer greater political contributions. These considerations eschew the idea of a benevolent, social-welfare maximizing lawmaker. The likelihood that a lawmaker will maximize a “political influence function” depends on the utility functions of relevant interest groups weighted by their political strength.\(^{643}\)

Interest groups impact political outcomes not solely by influencing the public opinion formation process, but also by acting directly on elected politicians (affecting their incentives or the quality of the information available to them). First, agents who wish to be re-elected cannot disregard interest group concerns and preferences; otherwise, they run the risk that special interests favouring alternative outcomes will engage in political efforts to remove them from power and to install someone willing to confer the desired regulatory benefit. Second, interest groups can distort the flow of information to the lawmaker, despite the public-spirited attitude of office-holders. This point, too, is well emphasized by Komesar:

> The now partially ignorant, but still public-interested, official must depend on others to provide information on the desires of his or her constituents and the relationships between those desires and policy alternatives. Assuming [...] that concentrated groups have significant advantages in understanding and effectively representing their viewpoints, the public-interested official will garner a distorted picture of constituency desires and public

\(^{643}\) See, Becker “A Theory” supra note 208 at 373-376. Furthermore, since politicians cannot disregard special interest groups, the overall set of agents is likely to be selected in such a manner as to be more responsive to the interests of those groups of agents who are most influential in the electoral process. (See Komesar, Imperfect Alternatives, supra note 7 at 63: “[…] overrepresented interest groups can alter outcomes through the election process”)

(iv) **Agency Implications of the Social Choice Problem.** The social choice problem is in many ways related to the agency problem. First, the absence (in certain situations) of electoral equilibrium results in an absence of equilibrium *strategies* among candidates, to which politicians respond by adopting a strategy of *ambiguity*. “Indeed, in itself, ambiguity represents breakdown in the connection between citizens’ preferences and candidates’ election platforms.” Second, as previously emphasized, social intransitivities generate room for agenda manipulation and misrepresentation of preferences. Third, equilibrium-inducing devices (e.g., calendaring rules, committees structures) raise the costs of passing bills, which exacerbates the influencing power of interest groups. Stearns and Zywicki argue that these institutional arrangements raise the costs of majoritarian politics, with the result that they increase the tendency of high *per capita* minorities to dominate in the legislative process. Fourth, vote-trading among legislators often “jeopardizes the accountability of legislators.” In many cases the complexity of the bargaining logic underlying trade choices is not reflected in voting

---


646 Aranson and Ordeshook, “Public Interest”, supra note 588 at 117.


648 See Cooter, *Strategic Constitution*, supra note 219 at 64.
records of elected representatives, and this might undermine incentives to vote trade so as to increase voter satisfaction.  

5.3.2.1.2. Determinants

The preceding considerations suggest agency slack is the normal state of affairs in real-world political processes. Here, I briefly mention the institutional features of the political process that reduce agency slack. It should be recognized that every mechanism that reduces the information, organization and monitoring costs voters face can be regarded as an agency-slack reducing mechanism. Lavine and Forrence have identified the following institutional features: (a) incumbent self-publicity, (b) political competition, (c) public-policy intelligentsia, and (d) news-media. For the purposes of this study, mechanism (a) and (b) must be emphasized. First, in some cases legislators can receive more political support by investing political resources to reduce agency-slack (and advocating for the general interest), rather than by investing in accommodating special-interest group demands. The choice between agency-slack reduction and interest-group accommodation depends on the relative costs and benefits of each activity. So, for example, if the issue becomes salient and the general-interest policy enjoys significant public attention and support, it might be more rewarding for politicians to invest in agency-slack reduction. Second, the degree of political competition plays a crucial role in determining agency costs. In general, the higher the degree of political competition, the greater the opportunities will be for monitoring legislators’ behaviour, publicizing, and promoting general-interest politics.

---

649 This point is well explained in Robert Cooter, The Strategic Constitution, Web Version, January 1999 at 109 through the following illustrative example: “[…] assume that two bills, A and B, are pending in the legislature. Also assume the constituents of a certain district favor both bills, but passing A is more important to them than passing B. If the district’s representative votes for both bills, her constituents will approve of her voting record. Assume, however, that both bills will be defeated unless she trades her vote on B to obtain a vote on A. Now the legislator faces a dilemma. Her constituents care more about A than B, so they would presumably want her to trade votes in order to secure the passage of A. If, however, she does trade votes, she will go on record as voting against B. Since her constituents favor B, her opponent in the next election will tell the voters that she voted against B. Anticipating this possibility, the legislator may refuse to trade votes, thus giving her constituents their preferred symbols, rather than their preferred legislation. In general, legislative bargains obscure the information provided to citizens by the legislator’s voting record.”.

650 Lavine and Florrence “Regulatory” supra note 637 at 187.

651 Ibid. at 186
Another set of “slack-reducing” factors includes those institutional mechanisms that increase the price of political pressure, thus reducing the resources invested in lobbying. Cooter identifies the following mechanisms capable of reducing the amount of resources invested in the long run by firms in lobbying: (a) the division of power, (b) the bicameral legislature, (c) the fragmentation of political parties, (d) burdensome rules governing legislative proceedings, (e) public financing of campaigns. As Cooter points out, these institutional mechanisms raise the cost of “purchasing legislation,” thereby generating the effect of reducing private actors’ investments in lobbying. However, it must be considered that raising the costs of political influence might produce the countervailing effect of exacerbating group dominance, thereby increasing outcome-allocative inefficiency. Maxwell Stearns contends that these institutional arrangements – referred to as “veto gates” or “negative legislative check-points” – “are in place to slow down or to stop legislation that benefits the public at large at a cost borne largely or entirely by a narrow interest group.” Whether the Cooter or Stearn effect prevails is likely to depend on many factors, including those identified by Lavine and Forrence.

5.3.2.1.3. Efficiency Implications

Agency slack affects interest groups differentially because legislators have incentives to shirk with respect to groups suffering from greater monitoring costs. As Mayhew has clarified, legislators seek to “claim credit” and to “avoid blame”. Thus, they have incentives to promote legislation that concentrates benefits on groups with greater monitoring ability, and to spread the costs on groups facing greater agency slack problems. This consideration elucidates a crucial point: ‘group dominance’ on the demand side and ‘agency-slack’ on the supply side drive the political process in the same direction. On the demand side, special interest groups face greater incentives to lobby either for or against proposed legislation when the related benefits and costs are concentrated on specific, well-identifiable interest groups. On the contrary, when

652 Cooter, Strategic Constitution, supra note 219 at 65
653 Ibid.
655 Stearns and Zywicki, Public Choice, supra note 500 at 71.
conferred benefits and imposed costs are widely distributed incentives to engage in lobbying activities are diminished. On the supply side, legislators can more effectively claim credit by providing narrowly concentrated benefits and avoid blame by imposing thinly spread costs. The pursuit of particularized legislation is more politically rewarding for any legislator than the pursuit of universalistic, general-interest legislation.

To summarize: under ordinary circumstances, high principal-agent slack plagues the relationship between voters and legislation; this reinforces the demand-driven tendency toward group dominance, and generates incentives for politicians to produce legislation that confers narrowly concentrated benefits and that imposes widely distributed costs.\textsuperscript{657} However, when (i) factors exacerbating group dominance are mitigated,\textsuperscript{658} and (ii) factors reducing agency-slack are at work,\textsuperscript{659} politicians have greater incentives to produce legislation conferring widely distributed benefits to the extent that the voting public becomes informed enough of benefits so as to reward legislators pursuing the public-interest. When widely distributed benefits are accompanied by narrowly concentrated costs legislators face a difficult political trade-off between claiming credit for potential beneficiaries and avoiding blame by prospective losers. In these cases, legislators are likely to strategically delegate law-making power to either courts or bureaucratic authorities. I will expand on this point subsequently in subsection section 6.2.2.

Ultimately, legislative outcomes are largely dependent on (1) the structure of the organization and information costs interest groups face – i.e., the relative intensity of the free-rider problems among interest groups, and (2) the distribution of the benefits and costs generated by the legislative outcome – i.e., the distribution of stakes.

\textsuperscript{658} See supra at 5.3.1.2.6
\textsuperscript{659} See supra at 5.3.2.1.2
5.3.2.2. The Augmentation of Transaction Costs

5.3.2.2.1. Mechanisms

The foregoing discussion has focused on monitoring and information costs originating from the institutional structure of the agency relationship between voters and legislators. These costs can be called “natural” or “unavoidable” agency costs. Since they are related to the structural features of the political process, agents cannot eliminate them even assuming transaction-cost minimizing behaviour. In a certain sense these are exogenous costs that “would remain [even] if everyone [in the political process] tried to minimize transaction-cost impediments to political exchange.”

Such natural transaction costs change as society’s technological, scientific and organizational know-how changes.

However, in real-word political processes legislators are not limited to capitalizing on the shirking opportunities provided by the high level of “natural” transaction costs. Politicians often use law-making institutional mechanisms to strategically manipulate the transaction costs faced by voters or political opponents in resisting governmental activities. Thus, while “natural” agency costs are exogenous, self-interested legislators endogenously produce these costs to inhibit resistance to their preferred policies. The deliberate, artificial increases in transaction costs by politicians take many different forms; Twight has developed, for explanatory purposes, the unifying category of “transaction-cost augmentation” to refer to the transaction that exceeds natural agency costs and that are strategically generated by governing authorities.

---

661 Ibid.
662 Ibid.
In this section, I borrow some of the analytical insights provided by Twight’s taxonomy of governmental transaction-cost augmentation, and adapt them to the specific context of political law-making. Twight identifies two forms of strategic manipulation of transaction costs: (i) manipulation of information costs – i.e., altering “people’s perception of the costs and benefits of governmental activities,” and (ii) manipulation of agreement and enforcement costs – i.e., increasing people’s costs of “taking political action based upon these perceptions.” In either case, legislators reduce the ability of their adversaries (voters and political opponents) to secure political change, thereby increasing their own ability to implement their personal goals and preferences. An example of technique (i) is the use of methods of taxation that alter tax-payers’ perception of the tax burden imposed on them, or the use of methods of distributing public resources that distort people’s perception of the associated costs and benefits. An example of technique (ii) is an increase in the costs of accessing the political arena (e.g., laws increasing the number of signatures to be collected by new political parties to participate in the electoral process), or an increase in the costs of challenging the legislative outcome via judicial or administrative processes. In addition, as I will discuss in the next chapter, a common technique is the “strategic delegation” of law-making authority to administrative agencies.

Although each law-making institution is potentially plagued by cost-augmentation, the political process suffers from comparative disadvantage versus adjudication and market processes. To understand why, it is useful to think of the transaction costs faced by principals/voters in controlling and monitoring agents/legislators as a “constraint” on the review of lawmakers’ activity. The efficiency of any agency relationship largely depends on the fact that constraints on the review of the agent’s activity are outside the

---

664 The taxonomy mentioned in the text is developed in Twight, “Political”, supra note 660 at 202-212.
667 Ibid.
668 Twight, “Political”, supra note 660 at 208.
669 Ibid. at 205.
670 Ibid. at 206.
671 This point is emphasized in Twight, “Government”, supra note 663 at 131
sphere of the agent’s control. Now, one of the most distinctive features of the political process is that, unlike market and adjudication, the principal/lawmaker can strategically manipulate the constraints on the review of his or her own activity. By increasing the review costs lawmakers increase their autonomy and broaden the scope for their opportunistic behaviour.

Two characteristics of the political process allow the lawmaker to engage in the practice of transaction-cost augmentation. First, there is often an incentive for interest groups to support legislator manipulation of transaction costs. The strategic “alliance” between the lawmaker and special interest groups is strictly correlated to the predictability of the impact of legislation on economic actors. As explained earlier, the legal rules produced through the political process are likely to confer narrowly focused benefits on specific portions of the electorate. In this context, the beneficiaries of interest-group legislation are able to appreciate ex ante (i.e., during the period preceding a bill approval) the personal impact of proposed legal rules. It follows that interest groups are willing to support the augmentation of transaction costs by politicians, thereby facilitating the passing of laws favorable to them and hindering opposition by the losing majority.

Second, an important potential constraint on transaction-cost augmentation is the degree of pluralism of the sources of law. As discussed at length in chapter 4, the centralization of the law-making process raises the exit costs for those affected by inefficient legal rules. These increased opt-out costs offer lawmakers a lot of opportunities to augment transaction costs because they increase the asset specificity problem for the people subject to law with respect to the existing rules. In short, the ex ante centralized structure of political law-making generates (i) asset specificity for voters by increasing exit costs and, (ii) the room for opportunistic behavior by legislators. That is, interest

---

672 I have already underlined the importance – emphasized by both philosophic and economic theory – of the ex ante symmetric institutional setting as a structural condition for ensuring the impartiality of social decision-making. See text accompanying note 160
673 Ibid. at 27.
674 Ibid.
675 See the discussion on the exit-downgrading effect in section 4.4.1.1.
group dominance exacerbates politicians’ shirking by supporting their transaction-cost-augmentation strategy.

5.3.2.2.2. Determinants
Following Twight, it is possible to identify some factors influencing legislators to engage in transaction-cost augmentation. A detailed analysis of each variable is beyond the scope of this discussion. However, it is important to briefly mention the most relevant of them, including: (a) the complexity of the issue, “since complexity serves as a credible ‘cover’ and excuse for policymakers, as well as a potential source of legislative mistakes”; (b) the enhancement of political job-security; (c) the pay-off to the lawmaker (e.g., campaign contributions, or any other economic advantage such as, for example, the promise of future employment in the private sector, and so on) financed by third parties receiving benefits from transaction-cost augmentation; (d) the “existence of an appealing justification for the measure”; (e) the perceived importance by constituents of the regulatory issue; (f) publicity of the issue. By contrast, transaction-cost augmentation tends to be reduced by: (g) media attention; and by (h) institutional mechanisms aimed at reducing the degree of legislator autonomy (e.g. separation of powers, judicial review, constitutional check and balances, and so on).

5.3.2.2.3. Efficiency Implications
The foregoing discussion yields some important conclusions. First, transaction-cost augmentation theory is consistent with the claim that interest group dominance is one of the main forces driving the political process. More precisely, the augmentation of transaction costs is concurrently a cause and a consequence of group dominance. It is a cause as the increased autonomy of legislators reinforces their ability to distribute divisible private benefits to interest groups in exchange for political support; it is a consequence as interest groups support legislators in their manipulating strategies. Second, the augmentation of transaction costs exacerbates the rent-seeking costs.

---

676 Twight discusses the determinants of government manipulation of constitutional-level transaction costs in Twight, “Government”, supra note 663 at 132-140.
677 Ibid at 135.
678 Ibid at 133
679 Ibid at 138.
680 For a discussion of some possible institutional arrangements that reduce the scope of transaction-cost augmentation see Crew and Twight, “On the Efficiency”, supra note 663 at 28-32.
identified above and the rent-extraction costs that I will illustrate shortly below. As Crew and Twight explain: “the more vulnerable an area of law is to such political manipulation of transaction costs, the greater the magnitude of net social costs that will be tolerated by the voting public, holding voter preferences constant.”

Third, the augmentation of transaction costs is a positive function of the level of “unavoidable” transaction costs: the higher the natural transaction costs, the larger the scope is for augmenting transaction costs.

Fourth, the political process is characterized by greater incentives to augment transaction costs compared to market and adjudication processes. This is a crucial point to be considered when comparing alternative law-making institutions in terms of agency efficiency: political lawmakers retain the possibility of changing constraints to the review of their activity to the detriment of the preferences and interests of a large part of the voting public.

5.3.2.3. Rent Extraction

5.3.2.3.1. Mechanisms

As previously noted, the magnitude of resources invested by interest groups in rent-seeking depends on the magnitude of rents available in the political market. Politicians can thus affect rent-seeking costs by determining the amount of resources available for political allocation. In particular, McChesney has demonstrated that even in the absence of rent-seeking efforts by interest groups, legislators “actively enter the market for rents with their own demands” not just by promising to confer the benefits sought after by competing pressure groups, “but by threatening to impose costs.” McChesney’s model captures an additional source of political process inefficiency

---

682 Ibid. at 27-28.
684 McChesney, “Rent Extraction” supra note 683, at 104 [Emphasis mine].
685 Ibid.
previously overlooked by traditional interest group models. In essence, the political process provides legislators with “rent-extraction” or “rent-extortion” opportunities.

According to the IGT, interest groups are assumed to be willing to invest in rent-seeking up to the presented discounted value of rent available in the political market. Likewise, private actors earning rents from their own activity or investments are assumed to be willing to invest resources to protect their rents from threats of expropriation (by the political authority) up to the value of the rent. The willingness of private parties to invest in rent protection creates the opportunity for legislators to extract private rents through the political process by threatening to impose costs on private actors. For example, the legislator might threaten to “deregulate” an industry sector by removing existing anticompetitive barriers, or to impose a costly new regulation. In such a scenario, private actors will surrender to “political blackmail” if the expected costs of the threatened deregulation (in terms, for example, of lost market shares) exceeds the loss of utility suffered by surrendering (e.g., financing political campaigns, promising private future employment in the private sector, and so on).

5.3.2.3.2. Determinants

The discussion in the last three subsections suggests that politicians have two strategies for maximizing the total return to themselves: political rent creation, and private rent extraction. Political rent creation consists in conferring monopolistic rents on rent-seekers in exchange for political and economic support. Rent extraction consists in threatening to impose costs on private actors (e.g., removing monopolistic rents) in order to obtain political and economic support. The relative advantages of rent creation versus rent extraction depend on the elasticity of the supply and demand with respect to change in consumer prices in the relevant industrial sector. The elasticity of industry supply affects the magnitude of producer surplus and, as a consequence, determines the magnitude of private rents available for political extraction. Thus, the more inelastic the industry supply, the more rent extraction is attractive relative to rent creation. In fact, when industry supply is inelastic, legislators can more easily capture part of the

---

686 Ibid. at 110.
producer’s surplus. By contrast, when industry supply is elastic, rent extraction is less attractive because the size of the producer’s surplus is smaller, and industries are more responsive to political threats. For example, the high specificity of capital stocks invested in industry generates greater incentives to extract rent; by contrast, the presence of more mobile invested capitals reduces the opportunity for rent extraction. The elasticity of the demand side of the relevant industrial sector affects the magnitude of consumer surplus approiprible by producers; that is, it determines the size of private rents that the lawmaker can distribute to pressure groups. Therefore, the more inelastic the demand side of the relevant industrial sector the more rent creation is attractive relative to rent extraction. On the contrary, when the market demand is highly elastic, there is no opportunity for political rent creation because monopolistic price strategies by rent-seekers are not practicable; as a consequence, political rent extraction is a relatively more attractive strategy. In this respect, the ease of market entry reduces the size of the rent that politicians can offer as consideration for political support, which makes political rent creation less attractive than rent extraction. To summarize, as McChesney puts it: “[c]onditions that make political rent creation relatively unattractive to politicians make private rent extraction more attractive.”

5.3.2.3.3. Efficiency Implications

The dreaded possibility of rent extraction by politicians generates serious allocative consequences for the economic system by distorting the ex ante incentives of private actors. First, it reduces the expected return from private investments, thereby reducing firms’ incentives to invest. Second, it generates a substitute effect by inducing firms to replace specific-capital investments, which expose them to the risk of rent-extraction, with more mobile-capital investments that increase supply elasticity by reducing rent extraction attractiveness to the lawmaker. This shift of economic resources is itself a distortive effect because the reallocation of resources is a function of the threatening presence of rent-extorting legislators rather than of the relative market profitability of private investments. In essence, the threat of rent-extraction generates welfare losses in terms of reduced investments in the threatened industry, independent of any rent-seeking

\[ \text{\textsuperscript{687} Ibid. at 117.} \]
effort by interest groups. Third, the threat of rent extortion generates *rent-avoidance* costs.\(^\text{688}\) In fact, those whose actual rents are threatened by the presence of extractive politicians will wastefully expend resources to counter the rent extraction activities of politicians (e.g., through investments in political information and influence). That is, the risk of rent extraction generates rent-avoidance costs inasmuch as the prospect of rent creation creates rent-seeking costs, as noted earlier.

Finally, there is an important feature of the rent-extraction phenomenon that should be emphasized. Its distortive effects are associated with merely being an industry subject to the sphere of influence of political authorities, even in the absence of any *actual* enactment of cost-imposing legislation. That is, the allocative inefficiencies are generated by the very presence of threatening political authorities, independent of the actual extraction of rents by politicians or the actual implementation of extortion practices. This aspect distinguishes significantly the political process from both the adjudication and market processes. First, the threatening potential of the lawmaker is superior in politics than in adjudication. A judge’s pronouncement is generally limited to cases selected and issues raised by private parties in the litigation process.\(^\text{689}\) By contrast, the political process defines for itself the boundaries of its institutional action.\(^\text{690}\) Second, the threatening potential of politics is also superior to that of the competitive markets, which are characterized by an absence of coercive power by centralized authorities.\(^\text{691}\)

### 5.3.3. Conclusion on Agency Efficiency

The analysis in this section has demonstrated that, from an economic perspective, the same factors driving the demand for legislation drive its supply.\(^\text{692}\) Interest groups prefer to invest resources in the demand for private, divisible, benefits rather than for collective, indivisible, benefits, because in this manner they can externalize a large part of the costs

---


\(^{689}\) Here I do not discuss the complicated problems related to strategic interactions between firms belonging to the same industry sector.

\(^{690}\) Constitutional limitations to the legislative power should be taken into account.

\(^{691}\) See Buchanan, “Individual Choice”, *supra* note 450.

\(^{692}\) Stearns and Zywicki, *Public Choice*, *supra* note 500 at 72.
and internalize a sizeable portion of the benefits. For their part, legislators prefer to enact particularized legislation providing private divisible benefits rather than general, public-interest legislation generating diffuse benefits. This is because claiming credit for enacting particularized benefit legislation is a rational and politically rewarding strategy for reelection-seeking legislators, while claiming credit for collective benefits “is not credible, not reliable, and potentially counterproductive.” As a result, the political process displays a fundamental tendency to produce particularized legislation offering private benefits to specific, well-identified interest groups and imposing a spread of collective costs with little or no regard for economic efficiency. Obviously, this does not mean that political law-making never produces general interest legislation that provides the voting majority with collective indivisible benefits. This type of legislation, however, is likely to be a by-product of interest-group politics rather than the result of a political process following the dictates of welfare economics (i.e., correcting market failures and providing public goods). The intensity of this tendency toward inefficient production of private benefits at collective costs varies depending on the characteristics of the regulatory environment, some of which I have identified above.

Political law-making entails a substantial increase in agency costs that is directly related to the tendency toward the public production of particularized benefit legislation. The foregoing discussion has identified four different categories of agency costs.

(i) Natural agency costs are unavoidably associated with the structure of the political process and are independent of the attitude of politicians toward efficiency. These costs are likely to be higher than those associated with adjudication, as they are related to the collective nature of the political process and to the collectivized nature of the political outcome. I have identified above some of the conditions exacerbating agency slack (e.g., issue complexity, issue visibility, appealing justification, and so on).

(ii) Strategically augmented costs are an important source of relative disadvantage in political law-making versus both market and adjudication. These costs are in fact associated with one of the most distinctive features of the political process: unlike private

693 Aranson and Ordeshook, “Political Bases”, supra note 588 at 22.
ordering and adjudication, the political lawmaker can strategically manipulate constraints to the review of his or her own activity.

(iii) Monopolistic-rent costs correspond to the loss of consumer surplus associated with the attribution via legislation of monopolistic positions to favoured interest groups.

(iv) Rent-seeking costs include the resources dissipated during the process of seeking to obtain monopolistic rents plus the deadweight loss imposed on the economy, to the extent that the efforts and resources invested in rent-seeking are diverted from more productive activities.

(v) The costs associated with the threat of rent-extraction by politicians. These include distortions in the allocation of resources by private actors to avoid the costs threatened by politicians plus the rent avoidance costs that are wastefully sustained by private actors to counteract the rent extraction activities of politicians.

5.4. Adaptive Efficiency

This section addresses the matter of whether the structure of the political process possesses some advantages over alternative source of law in facilitating the transition toward more efficient legal regimes. This discussion is related to sub-sections 4.4 (decentralized processes) and 7.4 (adjudication).

5.4.1. The Advantages of Political Law-making

Political law-making exhibits efficiency advantages such as (i) a coordinating device and (ii) a loss-mitigating mechanism. Let me clarify these two points.

5.4.1.1. Centralization as a Coordinating Device

The first significant advantage of the political process is related to its centralized nature. As explained in chapter 4, centralized institutions help us to overcome the dilemma associated with the collective migration toward superior legal regimes by disciplining people’s expectations about the behaviour of others. In many cases, the political lawmaker acts as a focal-point provider by reducing the cost of informing those who must decide whether to migrate or to stick to the status quo. This is certainly a major
advantage of political law-making versus private ordering and decentralized adjudication processes.

5.4.1.1.2. Legislator Offsetting Capabilities

The mitigation of losses is often a precondition for the political practicability of legal change. In this respect, politics has a greater institutional capacity to mitigate the losses associated with policy changes as compared to adjudication. While, as we will see, the dichotomous nature of the adjudication outcome limits the judge’s ability to compensate losers, within the legislative arena opposing groups can engage in bargaining practices (e.g., vote-trading and logrolling). The possibility of trading between intense minorities and less passionate majorities can potentially lead over time to a “tendency towards equalization in the gains and losses from government action” among opposing groups. This advantage should not be overestimated. First, as I have emphasized in section 5.2.5., socially efficient vote-trading is subject to severe structural limitations and restrictive utility requirements. Second, as Trebilcok observes, there are some type of losses that cannot be compensated through logrolling among political representatives. When “losses are both large in relation to the losers’ net worth and substantially larger than the losses that usually result from legislative decisions,” then the offsetting effectiveness of political bargaining is likely to be insufficient. The expropriation of a private residence for public construction is an example. Finally, in many cases the incentive structure of both losers and winners make transition-cost mitigating strategies politically impractical. In general, economic actors (people, firms or other organizations) weight losses substantially more than gains. Due to the “endowment effect,” people demand significantly more to give up an entitlement that they already own than they would be willing to pay to acquire it. This loss aversion is a cause of stickiness of the status quo. Additionally, gainers are likely to be resistant to legal change if they have to bear the costs of the offsetting strategies necessary to facilitate the regulatory change. When faced with the choice between actual compensation costs and uncertain future benefits, they are likely to discount heavily the advantages from legal change due to the time inconsistency problem. Obviously, in many cases the winners will attempt to ensure that the costs of the

694 See Trebilcock, Dealing with Losers, supra note 37.
695 Ibid. at 29.
regulatory change will not translate into high *per capita* costs in order to diminish the incentives for opposing the legal change. But this is not a guarantee of efficient legal change.

**5.4.2. The Disadvantages of Political Law-making**

Despite the two advantages identified above, political law-making is plagued by severe disadvantages in terms of adaptive efficiency. The next three subsections outline three major problems hampering the adaptive efficiency of politics.

**5.4.2.1. Ex Ante Centralization and Outcome Collectivization**

First, legislators do not directly capture the benefits, and they do not bear most of the costs of enacted legislation. This contradicts an essential condition for the promotion of *efficient* legal change, which is generally achieved where actors bear all the real costs and benefits of their decisions. Political law-making is a *centralized* process; as such, it entails a separation between lawmakers and the people subject to law. Furthermore, centralization involves the economic calculation problem: legislators do not have the institutional capacity to estimate relative prices and to measure choice-influencing costs.

Second, the *ex ante* dimension of political law-making entails the problem of incomplete political output. Legislators cannot foresee all the situations requiring a specific regulation. The higher the level of specificity of legal rules, the higher the maladaptation costs likely to be associated with changing conditions in the regulatory environments.

Third, the degree of centralization of the political process, as previously observed, is reflected in the *collectivized* nature of the political outcome. While adjudication outcomes generally produce lower and less visible, immediate, economic consequences, the general, prospective nature of enacted rules imposes a *non*-incremental direction on legal change. Namely, legislative rules entail higher immediate stakes and a consequent larger resistance by losers. In addition, due to the collectivized nature of political outcomes, the costs of inefficient legislation are much more dispersed than the costs imposed by inefficient judicial decisions. As Twight observes: “Costs of inefficient statutory rules are typically more dispersed than the costs of inefficient judicial ruling (with the significant
This implies that in the adjudication context “only a few minds need to be persuaded of the rectitude of one’s case. In vivid contrast, a change in statutory law – particularly federal statutes – requires that a great many minds be swayed, and transaction costs rise accordingly.” The more centralized the process, the higher the number of potential opponents to the collectivized political outcome.

5.4.2.2. The Dynamic Implications of Agency Costs.

The extremely high agency costs associated with political law-making has obvious, dynamic implications. First, group dominance might hamper the transition toward efficiency. Assume that a superior legal regime is available. The opposition of high per capita stakes interest groups whose interests are affected by the legislation might impede the transition to socially efficient legal regimes. This is the obvious consequence of the tendency of higher per capita stake minorities to dominate low per capita stake majorities, even when the aggregate benefit enjoyed by the majority more than compensates for the aggregate costs sustained by the minority. As Olson has demonstrated, modern democratic societies will accumulate an increasing number of special-interest organizations as time goes on. This implies that the number of “veto players” tends to increase over time. In turn, this significantly increases the costs of legal change because the number of actors whose assent is required and whose particularized demand for private benefits needs to be satisfied increases. This effect is exacerbated by the fact that, from the strategic perspective of a single interest group, acting as a veto player is generally a politically more rewarding strategy than advocating for legal change. Consider also that the number of veto players is partially a function of the institutional environment. In this respect, it is important to underline that the adjudication process has a comparative advantage over adjudication in reducing the number of veto players blocking legal change. Trebilcock illustrates the problem of veto players as a factor hampering legal change by examining the case of farmers’

697 Ibid.
698 Olson, Rise and Decline, supra note 623
699 Tsebelis, “Veto Players”, supra note 629
organizations in the agricultural policy sector, and that of labour unions on matters of immigration policy.\textsuperscript{700}

Second, agency costs are likely to influence the direction of legal change; in fact, rent-seeking organized interest groups are likely to promote socially inefficient political outcomes. The preceding discussion on rent-seeking demonstrates this point. The supply of legal change depends critically on the balance in the influence structure among vested interest groups in society. Socially inefficient legal changes may occur if the expected gains to dominant interest groups exceed the gains to society. Furthermore, interest groups determine an “entry-limiting” effect. Namely, interest groups might limit the capacity to influence in the political arena of emerging interests through an instrumental use of “veto gates” and “check points” disseminated along the path of legislative change.

Finally, it might happen that under the influence of powerful interest groups a majority imposes major costs of a policy-legal transition on unorganized minorities. As Trebilcock observes, “this might lead to the adoption of policies that are not socially optimal because they confer modest benefits on the majority at great expense to the minority.”\textsuperscript{701} An example is provided by the taking of private residences to enable the construction of economic-development takings.

To summarize this important point, recourse to the political process entails numerous “brakes” on efficient legal change: (i) an increasing number of veto-players might block efficient transitions; (ii) an increasing number of rent-seeking groups have incentives to promote inefficient transitions and; finally, (iii) outsider minorities unable to block inefficiency might be forced to bear the major costs of innovations. This explains why in regulated environments where the push for change towards superior legal regimes is strongly due to continuous technological innovation, and where high per capita private, economic interests harmed by the socially efficient transition are huge, the political process is rarely the institutional instrument triggering legal change. More often, it

\textsuperscript{700} See Trebilcock, \textit{Dealing with Losers}, supra note 37 respectively, chapter 6 and 7.
\textsuperscript{701} \textit{Ibid}. at 12.
happens that legal change begins in other institutional arenas and is later accompanied by statutory reforms.

Examples corroborating the preceding considerations abound. Kearney and Merrill provide a useful analysis of the regulatory changes that occurred in different, relevant industrial sectors in the US beginning in the 1970s. They analyze the move from an old paradigm of regulation, based on the protection of end-users and the effort to ensure reliability of service, to a new paradigm based on competition among suppliers as a way to enhance consumer welfare. One of the conclusions of their work is that “Congress has generally not made the first move in opening up regulated industries to greater competition.”702 For example, in the regulation of the telecommunication industry, Congress maintained a passive role until 1996, by which time the court and the Federal Communication Commission (“FCC”) had made significant legal innovations in the direction of the new regulatory paradigm.703 This does not mean that the role of Congress can be considered of secondary importance: “there can be no question but that the [Telecommunication] Act sets forth a comprehensive blueprint for a new order far more sweeping than anything that would have been produced by continued litigation and administrative action.”704 Once the wheels of change had been set in motion, the statutorification of legal changes undoubtedly helped the transition toward efficiency, thus increasing the overall efficiency of the system. However, it must be recognized that “the central disputes resolved by the Telecommunication Act were defined by previous fights among contending interests in court and before the FCC[...]; consequently, “it would be difficult to say that Congress has played the role of the initial policymaker’s body with respect to most of the great transformations.”705 “Both in the early 1980s and for substantially more than a decade afterwards, Congress was unable to come to a

702 Kearney and Merrill, “Great Transformation” supra note 613 at 1380 [emphasis mine].
703 Ibid. at 1381 (“The entire opening up of the manufacturing and long-distance telecommunications markets during these years was the result of efforts by the FCC (prodded at times by the courts) and the Department of Justice. The most that can be said of Congress in this regard is that it resisted—or at least did not act upon the Bell System’s pleas that it be given protection from the Department of Justice’s antitrust suit that led to divestiture [...]").
704 Ibid.
705 Ibid. at 381 [Emphasis mine].
consensus concerning whether and what type of significant telecommunications legislation should be passed.”

In conclusion, due to the high agency costs affecting the political process parliaments are rarely the promoters of transitions toward legal efficiency. Kearney and Merril describe the role played by Congress as “reactive, irregular, and often ponderously slow.” This description can be generalized to identify characteristics of legal change by statute. Politically induced legal change is often a by-product of the struggle among rival interest groups, and as such, it is burdened by a high level of agency costs, which slow legal change. On the basis of the analysis conducted in this chapter it seems plausible to conclude that legislatures can be the promoter of legal change only if (i) the balance of power among affected interest groups is in favour of socially efficient legal change, (ii) or the visibility of the issue acts as a catalyst for forces driving legal change.

5.4.2.3. Politician Shortsightedness and The Transitional Gain Trap.

Before concluding the discussion in this chapter, it is worth mentioning two specific problems associated with the political production of law. First, “the political process is biased toward the adoption of short-sighted policies, and against the selection of sound long-range policies, when they involve observable [actual] costs […].” This implies that the political process is likely to lead to an inefficient inter-temporal allocation of resources. Second, once short-term rent-seeking efforts have succeeded and an

---

706 Ibid. at 1382. Merrill and Kearney provide other examples of statutorification of pro-market legal changes previously promoted by courts or agencies, including the enactment of the Natural Gas Wellhead Decontrol Act of 1989 (which has statutorified the legal innovations introduced by the Federal Energy Regulatory Commission) and the Negotiated Rates Act of 1993 (which was passed in response to previous decisions by Courts). See also Clifford Winston, “Economic Deregulation: Days of Reckoning for Microeconomists.” (1993) 31 Journal of Economic Literature, 3, 1263 (“[C]ongressional action was not the sole source of the […] deregulation movement and, in fact, was often the last step in the process.”). 707 James D Gwartney and Richard E. Wagner, “Public Choice And The Conduct Of Representative Government,” in James D. Gwartney and Richard E. Wagner, Public Choice And Constitutional Economics (Vol. 6. Jai Pr, 1988) at 13. 708 See the book by Alan M. Jacobs, Governing For The Long Term: Democracy And The Politics Of Investment (Cambridge: Cambridge University Press, 2011). Jacobs explores this problem in depth. He identifies the following three problems preventing elected governments from investing in long-term social benefits at short-term social cost: (1) “electoral risk,” caused by scarce voter attention to long-term policy issues; (2) “prediction”, associated with the difficulty of appreciating the complexity of long-term policy effects; and (3) “institutional capacity,” rooted in “interest groups’ preferences for distributive gains over inter-temporal bargains.”
inefficient regulation has been enacted, the future cancellation of the conferred monopolistic rents become more and more difficult to implement politically, despite the fact that in the meantime the gains to successors to the original beneficiaries have been eroded by market adjustments. Tullock has called this mechanism the “transitional gain trap.”709 Once rents have been captured, they cannot be easily removed, and society remains stuck in a sub-optimal equilibrium. More recently, Trebilcok has made this point clear:

[…] the longer a program has been in place, the more likely it will have come to form an integral part of beneficiaries’ expectations and the more likely those individuals are to identify with one another and mobilize to defend what they perceive as their due. As individuals are exposed to particular regulatory environments, they are not only more likely to adapt to these situations, but are also more likely to accept them as appropriate. This, in turn, is likely to engender a greater reliance on the policy. Individuals may also form identities based on their shared reliance or experiences, which might also be sufficient to overcome collective action problems.710

The combination of political short-sightedness and pervasive “transitory gain traps” generates dramatic lock-in effects. One example is pro-labour legislation, enacted in some countries under the pressure of labour unions that exceedingly raise the costs or lower the productivity of labour (e.g., raising the cost of firing unproductive workers). From a social efficiency standpoint, the economic advantage obtained by current unionized workers is likely to be “transitory,” in the sense that future workers will not receive significant benefits from it. The reason is that, due to the increased costs and diminishing productivity of labour, employers will reduce their demand for labour. They will invest in labour-saving technologies, they will delocalize their production plants to low labour costs economies, and so on. As a result, new workers will encounter increasing difficulties in entering the labour market. Legislators confronted with a no longer sustainable labour market (and with the impossibility of reforming the existing legislation due to widespread protests and strong oppositions by labour unions) introduce new, more flexible, types of contractual agreements in order to increase incentives for employers to hire new workers. As a result, new workers will spend a good deal of time (even years) working intermittently. Market labour comes to be regulated by two legal

710 Trebilcock, Dealing with Losers, supra note 37 at 27
regimes. One is inefficient because of its excessive rigidity, which discourages new employers. The other is inefficient because of the lack of worker protection. As time goes on, the number of workers under the second regime increases because employers opt for the new, more flexible legislation. The result is increasing job insecurity, lower wages and increasing emigration of the more skilled labour force to other economies. The transitional s obtained by original beneficiaries of rigid pro-labour legislation have been almost completely eroded, but the inefficiencies imposed on the labour market remain.

Conclusions
The unique structural feature that distinguishes democratic politics from other law-making processes is that they give some people the right to make themselves better off at the expense of others. Winners’ decisions are legitimate and binding for all members of the political community. As such, they impose involuntary costs on dissenting groups. This involuntary redistribution effect, which makes some better off and others worse off, is foreign to private markets, which are based on voluntary exchange. It is also alien to adjudication processes, which are based on parties’ voluntary decision to litigate. Unlike market and adjudication processes, in politics some individuals can be forced by other individuals to bear costs for the public provision of goods.

Normative democratic theory tells us that this state of affair is justified by the principle of equality that lies at the basis of majoritarian decision-making. Since people have the same equal right to count in the political process, the legitimate criterion for making social choices is based on the arithmetical strength of alternative options measured by the number of votes expressed. However, positive economic theory suggests that majoritarian decision-making based on equally weighted votes is a highly problematic mechanism for identifying efficient political outcomes.

The collectivized nature of the political outcome and the collective nature of the political process generate a striking number of collective action problems that prevent political law-making from reaching outcome efficiency. First, from a productive efficiency
standpoint, voters face incentives to remain ignorant and to behave irrationally, while politicians confront the economic calculation problem preventing them from identifying efficient legislative rules. Second, from a social choice standpoint, the major problem is that majority rule does not capture the intensity of people’s preferences. The inability to register voters’ cardinal utilities, coupled with the indivisibility of the political outcome, generates large political externalities. Furthermore, given the multidimensionality of many decision issues, it is impossible in many cases to identify a stable majority; the majoritarian equilibrium amounts in most cases to a legal fiction based on equilibrium-inducing institutional mechanisms. The structural-induced equilibrium is unlikely to correspond to the efficient outcome, and it is likely to be manipulated by the agenda setter. Vote-trading can potentially mitigate the insensitiveness to preference intensity, but it is subject to strong structural limitations and restrictive utility requirements. Third, from the agency efficiency perspective, the political process is characterized by an inbuilt institutional bias that structurally advantages the claims of minoritarian interest groups. Furthermore, the agency problem and the social choice problem reinforce each other by favoring the political production of private divisible benefits and the externalization of costs. Finally, from the adaptive efficiency standpoint, the political process is unlikely to promote efficient legal changes, and it is likely to lock in inefficient policies generated by rent-seeking pressures.

The lesson to be drawn from the preceding discussion is that democratic politics is a relatively inefficient mechanism for the production of law. The beneficial effects of political law-making are mostly related to its ex ante centralized structure. They consist in the reduction of information costs in the case of a highly homogeneous and frequent legal demand. In addition, under appropriate conditions, legislation can enhance legal certainty and facilitate legal change promoted by previous litigation or administrative actions. However, these informational advantages should be weighed against the harmful effects related to the mechanism of political representation and to the majoritarian decision-making process. Political law-making results in many cases in an inefficient production of law, it generates political externalities to the detriment of low-per capita majorities, and it entails the risk of blocking legal change, thus preventing society from
moving toward the production possibility frontier. Stripped to its essence, democratic representative politics *seems to work much better as a limit or the abuse of political power rather than as a mechanism for the efficient production of legal rules.*
This chapter examines the structure of costs associated with bureaucratic law-making. The aim is to identify the comparative advantages in production and agency efficiency of bureaucracies and agencies (here used as synonyms) as producers of legal rules. Social choice and adaptive efficiency of bureaucracy are not discussed; these are important issues, but they are not essential for the general purposes of this study, so I leave them to future research.

Politics and bureaucracy are ex-ante centralized systems; as such, they share the relative efficiency advantages and disadvantages identified earlier. While political law-making relies upon generalist, elected representatives to make policy decisions, bureaucratic decision-making is entrusted to skilled experts, without a direct mandate by voters. The reliance of bureaucratic law-making on politically insulated, specialized bureaucrats, instead of on re-election seeking and generalist legislators, enables it to overcome certain inefficiencies typically associated with the electoral-representative mechanism. In addition, dependence on electorally disconnected experts enables a better utilization of some of the informational advantages of ex ante-centralized law-making. At the same time, however, it must be recognized that recourse to technocratic law-making exacerbates some of the production and agency costs associated with centralized law-making. Thus, to fully appreciate the conditions under which bureaucracy can be regarded as an efficient source of law, we should be careful to distinguish between what politics and bureaucracy have in common (ex ante centralization) and their differences.
(generalist political representatives versus insulated bureaucratic experts). The line of inquiry pursued in this chapter is the economic trade-off between the lawmaker’s “bureaucratic insulation” and his or her “electoral connection.”

Introduction: Insulation from the Political Process

Since the first half of the twentieth century, the politicization of law has been paralleled by the growing importance of “bureaucratic” law-making (or “technocratic” law-making; the terms are used as synonyms here). This tendency reflects a growing awareness of the limitations of representative bodies in producing appropriate legal rules to govern increasingly complex contemporary societies. This has led to the idea that technocratic law-making is an efficient and effective way to remedy the shortcomings of political law-making. Since the second half of the twentieth century the administrative production of legal rules has dramatically risen throughout the US and Western Europe. To the present day the delegation of legislative authority to regulatory agencies has come to be regarded as an efficient and effective method for managing the complexity of modern economies.

However, the idea of remedying the inefficiencies of political representative bodies through technocratic law-making processes is based on a confused understanding of the process of law creation. Namely, the idea that bureaucratic law-making has the capacity to rectify the failures of political law-making has not been always supported by informed distinctions between deficiencies attributable to the mechanism of “political representation” and those related to “ex ante centralization.” Technocratic law-making may well counteract some deficiencies of political representation, but it certainly cannot make up for the inefficiencies associated with the ex ante centralized production of law. Ironically, the conflation of different structural dimensions of law creation – together with the emphasis on the virtues of technocratic law-making – have resulted in a self-reinforcing linkage between increasing “statutorification” and the “bureaucratization” of law-making with a consequent growth in the total amount of regulation.711

---
Nothing in these considerations denies the comparative advantages of agencies and bureaucracies versus those of political institutions. As repeatedly emphasized throughout this study, the political process often poorly predicts the overall impact of legal rules on the regulated environment, and it often lacks the cognitive resources required to monitor the costs and benefits of enacted legislation. By comparison, agencies often have superior analytical competence, larger informational resources, and a better understanding of the regulatory issues. In general, as the technical complexity of the regulatory environment increases so does the relative informational advantage of agencies versus political institutions. From this perspective, the delegation of law-making power to agencies constitutes a rational response to the informational inefficiencies of political processes.\footnote{Mathew D. McCubbins, “The Legislative Design of Regulatory Structure” (1985) 29 American Journal of Political Science, 4, 721, at 737 [“Legislative Design”] (“Congressmen and congressional staffs are capable of eliciting information concerning the various policy alternatives. Such information retrieval, however, may have high opportunity costs, taking away time and resources from other electorally oriented activities. On the other hand, the problem of choice, and therewith the costs of eliciting information, can be delegated to an administrative agency. Information uncovered by the agency will be available, albeit indirectly, to congressmen. Inasmuch as Congress can exercise substantial control over the activities of regulatory agencies, congressmen would prefer to delegate the regulatory choices, and therewith a large portion of the information costs, to the agency and sit back in an oversight role [a?]waiting clarification of the issue. This strategy costs congressmen little in terms of policy control and enables them to pass the costs of decision making onto the agency.”)} However, despite these informational advantages, the insulation of experts from the political arena and from market processes carries significant law-making costs. “Unlike legislators, bureaucrats do not face direct electoral constraints, and unlike firms, they do not face external market pressures.”\footnote{Stearns and Zywicki, Public Choice, supra note 500 at 340.} Their insulation from politics entails a separation between information and political control that remains problematic in many respects. In addition, the insulation of bureaucrats from market forces affects their incentive structure and the flow of information available to them.

To summarize, insulation is a useful structural device for (i) exploiting informational advantages of \textit{ex ante} centralization, and (ii) correcting deficiencies of political representative mechanisms. However, under some circumstances, (iii) insulation exacerbates the disadvantages of centralized law-making in terms of production and agency efficiency. This chapter elucidates this intricate issue and identifies conditions for the efficient delegation of law-making authority to agencies and bureaucracies.
6.1. Productive Efficiency

This section examines some important departures of bureaucratic processes from productive efficiency. The discussion is organized around the input-output sequence. The first line of inquiry focuses on the tendency of bureaucratic agencies to expand the volume of supplied service beyond the point where marginal costs equal marginal benefits. Subsequently, analysis is extended to the bureaucratic tendency toward increasing law-making production costs. Finally, the discussion emphasizes the impact of the bureaucrat’s incentive structure on the quality of the bureaucratic outcome (i.e., effectiveness).

6.1.1. The Production Function

Defining the bureaucratic production function poses exceptional methodological problems that cannot be discussed here. The analysis in this subsection assumes that a physical measurement of output and some quality indicators are defined and available. A few preliminary working definitions are useful for the following discussion. First, productive efficiency is attained when the bureau achieves the maximum result (measured in physical output units) at the minimum cost. An alternative perspective is that of X-inefficiency, which occurs when inputs are not used efficiently, including the

---

714 In the interest of brevity, I adopt an encompassing notion of bureaucracy and ignore some important distinctions that would be central in a more comprehensive economic analysis of bureaucracy. In particular, I focus on agency “rulemaking” while ignoring agency “adjudication” activity. Second, I ignore the distinction between the executive administration and independent agencies. Finally, I do not distinguish between different professional levels in the organization of the workforce of bureaucracies and administrations.

715 Many economic models of bureaucracy have been developed in the last few decades. These are based on different sets of assumptions about the bureaucrat’s preferences, incentive, and constraints, which yield different efficiency implications. These models should not be regarded as mutually exclusive, rather they often provide complementary explanations of bureaucratic behavior, and all considered allow for a sufficiently elaborated understanding of bureaucratic efficiency.

716 For example, the absence (or, at least, the strong mitigation) of the “electoral connection” between lawmakers and recipients of law poses the question of whether the output should be still defined (i) in terms of induced compliance with the legal rules by the people subject to law, or, alternatively, (ii) in terms of compliance by bureaucrats to delegation by legislators. In addition, as I will briefly mention in sub-section 6.1.2.3, complex expectations of bureaucratic law-making, coupled with the absence of price signals from the market, poses exceptional problems to defining the output and to monetizing the costs and benefits of regulation.

following circumstances: (i) excessive use of production factors (e.g., excess personnel or excess capital); (ii) excessive price for production factors (e.g., excessive cost of labour); (iii) poor organization or technological backwardness of the administration. Finally, *effectiveness* expresses the relationship between bureaucratic output (i.e., regulation) and the level of utility generated for output recipients (i.e., the net utility of regulation for people subject to law). Effectiveness pertains to the *quality* of output in relation to the needs that production is designed to satisfy. Efficiency is related to effectiveness because – holding all else equal – the more efficient the use of production factors the more effective the output. For example, holding constant the quality of administrative action, maximizing the number of administrative acts issued by the public administration in a given period of time, and minimizing unit costs, results in maximal total utility by the public administration.

6.1.1.1. The Budget-Maximizing Behavior of Bureaucrats

Earlier economic models identified the major cause of bureaucratic inefficiency in an elevated number of activities, that is, in an excessive amount of output. In particular, Niskanen developed a utility maximization model of bureaucratic behaviour that is still considered the benchmark for economic analyses of bureaucracy. Niskanen’s model is based on the following simplifying assumption. (a1). A bureaucrat’s utility function (e.g., power, prestige, career advancement, income level) is a positive monotonic function of *budget* maximization.\(^{718}\) (a2). Bureaucrats possess the *monopoly* power of a perfect price discriminator,\(^{719}\) and enjoy a strong information advantage over their principals (i.e., legislators). This creates some leeway for discretionary decisions, which allows them to pursue their own interests. (a3) Rather than bargaining on a per unit basis like firms in private markets, bureaucrats trade a *total* output for a *total* budget. This means that fiscal purchasers are constrained to buy at all-or-nothing prices – i.e., they cannot reduce the quantity they willing to purchase as the price increases. Both these elements confer monopoly power on bureaucrats, which enables them to extract the total consumer

---


\(^{719}\) Here I refer to the basic model of bureaucracy developed by Niskanen in chapters 1-6. Subsequently in the book, Niskanen relaxes the assumption of bureaucratic monopoly", and discusses the beneficial effect of competitive environments on bureaucratic behavior.
surplus from the public goods they supply. (a4). The output level that maximizes budgetary expansion is produced at minimum costs.\(^{720}\)

The prediction emerging from Niskanen’s model is that bureaucratic output will exceed the Pareto optimal level. Excessive output will occur because bureaucratic agencies are budget-constrained rather than demand-constrained. This means that for bureaucrats the output-maximizing level does not correspond to the point where the supply of public goods intersects with citizen demand for such goods.\(^{721}\) Rather, bureaucratic agencies will produce up to the point where the total production costs equal the total utility to citizens. At the budget-maximizing output level, the marginal production cost exceeds the marginal benefit for citizens, which entails a utility loss equal to the size of consumer surplus obtainable at the social-welfare maximizing point. Namely, by producing up to the point where the budget (rather than social welfare) is maximized, bureaucrats extract the total consumer surplus from the supply of public services. Crucially, Niskanen assumes that citizens (and their representatives) will not react to the oversupply of public goods up to the point where the demand for public goods is globally satisfied without a net loss of social welfare. Only beyond this point, once the total cost begins to exceed the available budget and public service operates at loss (i.e., in deficit), will citizens react against productive inefficiency in bureaucratic behaviour.\(^{722}\)

\(^{720}\) This assumption is unrealistic, because, as I will discuss subsequently, one of the central tendencies of bureaucratic processes is the inefficient utilization of input factors. However, Niskanen implicitly assumes cost-minimization and focuses on the distortive effect generated by the tendency of bureaucratic processes toward budget maximization and over-supply of public goods.

\(^{721}\) That is, the output-maximizing level does not correspond to the social-welfare maximizing point where marginal production costs equal marginal benefits.

\(^{722}\) In the Niskanen model the only constraint on budget maximization is the citizen demand curve. However, citizens (and their representatives) are not expected to know the true costs of the bureau’s services; citizens do not fully realize that bureaucrats are spending resources to produce public goods in excess of the quantity demanded by the public. Therefore, there is a region of production in which output exceeds the socially optimal level (in which supply and demand intersect), yet citizens do not react because their demand for public goods is satisfied, and they are not immediately affected by the excess production. Citizens react only when the total cost of a bureau’s services exceeds the total value to them; that is, they respond only when the bureaucratic process produces financial deficits (i.e., the total costs of production exceed the available budget). At this point citizens react because (since public deficits are reflected in higher taxes) they incur the costs of the inefficient production of public goods. For this reason, bureaucrats are better off by choosing output levels that exceed what citizens are willing to pay under perfect information, up to the point where the total cost of production equals the available budget. For a critique of the assumption concerning consumer reactions to the bureaucratic over-supply of public goods, see Albert Breton and R. Wintrobe, “The Equilibrium Size of a Budget-Maximizing Bureau”
Although Niskanen’s assumptions are overly simplifying, the model captures an important tendency of bureaucratic processes: the oversupply of public goods. Applied to the production of legal rules, Niskanen’s model provides an economic explanation for the constant increase in regulation by bureaucratic agencies in the last few decades. From this latter perspective, expansion of the jurisdictional domain of administrative regulation, which characterizes the evolution of contemporary legal orders, can be explained, at least in part, by the budget-maximization attitude of regulators. Bureaucrats might increase the number of legal rules as an attempt to cover the maximum share of the regulation market. Regulatory agencies have strong incentives to expand the area they cover so as to eliminate potential competitors – i.e., other agencies that could be more efficient and effective in providing regulatory services. For example, the recent approval of the Dodd-Frank Act, which dramatically expands the scope for agency intervention in the financial market, is likely to promote a race among various agencies to regulate first. Such competition might be beneficial in many respects, as previously emphasized, but it might also entail excessive and inconsistent regulation.

The overriding complexity of regulations might be explained, at least in part, as an attempt by bureaucrats to structure them so as to artificially boost the production of their services, and thereby increase the size of their budget. Bureaucrats maximize the complexity and level of detail of legal rules as a form of Niskanen output-maximization. “[B]y writing highly complex, detailed, and specific regulations, bureaucrats can build ‘obsolescence’ into regulations, thereby requiring their ongoing services to update them.” This means that regulators artificially expand legal rules through a strategic provision of complexity and detail. Zywicki provides the example of economically

---

(1975) 83.1 Journal of Political Economy, 195. See also Forte, Principles, supra note 621 at 225-226 (proposing to integrate the Niskanen’s model with Baumol’s constrained managerial model, so that the equilibrium size of the budget is reached at the point where a minimum level of “rent” to consumers is ensured).


725 Stearns and Zywicki, Public Choice, supra note 500 at 365.
inefficient environmental regulatory tools (e.g., command-and-control regulation) that are usually preferred by regulators to more efficient decentralized market-based regulatory methods.\textsuperscript{726}

6.1.1.2. The Appropriation of Fiscal Residuum and Increasing Marginal Costs

Niskanen’s original model assumes that all the dimensions of the bureaucrat’s utility are positively related to budgetary expansion. This assumption hinges on the supposition that bureaucrats derive zero utility from the fiscal residuum. That is, bureaucrats are seen as not appropriating the difference between revenues and costs; consequently, they do not have any incentive to maximize the fiscal residuum. Since they are assumed not to appropriate the fiscal residuum, the only way for bureaucrats to extract wealth from public services is to expand production beyond the point where marginal costs equal marginal benefits. Therefore, they use all expenditures to increase production and to generate the largest possible budget at minimum costs. The Niskanen bureaucrat “pursues the productive efficiency of the competitive firm without any regard for taking rents from the bureau for himself.”\textsuperscript{727} “Paradoxically, the bureaucrat in Niskanen’s model is as production efficient as a firm in a competitive market but takes zero rent for his effort.”\textsuperscript{728}

The zero utility assumption constitutes the major limitation of Niskanen’s model. Indeed, it has been observed that in real-world situations bureaucrats derive utility from the budget discretion that enables them to enlarge the prerequisites of their office and to obtain other indirect benefits from the status of the agency. Thus, it is probably closer to reality to assume that bureaucrats spend less than the total budget on production, and retain a “discretionary margin” added to the minimum (i.e., optimal) cost of producing


\textsuperscript{727} Charles Rowley and Robert Elgin, “Toward a Theory of Bureaucratic Behavior” in D. Greenaway and G. K. Shaw (eds.) Public Choice, Public Finance and Public Policy (Oxford: B. Blackwell, 1985) at 33 [Rowley and Elgin“Toward a Theory”]. [ I will stop editing these and allow you to standardize them according to this model (or another one if I have missed something here). The important thing is that they all follow the same format, that titles are capitalized, and that quotation marks and italics do not occur with the same title.

public goods. Based on this consideration, Migue and Belanger developed a model in which the bureaucrat’s utility function depends on a chosen combination of output and residuum. The residuum is the difference between the tax revenue collected for a public service and the minimum cost of producing it. Bureaucrats appropriate part of the residuum, from which they extract utility in the form of office prerequisites and other benefits. In addition, in the M-B model bureaucrats maintain the power of perfect price discrimination, which enables them to capture the total surplus to the detriment of their principals (i.e., politicians and, ultimately, voters).

The inclusion of both the output and the residuum in the objective function has two important implications. First, while the Niskanen bureaucrat produces at minimum costs, the M-B bureaucrat extracts utility from the public bureau by increasing the marginal cost above the minimum level. In short, bureaucrats produce above minimum costs. Second, holding all other things equal, if the marginal cost curve moves upward, the output level falls compared to the output predicted by the Niskanen model. That is, both the Niskanen and the M-B bureaucrat produce beyond the point of intersection between the demand for and the supply of public goods; however, while the Niskanen bureaucrat maximizes the output given the budget constraint, the M-B bureaucrat produces less output at a greater per unit cost. In essence, the M-B bureaucrat exhibits lower productivity than the Niskanen bureaucrat. While the latter is exchange inefficient, the former is both exchange and production inefficient. Thus the M-P model identifies an additional tendency of bureaucratic behaviour, namely, “that of increasing the production costs as compared to the cost that would arise with a medium-efficient market enterprise.”

To summarize: the Niskanen and the M-B models capture two important characteristics of the bureaucratic production function: (i) a tendency to produce a global activity volume exceeding the Pareto-optimal level, and (ii) a pattern of lowering productivity in response to increased marginal costs resulting from the bureaucratic appropriation of the fiscal residuum.

---

729 That is, beyond the point where the fiscal residuum is maximized.
730 Forte, Principles, supra note 621
6.1.1.3. X-Inefficiency

Previous economic models neither identify nor explain the mechanisms whereby bureaucrats extract wealth from the supply of public goods. A useful line of investigation is to encapsulate bureaucratic input preferences in the model; this enables us to better specify how the discretionary margin is diverted from the supply of public goods to the bureaucrat’s own purposes. Economic theory has focused on: (i) excessive employment of production factors and, in particular, a non-optimal labour/capital ratio; (ii) higher payment to production factors; (iii) lower productivity levels. These inefficiencies are often lumped into the broad category of X-inefficiency.731

6.1.1.3.1. Labour/Capital Ratio

In early studies of bureaucratic behaviour, C.N. Parkinson emphasized that bureaucratic organizations exhibit the tendency over time toward expanding the workforce despite declining production levels. That is, the growth rate of the bureaucratic workforce tends to be positively constant over time, independent of the fact that the bureau may operate in an “uneconomic” region where the marginal productivity of labour is negative.732 Subsequent studies have confirmed that bureaucrats exhibit a preference for labour intensive production processes.733

Based on the preceding considerations Orzechowski has developed a model of the bureaucratic production process that incorporates the bureaucrat’s input preferences.734 He maintains both Niskanen’s assumption of monopoly power and M-B’s assumption of

---

731 Ibid. at 227 and 399 distinguishes between X-inefficiency in the strict sense (including excess of personal and other production factors) and X-inefficiency in the broad sense (including excessive prices for production factors and poor organization and technological backwardness of the public administration).

732 This aspect of Parkinson’s theory is discussed in Albert Breton and Ronald Wintrobe, “An Economic Analysis of Bureaucratic Efficiency” (1983) University of Toronto, Law and Economics Workshop Series, at 5.


the bureaucrat’s preference for above-minimum-cost production. In addition, he assumes that both output levels and quantities of labour are direct arguments of the bureaucrat’s utility function, and that the bureaucratic choice of output level is subject to the constraint of total input costs being equal to the available budget.\textsuperscript{735} This means that labour is at the same time an argument in the utility function and a factor of the constraining equation. The fact that labour is simultaneously an input and a preference variable\textsuperscript{736} distorts the optimal usage of inputs. To achieve an optimal input combination in a profit-maximizing firm the ratio of the marginal product of labour and capital must equal the ratio of their prices. However, since bureaucrats derive utility from labour distinct from its contribution to production, they will display a tendency toward inefficient combinations of labour and capital resulting in an overabundance of staff. This implies that as compared to a profit-maximizing firm “a bureau supplying the same service faced by the same demand condition and subject to the same input-price ratio will use a greater labour-capital ratio.”\textsuperscript{737} In essence, bureaucrats tend to choose an amount of labour beyond the point where the marginal product of labour equals its price.

6.1.1.3.2. Laziness

Peacock introduces into the model of bureaucracy “two widely recognised features of bureaucratic leadership – caution and laziness.”\textsuperscript{738} He challenges the assumption that labour is always positively related to the bureaucrat’s utility function, and contends that an increase in the number of administrative officials has a twofold effect on the senior bureaucrat’s utility function. On the one hand, it positively affects utility as it entails (i) greater opportunities to delegate duties, (ii) greater power and prestige, and (iii) greater political influence of the bureau. The positive impact of labour on the bureaucratic utility function is captured by the assumption that this function depends upon three independent variables $U (N, L, S)$. That is, bureaucrats derive their utility from $N$ denoting the number of officials under his command, $L$ denoting “on-the-job-leisure” (i.e., caution and

\textsuperscript{735} That is, the utility of senior-bureaucrats is a positive function of both output and labour, and is subject to a budget’s constraint given by the sum of the total cost of labour and the total cost of capital.

\textsuperscript{736} \textit{Ibid.} at 244

\textsuperscript{737} \textit{Ibid.} at 245.

laziness), and $S$ denoting the “surplus,” intended as the margin “over and above the wages and other emolument payable to administrative grade staff.” On the other hand, larger staffs increase the bureaucrat’s responsibilities, thereby diminishing bureaucratic utility.

Peacock incorporates these considerations into his model through the construction of two budget constraints. The first constraint is given by the sum of the total labour cost (with exogenous salary rate) and the surplus. A second constraint is represented by the leisure function $L(N)$, with $L'(N) < 0$, which summarizes the negative relationship between the on-the-job-leisure and the number of staff members. If the leisure function is substituted for $L$, the utility function can be expressed as $U(N, L(N), S)$, which means that $N$ affects $U$ as both a direct argument in the utility function and as an argument in the on-the-job leisure function. In other words, an increase in staff members produces two opposing effects on utility: (i) it increases utility, because the number of officials is positively related to utility; (ii) it decreases utility, because the number of officials is a negative argument in the on-the-job leisure function that is an argument of $U$.

Peacock observes that the above described twofold effect “places an upper limit on the expansion of output which […] makes it possible for the output level to be below the output level produced by the profit maximizing firm producing under conditions of pure competition.” This implication distinguishes the Peacock bureaucrat from both the Orzechowski and the Niskanen bureaucrat, who instead tend (to different extents) to over-supply public goods. The reason for this different predictive result is easily understood. In the Niskanen model the only upper limit on output is the budget constraint; likewise, in the Orzechowski model the upper limit on output is provided by the size of the budget given the cost of labour. By comparison, in Peacock’s model the output level chosen by bureaucrats is limited not only (i) by the size of budget, but also (ii) by the constraint on the bureaucrat’s demand for labour brought about by the loss of leisure associated with increasing staff. The Peacock model produces less output because

---

739 Ibid. at 128.
740 Ibid. at 130-131 [emphasis mine].
the bureaucrat’s preference for labour is less intense than the Orzechowski model assumes due to the negative effect of increasing staff on the bureaucratic utility function.

6.1.1.3.3. Risk Aversion and X-Inefficiency

The preceding economic models illustrate the impact of alternative utility functions on the production and cost functions underlying the bureaucratic process. They identify a twofold bureaucratic tendency toward over-expanding the quantity of public services, and increasing production costs as compared to what would be the case in profit-maximizing enterprises. For the purposes of process efficiency analysis, these models suffer from a significant limitation: they fail to identify the institutional mechanisms affecting the bureaucrat’s incentive structure. In particular, they do not explain (i) why bureaucrats exhibit a preference for an excess of inputs, and (ii) whether the bureaucratic environment plays any role in shaping a bureaucrat’s preferences. I will provide some useful insights into these issues subsequently when examining agency efficiency. In this subsection, I discuss one of the most critical elements to be taken into account when comparing bureaucracy with other institutional processes: the bureaucrat’s attitude toward risk.

Forte and Di Pierro have developed a model of what they call the “Weberian” bureaucrat’s behaviour, which they compare and contrast with “entrepreneurial” behaviour. The authors assume that bureaucrats and entrepreneurs differ with respect to their attitude toward risk, which is affected by the institutional environment in which each operates. The Weberian bureaucrat is expected to conform to rules prescribing ext- ante, well-specified patterns of action. “[I]n his professional life game any deviance from the rule is a ‘negative outcome’ while ‘full observance of the rules’ is the only

---

741 I am referring to X-inefficiency in the broad sense, see supra note 731
743 The activities of bureaucrats are governed by general, abstract and clearly defined rules; the bureaucratic structure exerts pressure “infusing group participants with appropriate attitudes and sentiments” (at…) to obtain conformity with prescribed pattern of action. Forte and Di Pierro observe that “these sentiments are often more intense than technically necessary” (ibid. at 92). In particular, the bureaucratic environment induces the tendency to shift the focus of the bureaucrat’s behaviour “from the aims of the organization onto the particular details of behavior required by the rules. Adherence to the rules, originally conceived as means, becomes transformed into an end-in-itself” (ibid.).
positive outcome.” Performing according to the rules is seen as an automatic and necessary behaviour, while failure to observe the rules results in disciplinary sanctions, and cannot be justified as the consequence of efficient high-risk behaviour. In short, the bureaucrat’s attitude toward risk is determined by the fact that in bureaucratic environments while compliance with rules is not rewarded, deviance is always penalized. “For these reasons the Weberian bureaucrat must develop an attitude to minimize the risk of deviation from the prescribed behaviour which leads him to employ an excess of inputs and to be reluctant to introduce innovations.”

In addition, the bureaucrat’s institutionally induced risk-aversion generates systematic outcome ineffectiveness. The Weberian bureaucrat pursues compliance “independently from the actual meaning of the results achieved.” This is in stark contrast with market environments in which risk is positively related to profits, so that the profit-maximizing entrepreneur has positive incentives to accept a degree of risk related to the expected level of profit. This profit-maximizing perspective induces “a tendency to optimize the relative input-output structure.” In sum, while the entrepreneur has a positive attitude toward risk and tends to optimize the input/output ratio, the Weberian bureaucrat has a negative attitude toward risk and tends to adopt a higher input/output ratio. For every given level of output, the bureaucrat accepts less risk and uses more input than the entrepreneur.

It is worth noting that the preceding prediction is largely consistent with the M-D, the Orzechowski and the Peacock models. However, the underlying explanatory analysis differs significantly. For Forte and Di Pierro bureaucratic X-inefficiency does not derive from a preference for “discretionary margin,” “excess use of labour,” or “on-the-job-leisure”; rather, it is caused by (i) the bureaucrat’s negative attitude toward risk, and (ii) the bureaucrat’s failure to take into account any comparison between the costs and the

---

744 Ibid. at 93
745 Ibid.
746 Ibid.
747 Ibid. at 97

Figure 22 provides a table summarizing the discussion of the bureaucratic production function.
Figure 22 Alternative Models of Bureaucratic Behaviour

<table>
<thead>
<tr>
<th></th>
<th>Utility Function</th>
<th>Constraint</th>
<th>Predicted Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Niskanen</strong></td>
<td>$U = f(Q)$</td>
<td>$B = aQ - bQ^2$</td>
<td>(1) Cost-Minimization (2) Budget–Maximizing Output Level</td>
</tr>
<tr>
<td><strong>M-B</strong></td>
<td>$U = f(Q, D)$</td>
<td>$B = aQ - bQ^2 - cQ - dQ^2$</td>
<td>(1) Costs Above Minimum Level (2) Output Above Optimal Level</td>
</tr>
<tr>
<td><strong>Orzechowski</strong></td>
<td>$U = f(Q, L)$</td>
<td>$R{f(K, L)} = wL + rK$</td>
<td>(1) Cost Above Minimum Level (2) Labour Above Optimal Level (3) Output Above Optimal Level</td>
</tr>
<tr>
<td><strong>Peacock</strong></td>
<td>$U = f(N, L, S)$</td>
<td>(1) $B = \bar{w}N + S$ (2) $L = L(N)$ with $L' &lt; 0$ and $L'' &lt; 0$</td>
<td>(1) Cost Above Minimum (3) Output Below Optimal Level</td>
</tr>
<tr>
<td><strong>Del Pierro - Forte</strong></td>
<td>Risk-Aversion</td>
<td>Institutional Constraints (inducing risk-minimizing attitude and tendency toward over-usage of inputs)</td>
<td>X-inefficiency</td>
</tr>
</tbody>
</table>
6.1.2. Output Ineffectiveness

It has been widely recognized that systemic output ineffectiveness plagues bureaucratic decision-making. With specific respect to the bureaucratic production of legal rules, I identify three main causes of output ineffectiveness: (i) risk-aversion (as already noted), (ii) lack of rigorous cost-benefit analysis, and (iii) biased decision-making.

6.1.2.1. Risk-Aversion and Output Effectiveness

Output effectiveness requires that a bureaucrat’s level of risk-aversion be aligned with the degree of social risk-aversion; however, because of the incentive-structure analyzed above, a bureaucrat’s risk-aversion is often higher than that required by strict cost-benefit analysis. Thus, in real-world situations law-making through bureaucrats allocates to people subject to law either higher or lower risks than the socially optimal level. According to process efficiency analysis, we need to identify the factual circumstances under which a bureaucrat’s risk-aversion is more or less likely to affect output ineffectiveness. A useful strategy for addressing this issue is to look at the detectability of a bureaucrat’s risk-aversion.

It should be recognized that deviance from risk neutrality is more easily detectable when the regulation provided by the bureaucrat entails a higher degree of risk than is socially acceptable. Conversely, when the risk occasioned by the enacted regulation is lower than that which is considered socially acceptable, the bureaucrat’s departure from risk neutrality is less easily detectable. The reason for this detectability differential is that when “higher” risk is allocated to people subject to law it often manifests in actual damages to society; by contrast, “lower” risk can be measured only through counterfactual assessments based on hypothetical scenarios. In short, “excessive” risk-taking in regulation is more easily detectable than “insufficient” riskiness. This has an important implication for process efficiency analysis: bureaucrats producing regulation in areas characterized by outcome riskiness and uncertainty will tend to estimate a greater
probability of imposing “excess risk” than of allocating “insufficient risk,” even if the expected economic costs of the former do not exceed those of the latter. 748

The effect of the bureaucrat’s risk-aversion on the level of regulation can be usefully explained through the concepts of Type I and Type II error. 749 In statistics, a Type I error is the incorrect rejection of a true null hypothesis, while a Type II error is the failure to reject a false null hypothesis. A null hypothesis is the statement that the observed phenomenon does not produce a given effect. In the regulatory context, an example of a null hypothesis is the statement “This activity (e.g., the commercialization of a new drug or food, or substance) has no effect on peoples’ health.” Thus, Type I errors (rejections of a true null hypothesis) occur when bureaucrats reach the conclusion that a safe activity is risky – i.e., a Type I error is a “false positive” with respect to the true null hypothesis that the activity is safe according to the standard of socially accepted risk. Conversely, Type II errors (failures to reject a false null hypothesis) occur when bureaucrats reach the conclusion that a risky activity is safe – i.e., a Type II error is a “false negative” with respect to the false null hypothesis that the activity is safe according to the standard of socially accepted risk.

From the perspective of individual bureaucrats, Type II errors engender more tangible harm than Type I errors. In fact, classifying a risky activity as safe (e.g., classifying a new dangerous drug or food as safe and approving its commercialization) entails the risk of tangible, easily detectable harms to people (e.g., poisoning, intoxication). This “threatens public condemnation, regulatory oversight, and potentially major negative career consequences.” 750 By contrast, classifying a safe activity as risky (e.g., classifying a new safe drug or food as risky and delaying its commercialization) “while also producing potentially significant social costs, might be harder to identify and thus be less

748 To be clear, two risks should be distinguished. One is the risk – confronted by regulators – of being punished for having enacted a regulation that resulted in utility losses for society; the other is the risk of harming people subject to law by the enacted regulation. The latter is the “social risk” associated with regulation. As regards social risk, departures from the social optimum level can be in the form of either “excess risk” or “insufficient risk.” Regulators calculate the risk of imposing an excessive level of social risk on the people subject to law as compared to the risk of allocating an insufficient level of social risk.
749 On this point see Stearns and Zywicki, Public Choice, supra note 500 at 358-361
750 Ibid. at 359.
likely to invite criticism.” In brief, in many cases, Type II errors involve more easily detectable societal costs, because they manifest in tangible harm to people; on the contrary, Type I errors engender social costs detectable only through a counterfactual assessment of foregone benefits that could have been derived from the safe activity that was not pursued.

In conclusion, the incentive structure for bureaucrats leads them to produce a regulation that is “systematically biased in favor of avoiding the more tangible harm associated with Type II error than the abstract and often unobservable harm from Type I error,” despite the fact the former costs might be equal to or even greater than the latter ones.

6.1.2.2. The Absence of Price Signals

Cost-benefit analysis requires that an agency consider all the costs and benefits of a particular regulation as compared to possible alternatives. However, the structural features of the bureaucratic process are often at odds with the possibility of conducting a sound cost-benefit analysis. Besides the problem of risk-aversion previously discussed, two further factors impair a full appreciation of the marginal costs and benefits of regulation. First, the “insulation” of bureaucrats from electoral pressure exacerbates the consequences of “cost-externalization” associated with the ex ante centralized dimension of bureaucratic law-making. As we have repeatedly emphasized, ex ante centralized law-making entails a separation between the lawmakers and the recipients of law – i.e., a separation between the entity creating the regulation and the individuals facing the marginal costs and benefits of regulation. However, while in the political process the problem of cost externalization is in part mitigated by the electoral connection between politicians and voters, the insulation of agencies deprives bureaucratic law-making of the feedback signals concerning the relative benefits and costs of regulation from the people subject to law - i.e., the costs of regulation are borne almost exclusively by private actors operating in the regulated environment, rather than also by rule-maker bureaucrats. The solution to this problem, as discussed earlier, depends in large part on the legislative design of appropriate controls of bureaucratic behavior and delegation strategies.

751 Ibid.
752 Ibid.
However, although important, these latter institutional arrangements remain highly imperfect and are plagued by numerous agency problems; as a result, bureaucratic law-making remains irrepressibly affected by an imperfect understanding on the part of bureaucrats of the marginal benefits and costs of regulation.\textsuperscript{753}

Second, even assuming away all imperfections in the price-signaling mechanism, the problem still remains of \textit{measuring} the benefits and costs of regulations. Stated in the language of process efficiency analysis, \textit{recourse to the bureaucratic process exacerbates the problem of ‘defining’ and ‘measuring’ the lawmaker’s production function}. On the one hand, we rarely have a clear and shared definition of “output” and attribution of “costs.”\textsuperscript{754} “The expectations of bureaucracies, both by employees and the clients they serve, are too complex to put in simple input-output ratios.”\textsuperscript{755} Thus, one major obstacle to measuring the costs and benefits of regulation is identifying efficiency criteria that encompass effectiveness and the overall impact of regulations on the cost and utility functions of the people subject to law. On the other hand, once the efficiency criteria have been identified, the ideal method of assessing the economic advantages and disadvantages of regulation is to conduct a cost-benefit analysis using “shadow prices” to estimate the public goods.\textsuperscript{756} Cost-benefit analysis tries to mimic a basic function of

\textsuperscript{753} As emphasized in chapter 2, agency problems are closely related to productive efficiency problems.

\textsuperscript{754} On the problem of defining the production function of bureaucracy, see Frank P. Scioli Jr, “Problems of Improving the Efficiency of Public Administration” in \textit{Anatomy}, supra note 738, 165.

\textsuperscript{755} \textit{Ibid.} at 166.

markets by seeking to replicate for the public sector a calculation that ordinarily happens in the private sector. However, exceptionally difficult problems are associated with monetizing the costs and benefits of regulation. Despite the fact that over time the ability of regulators to produce scientifically robust cost-benefit analyses has increased with technological advances, recourse to proxies for marginal costs and benefits of regulations remains riddled with conceptual, ethical, and computational problems and often leads to inconsistencies in decision-making outcomes. Finally, the regulated environment in which agencies operate is often fraught with pervasive scientific uncertainty and value judgements. As previously explained, the higher the complexity the more legislators are willing to delegate discretion to technocratic decision-making. However, this raises the problem that under conditions of great uncertainty and complexity the scope for value judgement (e.g., the degree of discretion in assessing the social acceptability of an unknown risk) increases, and this is at odds with the technocratic nature of the decision-maker.

A comprehensive discussion of the relationship between cost-benefit analysis and agency decision-making is beyond the scope of this study. What is important to emphasize here is that in calling for cost-benefit analysis as a tool for imposing consistency and accountability on the bureaucratic decision-making process we need to recognize the

757 Frank Ackerman and Lisa Heinzerling, “Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection” (2001) 150 U. Pa. L. Rev., 1553 (Ackerman and Heinzerling, “Pricing the Priceless”) at 1553 “[cost benefit analysis] demands that the advantages and disadvantages of a regulatory policy be reduced, as far as possible, to numbers, and then further reduced to dollars and cents”).


760 See the bibliographic references supra in note 756 and infra in note 761.

incentive structure faced by the bureaucrats who conduct cost-benefit analyses as well as all the other institutional features of bureaucracy discussed above (e.g., insulation, institutionally induced risk-aversion, and so on). If issues of institutional design are not carefully considered, what appears to be a solution (cost-benefit analysis) might turn out in some cases to be the source of many efficiency problems.\(^{762}\)

### 6.1.2.3. Decision-Making Biases

Despite its insulationist structure, the bureaucratic process is not immune from external pressures and influences. According to U.S. Supreme Court Justice Stephen Breyer, the inefficiencies of U.S. risk regulation are the result of a “vicious circle” involving three factors: (i) the public’s risk perception, (ii) congressional action and reaction, and (iii) the uncertainties inherent in the technical regulatory process. The dynamic interplay among these three elements can be summarized as follows. The public’s perception of risk – which is often seriously flawed with the systemic overestimation of some risks and underestimation of others – directly influences the demand for regulation. The general public exerts pressure on Congress on the basis of \emph{biased perceived} risk. Congress in turn pressures agencies to engage in risk-regulation. Within the regulatory process, uncertainty \textit{invariably} arises concerning many aspects of risk assessment (e.g., uncertainty about the probability of risk, its magnitude, and the costs of controlling its consequences). When public opinion is influenced by the uncertainties emerging from the regulatory process (through sensationalistic media), it becomes more anxious and increases pressure on Congress. The circle is “vicious” because the bias in the public’s perception of risk is magnified by the uncertainties of the regulatory process, and the increased pressure on agencies is likely to be reflected in biased decision-making.

Breyer suggests that regulatory inefficiency – generated by the vicious circle of public perception, politics and the regulatory process – manifests in three serious biases. First, agencies exhibit a tendency to focus narrowly on their regulatory mission at the expense of other policy goals (i.e., “\textit{tunnel vision}”). Tunnel vision is harmful because it discourages bureaucrats from engaging in robust cost-effectiveness analysis of regulation.

A single-minded pursuit of their regulatory task discourages bureaucrats from assessing relative costs and benefits of alternative programs for achieving the greatest result per unit cost. Agencies that develop tunnel vision tend to overspend resources in their attempts to achieve their regulatory goals; the result is excessive regulation, excessive costs and output ineffectiveness.

Vogel, in his study concerning some critical aspects of the U.S. chemical regulatory policy, provides a useful example of the pernicious effect of tunnel vision by regulatory agencies. In 1996 Congress mandated the U.S. Environmental Protection Agency (EPA) to develop a program to test certain chemicals for endocrine-disrupting effects and then to regulate them. The enabling legislation was contained in the Food Quality Protection Act of 1996 (FQPA), which demanded the agency protect public health from endocrine disruptor chemical pollutants (EDCs). In particular, Congress mandated the agency to “develop a screening program, using appropriate validated test systems and other scientifically relevant information, to determine whether certain substances may have [...] endocrine effects.” To address the legislative mandate the EPA developed the Endocrine Disruptor Screening Program (EDSP). For technical reasons (well documented in Vogel’s article) the complexity of endocrine disruption impedes the scientific testing and regulation of EDCs under the EDSP. Because of difficulties associated with this complexity in the regulatory environment, the EPA faced many challenges in realizing the EDSP and it repeatedly missed the implementation deadline. The EPA encouraged delays and slow-paced research to accomplish the instrumental goal of developing and validating screening and testing procedures for identifying EDCs. According to Vogel, one of the principal reasons for the failed program implementation lies in the inadequacy of the scientific testing and regulation paradigm. “This paradigm fails to protect public health, to account for modern social values, and to deal with the complexities and uncertainties of endocrine disruption science.” However, despite the existence of

764 (FQPA, P.L.104-170: Title IV sec. 408p)
765 The EPA organized the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) to provide recommendations on how to develop screening and testing procedures for EDCs.
766 Vogel, “Tunnel Vision,” supra note 763 at 287
viable alternative regulatory strategies that could help reduce some of the health consequences of chemical exposure or improve the current state of knowledge, the agency has remained stuck in an uncritical acceptance of the conventional testing and regulation paradigm without considering the potential efficacy of alternatives that depart from it. As a result this single-minded pursuit of the agency regulatory task has “undermine[d] the end goal to protect public health by distracting attention and other resources from non-testing regulatory alternatives (e.g., mandated pesticide use reduction), and non-regulatory alternatives (e.g., subsidies for organic farming) even in the face of an ineffective regulatory regime.”767 This example illustrates how tunnel vision can lead agencies to overlook alternative regulatory strategies, thereby leading to inefficient regulation.

The second administrative bias constituting the vicious circle is the agency tendency toward bias in the selection of issues for their regulatory agenda. Breyer calls this regulatory bias “random agenda selection”: agencies do not prioritize the most significant problems; rather their regulatory priorities are influenced by public misperceptions of risks. Biased agenda selection generates inefficiency because in many cases it results in the production of bureaucratic outputs that are not in line with recipient preferences. Breyer illustrates the example of an EPA report, issued in 1987, in which experts of the Agency provided their priority rankings. Risk managers ranked hazardous waste sites as generating a “medium-to-low” risk to the public and indoor air pollution as posing a “high” risk. Yet public perception has driven policy to focus on hazardous waste sites (the lower risk) rather than on indoor air quality (the higher risk). Another example is provided by the differences in regulatory agenda of both the FTC and the United States Department of Agriculture (USDA) within the agricultural policy-making process.768 While the FTC focuses its efforts on improving consumer economic welfare – by favouring increased competition to the benefit of consumers – the USDA is mostly concerned with favouring the economic interests of farmers, which represents their primary constituency, even if this might result in decreased competition and higher prices.

---

767 Ibid. at 297
768 Stearns and Zywicki, Public Choice, supra note 500 at 364.
of agricultural products. The agenda selection bias is in line with the prediction of the interest group theory, according to which regulators often pursue the interests of the most influential interest groups.

The last example introduces the third problem identified by Breyer: “inconsistency” among different regulatory agencies in the assessment of risk and the implementation of regulations. Several agencies may possess overlapping jurisdictions for the same regulatory issue, which creates the risk of inconsistent regulation when bureaucratic objectives and preferences are not aligned. As Stearn and Zywicki observe, “each agency tends to emphasize its particular regulatory mission and constituencies, even if those objectives conflict with other agencies.”769 The efficiency implications of outcome inconsistency are significant. Inconsistent regulation results in the misallocation of resources both within and across different regulatory programs (e.g., health and safety). It also may lead the public to ignore significant risks.

6.2. Agency Efficiency

The preceding models assume that legislatures play a passive role: “bureaus are strategic actors driven to achieve their own ends, while the legislature is passive – it sits by idly as the treasury is looted.”770 A more realistic view of bureaucratic power must recognize that “the legislature has authority over the bureau,”771 and that “[t]heir relationship is not simply one of bilateral monopoly. It is an authority relationship in which the legislature has the legal right to tell the bureau what to do. The legislature is the principal, the bureau the agent.”772 In this section I apply this methodological perspective to the examination of costs associated with the creation of agencies and the delegation to them of law-making

769 Ibid.
771 Moe, “Positive Theory,” supra note 770 at 460
772 Ibid.
powers as an alternative to passing laws through normal legislative procedures and allowing courts to oversee their enforcement.

It must be emphasized at the outset that legislators act simultaneously as voter agents and bureaucrat principals. Hence, the inquiry into agency costs can be subdivided in two parts. First, it is assumed that legislators act as faithful agents of their constituencies – i.e., legislator preferences reflect voter interests and goals. This enables the identification of agency costs associated with the principal/agent relationship between politicians and bureaucrats. Subsequently, this assumption is relaxed and analysis focuses on legislative delegation to agencies as a means legislators strategically employ to maximize their utility. This latter analytical perspective facilitates the capture of agency costs associated with the principal/agent relationship between politicians and voters. Hence, the discussion in this section is organized as follows. It first examines bureaucrat shirking of legislative delegation (and assumes the legislator is a faithful voter agent) then it focuses on legislator shirking of the electorate mandate through delegation (and assumes the bureaucrat is a passive actor).

6.2.1. Legislators as Faithful Agents

Assume that legislators behave as faithful agents of voters, and, as such, they decide to create agencies and to delegate law-making authority to them based on the voter utility maximization calculus. Agency theory tells us that the costliness of monitoring and controlling the agent entails substantial slippage: often the costs and benefits of agent choices are suboptimal if considered from the principal’s perspective. This elementary insight can be applied to the legislator-bureaucrat relationship to analyze the “delegation dilemma” confronted by politicians. On the one hand, delegation enables politicians to capture the benefits of specialization and division of labour among the sources of law; agencies can gather the necessary information at lower costs and thus develop superior analytical capacities. On the other hand, legislative delegation entails significant agency costs that result from three critical factors:

(i) *asymmetry* of information between legislators and agencies (bureaucrats act on the basis of information that is only imperfectly available to the principals);
(ii) *conflicts* of interest between the bureaucracy and the legislators (bureaucrats are driven by their own interests, including maximizing their budget, increasing their jurisdictional domain, and so on);

(iii) *imperfect* monitoring of bureaucrat behaviour (politicians incur prohibitive costs for monitoring, rewarding and sanctioning bureaucrats).

Stripped to its essence, the delegation dilemma confronted by legislators amounts to the problem of reaping the benefits of technocratic decision-making without abdicating legislative *control* over the law-making process. *Efficient* delegation maximizes the efficiency gains from agency informational superiority and minimizes the losses arising from the bureaucrat’s opportunistic behaviour. As Epstein and O’Halloran put it: “when delegating to executive branch actors, legislators trade off *informational* gain against *distributive* losses.”  

The following pages examine the structure of the delegation dilemma and the institutional mechanisms available to politicians to maximize the expected net benefits of delegation.

For expository convenience, it is useful to adopt a simple formalization of the delegation dilemma. Let $E$ denote the net benefit for legislators of delegating legislative authority to agencies. Assume that $E$ is a function of $d$, which denotes the degree of *discretion* given to agencies by legislators.
$B$ represents the gross benefits of delegation including, for example, technical expertise, superior analytical capabilities of agencies, and all the other benefits previously identified. It is reasonable to assume that $B$ is an increasing function of $d$: the broader the scope of the legislative delegation the more the lawmaking process benefits from larger informational resources at lower cost – i.e. $\frac{dB}{dd} > 0$. I assume that $B$ is concave (i.e. $\frac{d^2B}{dd^2} < 0$), meaning that the marginal benefit accruing to politicians from delegation is decreasing. $S_B$ denotes the agency slack (i.e., outcome inefficiency) generated by bureaucratic opportunistic behaviour. I assume that $S_B$ is an increasing function of $d$: the broader the scope of the legislative delegation the more politicians lose control over agency rulemaking, thereby increasing the risk and the cost of opportunistic behaviour.

Not only: it seems reasonable to assume that $S_B$ is convex, i.e. $\frac{d^2S_B}{dd} > 0$. Indeed, a higher degree of delegation to agencies makes slackness costs rise more rapidly. Finally, $C$ denotes the costs incurred by legislators to control the behaviour of agencies. The expression (1) means that the efficient delegation wishes to maximize the difference between the benefits from delegation and the sum of the outcome agency costs and controlling costs. Finally, $C$ denotes the costs incurred by legislators to control the behaviour of agencies.

Once they have delegated law-making authority, legislators control agency discretion thorough either ex ante controls or ex post oversight. Let $C_{ante}$ denote the costs of ex ante control, and $C_{post}$ denote the costs of ex post oversight, so that:

$$C(d) = C_{ante}(d) + C_{post}(d)$$ (2)

It is reasonable to assume that both $C_{ante}$ and $C_{post}$ are functions of the degree of bureaucratic discretion delegated to agencies. As I will explain subsequently, ex ante and ex post controls have different cost structures. While ex ante control costs are a
decreasing function of the degree of bureaucratic discretion, the *ex post* control costs are an increasing function of it. That is:

\[
\frac{dC_{\text{ante}}}{dd} < 0 \quad \text{and} \quad \frac{dC_{\text{post}}}{dd} > 0
\]  (3)

The two controlling strategies can be seen as substitutes; circumscribing or channelling *ex ante* agency discretion, i.e. increasing *ex ante* control, reduces the need for costly *ex post* oversight. We add the further assumption that \( C_{\text{post}} \) is convex (i.e., \( \frac{d^2C_{\text{post}}}{dd^2} > 0 \)). This is quite customary as regards costs in general: the marginal cost increases when the level of activity (in this case delegation) increases. Finally, we also assume that \( C_{\text{ante}} \) is convex (i.e., \( \frac{d^2C_{\text{ante}}}{dd^2} > 0 \)): the marginal cost of *ex ante* control decreases at a decreasing rate as the degree of bureaucratic discretion increases.

By substituting (2) in (1) we obtain:

\[
\max_{d} E(d) = B(d) - S_B(d) - C_{\text{ante}}(d) - C_{\text{post}}(d)
\]  (4)

which summarizes the cost and benefit functions considered by politicians when choosing the preferred level of discretion to confer on agencies, including the information benefits, bureaucratic slack, and the relative costs of *ex ante* versus *ex post* control.

The politicians’ preferred level of discretion, or delegation, must satisfy the first-order condition for a maximum. The second-order condition is satisfied by our previous assumptions: indeed, given concavity of \( B \) and convexity of \( S_B \), \( C_{\text{ante}} \) and \( C_{\text{post}} \), the function \( B(d) - [S_B(d) + C_{\text{ante}}(d) + C_{\text{post}}(d)] \) is clearly concave in \( d \). Hence, there is a unique level of \( d \) such that (4) is maximized, and this level, call it \( d^* \), must solve the first-order condition \( \frac{dB}{dd} - \frac{dS_B}{dd} - \frac{dC_{\text{ante}}}{dd} - \frac{dC_{\text{post}}}{dd} = 0 \), i.e. \( \frac{dB}{dd} = \frac{dS_B}{dd} + \frac{dC_{\text{ante}}}{dd} + \frac{dC_{\text{post}}}{dd} \).
From the graphical point of view, the left-hand side of last expression, the marginal benefit \((MB)\) is decreasing (concavity of \(B\)), while the sum on the right-hand side, marginal cost \((MC)\) is increasing (convexity of all costs). The following figure illustrates our result.

**Figure 23. The Preferred Delegation Degree, \(D^*\), Given Marginal Benefit And Cost**

In the following subsections I proceed as follows. First, I identify the different control strategies available to legislators to ensure bureaucrat compliance. Then, I identify the factual conditions under which politicians are likely to delegate more versus less discretion (i.e., the optimal mix of \textit{ex ante} versus \textit{ex post} control). Finally, I examine how alternative delegation patterns differently impact outcome efficiency.

### 6.2.1.1. \textit{Ex Post} Control

During the 1980s political scientists developed the so-called “congressional dominance” theory.\(^{775}\) Although the theory focuses on the constitutional system of the United States,

it provides some generalizable insights that are useful for the purposes of process efficiency analysis. The core thesis is simple: Congress successfully controls bureaucracies through effective ex post oversight. Because legislatures and bureaucracies are repeat players – and since inefficient bureaucratic outputs can negatively impact the probability of a politician’s re-election – it is likely that re-election-minded members of legislatures would not tolerate ongoing bureaucratic inefficiency. In particular, supporters of the congressional control theory optimistically contend that bureaucratic behaviour can be explained “by modelling the determinants of individual and committee choice within the legislature.”

To be clear, the theory of congressional dominance recognizes that explicit monitoring is prohibitively costly for legislators, and that the latter do not have incentive to engage in direct and continuous ex post control of agency decisions and administration. This is coherent with our hypothesis that $C_{\text{post}}$ is convex. However, the theory contends that since legislators seek re-election they have strong incentives to ensure that the bureaucratic output meets the demand of their constituencies. Several instruments enable politicians to control bureaucracies. First, Congress can influence the bureau through the budgetary process by directing financial resources to those agencies that provide higher marginal political benefits. Contrary to the assumption of the Niskanen model, many agencies are subject to a competitive budgetary process. That is, they do not enjoy of a position of bilateral monopoly with Congress, rather “bureaucrats know that their success (e.g.,


larger budget [...] depends on satisfying congressional interests." Second, Congress is able to exert significant influence over bureaucratic behaviour through the committee system, which is seen as a specialized institution for mitigating agency shirking. The legislative committee system allows legislatures to create “a system of mini-monopolies that oversee particular agencies” thereby reducing the information asymmetries between legislatures and agencies. “Congressional committees possess numerous avenues for regulatory intervention, many of which have negative effects on bureaucratic careers. Ex post sanctions, therefore, create ex ante incentives for bureaucrats to serve congressmen.” Third, through the committee system Congress can influence the appointment of bureaucrats, thereby influencing the interests represented in bureaucratic processes. These institutional mechanisms provide an effective incentive system that substitute for direct monitoring in aligning bureaucratic interests with congressional goals.

It must be emphasized that – according to congressional dominance theory – legislators exert control over agency decision-making by listening to constituency reactions to bureaucratic outputs. The chief concern of legislators is re-election, consequently, they take interest in bureaucratic behaviour only to the extent that it meets the expected net benefits of their constituencies. It follows that legislators have little or no incentive to engage in a “police patrol” type of oversight; rather they have substantial incentive to respond to the “fire-alarm” triggered by unsatisfied constituencies.

Moe vividly summarizes this point:

Regardless of how ineffectively an agency pursues its mandate or how sloppily it is managed, legislators are satisfied as long as constituents do not complain. When they do complain – when the “fire alarm” goes off, in McCubbins and Schwartz’ [...] terminology – legislators spring into action, using their formidable array of weapons to bring errant agencies back into line. [...] Because agency officials seek to avoid sanctions and can anticipate the negative consequences of getting out of line, they generally stay within the limits set by Congress.

777 Barry R. Weingast, “Congressional-Bureaucratic System”, supra note 775 at 156.
778 Stearns and Zywicki, Public Choice, supra note 500 at 352.
779 Barry R. Weingast, “Congressional-Bureaucratic System,” supra note 775 at 155-156.
780 McCubbins and Schwartz, “Congressional Oversight”, supra note 775.
To summarize, congressional theory envisages a division of labour between legislators and their constituencies. Constituency groups (which have both the incentive and the expertise to monitor bureaucratic output) exercise effective monitoring of bureaucratic behavior, while legislators (who possess the authority to inflict sanctions) ensure the sanctioning of non-compliant bureaucrats.\textsuperscript{782}

The literature provides some evidence for the effectiveness of legislative \textit{ex post} control. For example, Weingast and Moran\textsuperscript{783} have shown how the nature of the cases considered by the Federal Trade Commission (FTC) during the period between the late 1970s and the early 1980s evolved in response to preference changes in the relevant oversight committees in Congress. Along the same lines, Leavens and Tollison have demonstrated that the case-selection activity of the FTC has been biased in favour of firms operating in the jurisdictions of members of congressional committees having important budgetary and oversight powers with respect to the FTC. Moe\textsuperscript{784} (1985) has provided a further example corroborating the congressional dominance theory by showing that changes of either or both the presidential administration and congressional committees play an important role in the determination of cases at the National Labor Relations Board (NLRB). Grier,\textsuperscript{785} who shows that the Federal Reserve Board is significantly responsive to its political principals, provides further evidence supporting the dominance theory. He identifies a “cycle effect” in money growth corresponding with presidential elections: it decelerates in the year after election, and accelerates in the subsequent three years as political pressure for an expansive monetary policy (by re-election-seeking presidents) increases.\textsuperscript{786}

\textsuperscript{782} For a critique of the congressional dominance theory, see \textit{Ibid}.
\textsuperscript{783} Barry R. Weingast and Mark J. Moran, “Bureaucratic Discretion”, \textit{supra} note 775.
As the previous examples confirm, the congressional dominance theory provides useful insights into the mechanisms underlying ex post legislative oversight. However, the theory suffers from one major limitation in that it overstates legislative power and treats bureaucrats as passive actors. In the next subsection more attention will be devoted to the substantive discretion enjoyed by bureaucrats and to the ex ante controlling instruments available to legislators.

6.2.1.2. Ex Ante Control

Monitoring, rewarding and sanctioning are unlikely to be effective strategies for inducing bureaucratic compliance. Ex post control strategies entail significant costs (which limits their effectiveness)\(^787\) and leave broad scope for bureaucratic shirking and slippage.\(^788\) To overcome these limitations, legislators have at their disposal alternative controlling strategies that enable them to achieve agency compliance with greater effectiveness and at lower costs. In particular, legislators can strategically define the content of the substantive discretion they delegate to agencies. In this subsection I expand on this point.

The substantive discretion legislators delegate to bureaucratic agents is determined by (i) the regulatory scope afforded to agents through legislative delegation, and (ii) the procedural requirements imposed on agency decision-making processes. Regulatory scope and procedural requirements are the two levers by which legislators control ex ante the substantive discretion delegated to bureaucrats.

\(^787\) First, monitoring imposes significant costs without enabling legislators to overcome the informational asymmetry between themselves and the bureaucrats. See, Mathew D. McCubbins, Roger G. Noll and Barry R. Weingast (“Mcnollgast”) “Administrative Procedures as Instruments of Political Control” (1987) 3 Journal of Law, Economics, and Organization, 243, at 250-251 [Administrative Procedures”]. Second, sanctions available to legislators are limited and often ineffective given the low detection probabilities caused by inefficient monitoring. Third, sanctions create high political costs for legislators, and the very act of imposing punishment entails high opportunity costs for both bureaucrats (who are distracted from delivering public service) and politicians (Ibid. at 252). See also T.J. Lowi, The End of Liberalism: The Second Republic of the United States, 2nd ed. (New York: W.W. Norton and Company, 1979) and Stearns and Zywicki, Public Choice, supra note 500 at 350.

\(^788\) On the distinction between shirking and slippage, see McCubbins, “The Legislative Design,” supra note 775 at 724.
Regulatory Scope. The regulatory scope defines “the domain of potential regulatory targets or problems the entity administering the regulation may address.” 789 By strategically defining the regulatory scope legislators can place “restrictions on the domain of policy alternatives available to the agent.” 790 This way, legislators can protect themselves against the risks of bureaucratic drift by eliminating from the regulatory scope some of their least preferred alternatives. The agent choice set is defined ex ante in such a manner as to be more in line with congressional intent.

In his foundational paper McCubbins 791 provides the example of the Toxic Substances Control Act (TSCA) in which the U.S. Congress delegated authority to the EPA by expressly limiting the regulatory scope to certain classes of chemicals. The Resource Conservation and Recovery Act (RCRA) provides another example in which Congress delegated to this same authority the substantive discretion to regulate waste materials and their management, but it precisely identifies the type of chemicals, the regulatory issues, and the regulatory objectives included in the regulatory scope. 792

Procedural Rules. Within the defined regulatory scope, legislators can reduce the agency’s substantive discretion and influence the final bureaucratic outcome by making strategic choices in defining procedural rules. Procedural requirements of the decision-making process include, for example: rules of standing and evidence; rules governing the allocation of burden of proof; the sequence of mandatory hearings to be followed by the agency; reporting and consultation procedures; access for third parties; standards for review; and so on.

Defining appropriate procedural rules represents a less expensive and a more effective means of solving the problem of bureaucratic compliance, as compared to ex post

---

789 McCubbins, “The Legislative Design,” supra note 775, at 726.
790 Ibid. at 731
791 Ibid.
mechanisms. Indeed, the strategic definition of procedural rules allows for three comparative advantages over ex post control. First, procedural rules enable legislators to deal with the information asymmetry problem by requiring that agencies gather and disseminate information. In general, informational procedural rules require “the agency to undertake steps to accumulate [and disseminate] information on potential policy alternatives.” Consider, for example, procedural rules requiring the agency to give notice that it will consider an issue, to solicit comments, to collect evidence, to construct publicly available records of information, and so on. Second, procedural rules enable legislators to affect the relative influence on the decision-making process of different interests affected by the bureaucratic outcome. Consider, for example, procedural rules that allow for the participation of external interests or for establishing evidentiary standards in the administrative process. Through these kinds of rules legislators can channel bureaucratic decision-making toward their favoured interests without directly evaluating outcome substantive efficiency. Third, procedural rules enable legislators to shift part of the enforcement costs to courts and constituency groups. For example, rules imposing public disclosure requirements reduce the agency’s informational advantage thereby increasing the efficacy of “fire alarms” by interest parties. Similarly, rules that establish standards for review or rules of standing can make it easier for constituents to bring suit thereby lowering enforcement costs for politicians. Finally, procedural rules affect the degree of agency responsiveness by influencing the costs to agencies of implementing policies.

To summarize, while the specification of regulatory scope explicitly limits the discretion of an agency to move outcomes from the status quo, the definition of ex ante

795 See, McNollgast, “Administrative Procedures,” supra note 787 at 256-261
796 Ibid. at 259-260
797 Ibid. at 263.
procedural rules affects the information available to the agent and the means for relating that information to the final decisions within the permitted set of feasible outcomes. In addition, procedural rules may reduce enforcement costs and, by favouring constituency controls, enable politicians to maintain their control and influence without their direct participation. By defining the regulatory scope and procedural requirements legislators minimize the risk of bureaucratic drift and concurrently avoid the problem of costly ex post oversight. This implies a trade-off between ex ante and ex post forms of control. All this is coherent with our hypothesis that a higher degree of delegation (supposing that the mix of ex ante and ex post stays more or less constant) reduces the ex-ante component of overall control costs \( \frac{dC_{ante}}{dd} < 0 \). However, one expects that this effect becomes smaller and smaller as the amount of delegation increases: in other terms, the ex-ante marginal contribution to the reduction of control costs decreases as delegation increases - \( \frac{d^2C_{ante}}{dd} > 0 \) - as we assumed beforehand.

6.2.1.3. The Determinants of Delegation

In the two previous subsections, I identified the control techniques available to legislators. When delegating powers to agencies and bureaucracies, legislators have three instruments for controlling the bureaucrat’s substantive discretion: (i) the breadth of the regulatory scope, (ii) the extensiveness of procedural requirements, and (iii) the intensity of the ex post oversight. With some simplification, we can identify four options available to legislators: (i) broad delegation, (ii) broad delegation with tight ex ante procedural constraints; (iii) broad delegation with intense ex post oversight (e.g., ex post veto on agency decisions); (iv) narrow delegation.

Let me turn now to the characteristics of the regulatory environments that most contribute to an efficient scope of delegation (i.e., more versus less discretion) and an efficient control strategy (ex ante versus ex post control). An efficient delegation strategy depends on the following factual predicates: \(^{799}\) (a) the degree of regulatory uncertainty; (b) the

---

\(^{799}\) For a survey of the most recent political science literature on political delegation strategies, see John D. Huber and Charles R. Shipan, “Politics, Delegation, and Bureaucracy” in Barry R. Weingast and
level of policy conflict between politicians and bureaucrats; (c) the asymmetry of information between politicians and bureaucrats. Depending on these factual variables the optimal mix of control strategies will take one of the four forms identified above.

6.2.1.3.1. Regulatory Uncertainty

Legislator willingness to delegate discretion is influenced by the level of uncertainty pertaining to the policy outcome. Two clarifications about the concept of uncertainty are useful. First, uncertainty is intended as “outcome-uncertainty,” that is, “the range of alternative policy choices about which the legislator has little or no information, and thus for which the legislator cannot discern a clear optimum.” Uncertainty makes the links between policies and outcomes less predictable ex ante. Second, the degree of uncertainty corresponds to the range of possible outcomes that could result from no agency action.

It is widely argued in the literature that as uncertainty increases legislators tends to prefer bureaucratic law-making to the legislative process. If we denote the degree of outcome uncertainty as $u$, then we may argue (in the terms of our model) that the benefit accruing to politicians from delegation increases with the amount of uncertainty.

This given, we can say that the politicians’ benefit is not only a function of the degree of delegation, it is also a function of the degree of uncertainty: hence, we write now $B = B(d; u)$, with the assumptions $\frac{\partial B}{\partial u} > 0$ and $\frac{\partial^2 B}{\partial d \partial u} > 0$: i.e. uncertainty makes delegation more attractive for politicians, in terms of both total and marginal benefit. Said differently, an increase in uncertainty raises the marginal benefit of delegation. The implication is that when uncertainty increases the preferred $d^*$ becomes greater, as the following figure shows.

Figure 24 The Effect Of Increased Uncertainty

---


802 See Epstein and O’Halloran, “Administrative Procedures,” supra note 798 at 704 (according to which uncertainty refers to “the outcome that would result if the agency took no action”).
Many factors can affect the degree of uncertainty of the regulatory environment. Uncertainty can be determined by the *complexity* of the regulatory issue (e.g., the pervasive unpredictable elements uncertainties in the assessment of exposure to hazardous substances). Complex issues require greater knowledge and more costly information, thereby increasing the marginal utility of delegating additional discretion to agencies and bureaucracies. For this reason, the conventional view holds that the greater the complexity of an issue and the more uncertain the decision-making, the *broader* the regulatory scope will be.

It is important to recognize, however, that the greater the uncertainty of the regulatory environment, the greater the need to control bureaucratic action. Thus, legislators are likely to respond to increasing complexity and uncertainty of the regulatory environment.

---

803 It often entails incomplete information regarding the costs of controlling risk, and the costs of expected losses (e.g., the risk of a new airline safety technology). One useful discussion about the uncertainty inherent to risk assessment and regulation is contained in Fraiberg and Trebilcock, “Risk Regulation,” *supra* note 759.


in two different ways. On the one hand, increased uncertainty leads legislators to enlarge the regulatory scope, thereby expanding the set of feasible alternatives for agency consideration. On the other hand – since the expanded regulatory scope coupled with increased uncertainty generates a greater need for legislative control – legislators will set up more constraining regulatory procedures to minimize the risk of bureaucratic drift. In short, increased uncertainty creates an incentive to delegate more (through expansion of the regulatory scope) and to control more (through increasing the level of procedural requirements).\footnote{McCubbins, “The Legislative Design,” supra note 775 at 736-42} McNollgast contend that the net effect of a broader scope coupled with tighter control reduces the degree of the agency’s substantive discretion.\footnote{McNollgast, “Administrative Procedures,” supra note 787}

Finally, it is worth noting that increased uncertainty involves increased (political and regulatory) risk for legislators. It follows that the impact of uncertainty on the decision to delegate is largely affected by the legislator’s attitude toward risk. The more risk-averse the legislator the more likely he or she will be to delegate discretion (because he or she will be less tolerant toward outcome-uncertainty). For example, in the area of drug safety (where informational complexity prevails over distributional concerns) legislators are likely to be more risk averse than in the area of tariff regulation; as a result, legislators are likely to delegate discretion in the former case than in the latter.\footnote{See Sharyn O’Halloran, Politics, Process, and American Trade Policy (Ann Arbor: University of Michigan Press, 1994).}

6.2.1.3.2. Conflict of Preferences between Politicians and Bureaucrats

Another decisive factor influencing legislators to delegate discretion to bureaucrats is the anticipated disparity in policy preferences between them. The existing literature emphasizes that legislator willingness to delegate discretion declines as agent preferences become more distant from the legislator’s ideal point. This is usually referred to as “the alley principle”: if legislators expect the bureaucratic outcome to be closer to their ideal point, then they will be more willing to delegate discretion; conversely, when politicians anticipate a bureaucratic outcome distant from their ideal point they will be less willing to
delegate discretion. Legislators assume that the greater the preference disparity between them and the agency the higher the bureaucrat’s incentive will be to behave to their detriment.

With some simplification the degree of conflict between the preferences of bureaucrats and those of legislators can be conceptualized in terms of the distance between two ideal points. The preceding considerations suggest that the attractiveness of delegation for legislators is a decreasing function of the distance between legislators’ and bureaucrats’ ideal points. This means that the outcome inefficiency deriving from bureaucratic drift is directly affected by the degree of policy conflict between bureaucrats and legislators. In short, a high level of policy conflict between legislators and bureaucracies reduces the benefits of delegation and increases agency slippage; as a consequence, higher policy conflict entails less willingness to delegate on the part of legislators.

Denoting the degree of conflict as $c$, we can formalise the preceding ideas by saying that when the degree of conflict increases the magnitude of the agency slippage is likely to increase. That is, $S_B$ is not a function of $d$ only, but also of $c$, and we assume:

$$S_B = S_B(d; c), \quad \text{with} \quad \frac{\partial S_B}{\partial c} > 0 \quad \text{and} \quad \frac{\partial^2 S_B}{\partial d \partial c} > 0$$

I.e., conflict raises both total and marginal slackness cost generated by delegation. The implication is that when conflict increases politicians reduce the preferred $d^*$. Of course, an efficient delegation strategy depends on all the characteristics of the regulatory environment. Thus, holding all other conditions constant, if high policy conflict is coupled with high distributional intensity, legislators will likely opt for narrow

---


810 Obviously the expected slippage is also higher, which increases the negative impact on $d$. In this model, the magnitude of the expected slippage is a function of the degree of preference conflict, which therefore implicitly recognizes the importance of the legislator’s attitude toward future bureaucratic drift. For simplicity, I prefer to assume constant risk aversion by legislators. I will endogenize the attitude toward bureaucratic drift in the next section when discussing strategic delegation.
delegation (or no delegation at all), as both of these features of the regulated environment reduce the expected benefit of legislation. By contrast, if high policy conflict comes with high regulatory uncertainty legislators are likely to opt for a broader delegation coupled with either tighter procedural constraints or more intense ex post legislative oversight. The net effect on the overall degree of delegation depends on which of the two sources, conflict and uncertainty, operates more strongly.

The conflict level between legislators and bureaucrats, on the other side, increases the costs of ex ante control, because legislators, anticipating a higher risk of bureaucratic drift, are likely to strengthen the procedural constraints. Hence we can write

$$C_{ante} = C_{ante}(d;c), \quad \frac{\partial C_{ante}}{\partial c} > 0$$

(6)

As we did before, we assume that $c$ raises also the marginal cost of delegation deriving from ex-ante control, i.e. $\frac{\partial^2 C_{ante}}{\partial d \partial c} > 0$. One must be careful in interpreting the meaning of this hypothesis: we assumed (expression 3) that an increase of the amount of overall delegation reduces the ex-ante cost ($\frac{dC_{ante}}{dd} < 0$); hence, we are now saying that this reduction becomes smaller and smaller as conflict increases. The increase of the marginal cost fostered by higher conflict induces politicians to narrow the preferred amount of delegation $d^*$. This is explained by the following figure.

---

811 This conclusion largely depends on the assumption of a non-strategic attitude on the part of the legislator. As I will explain in the next sub-section, a strategic legislator will tend to either delegate or remain inert in cases of highly conflictual distributional issues.
Once the scope of delegation has been determined according to the characteristics of the regulatory environment, the choice of an efficient mix of *ex ante* and *ex post* controls largely depends on the informational disparity between legislators and bureaucrats. The next subsection briefly elaborates on this point.

### 6.2.1.3.3. Asymmetry of Information

The degree to which information is verifiable *ex post* by legislators affects their choice of regulatory scope. If the information asymmetry between legislators and bureaucrats is low, legislators will be *more* likely to choose high-discretion delegation; by contrast, when the asymmetry is large legislators are *less* willing to delegate discretion, so they will narrow the scope of agency discretion through statutory precision.

As repeatedly emphasized, low-discretion delegation results in lower effectiveness because it reduces the agency’s ability to adapt to future changes in the regulatory environments and to fully employ its expertise, thereby diminishing the benefits of delegation. If we denote the degree of information inequality between legislators and bureaucracies as $\omega$, the relationship between information asymmetry and the benefit of delegation can be denoted as follows:
\[ B = B(d; u, \omega), \frac{\partial B}{\partial \omega} < 0; \frac{\partial B}{\partial u} < 0 \] (7)

which means that the larger the information imbalance between bureaucrats and legislators the less able legislators will be to make use of the benefits of delegation. We assume also that an increases in information asymmetry reduces the marginal benefit of delegation, that is \( \frac{\partial^2 B}{\partial d\partial \omega} < 0 \).

On the other side, the presence of asymmetric information produces a twofold effect on control costs. First, information asymmetry increases the total and marginal cost of control because both \textit{ex ante} and \textit{ex post} control becomes more costly per unit of delegation. The marginal costs of \textit{ex ante} control increase because legislators employ larger resources to specify and enforce the procedural constraints designed to deal with the information imbalance. The marginal cost of \textit{ex post} legislative oversight increases because, as discussed above, the asymmetry of information raises the costs of monitoring, rewarding and sanctioning. Hence, we can now write:

\[
\begin{align*}
C_{\text{ante}} &= C_{\text{ante}}(d; c, \omega); \quad \frac{\partial C_{\text{ante}}}{\partial \omega} > 0; \quad \frac{\partial^2 C_{\text{ante}}}{\partial d\partial \omega} > 0 \\
C_{\text{post}} &= C_{\text{post}}(d; c, \omega); \quad \frac{\partial C_{\text{post}}}{\partial \omega} > 0; \quad \frac{\partial^2 C_{\text{post}}}{\partial d\partial \omega} > 0 \\
\end{align*}
\] (8)

The preceding assumptions imply that a higher amount of information asymmetry raises the marginal cost of delegation, inducing a lower preferred level thereof. Figure 25 above can be used to illustrate this effect, since the main qualitative feature (an increased marginal cost) is the same.

Second, the presence of a large inequality in information, however, may change the relative costs of \textit{ex ante} versus \textit{ex post} control. In particular, information asymmetry raises the costs of \textit{ex post} oversight, thereby increasing the value for legislators of \textit{ex ante} control. As the information gap increases legislators will augment the intensity of procedural constraints designed to reduce informational inequality (e.g., by mandating agencies to collect and disseminate information, or by imposing reporting and consulting procedures on the agency). In short, the presence of informational asymmetries between
legislators and politicians generates a substitute effect as the relative cost-effectiveness of \textit{ex ante} controls increases.

In our setup, the preceding observations might take the following form: we might assume that the marginal effect of asymmetric information on \textit{ex-post} delegation costs is higher than the marginal effect on \textit{ex-ante} ones, i.e. \( \frac{\partial C_{\text{post}}}{\partial \omega} > \frac{\partial C_{\text{ante}}}{\partial \omega} \). In addition, we might assume that also the second-order effect of unequal information on delegation costs is greater in the case of \textit{ex-post} costs, i.e. \( \frac{\partial^2 C_{\text{post}}}{\partial d \partial \omega} > \frac{\partial^2 C_{\text{ante}}}{\partial d \partial \omega} \). In a full-fledged model, one should consider explicitly the presence of two instruments available for politicians, namely the \textit{ex-ante} degree of delegation \( (d_{\text{ante}}) \) and the \textit{ex-post} degree of delegation \( (d_{\text{post}}) \). We leave this more sophisticated setup for future research.

Finally, it is worth noting that the preceding discussion does not necessarily hold for the highest levels of regulatory issue \textit{complexity}. In this latter case, the \textit{ex ante} limitation of agency discretion might excessively reduce the benefits of delegation, and impose prohibitive costs on agency decision-making (e.g., exceedingly rigid and cumbersome procedures for the agency to follow). In this latter case, the \textit{ex post} control of the regulatory outcome might be more cost effective despite all the limitations emphasized above.

The foregoing discussion of the factual predicates affecting the structure of efficient legislative delegation can be summarized as follows. The problem faced by politicians is

\[
\max_d E(d;u,c,\omega) = B(d;u,\omega) - S_B(d;c) - C_{\text{ante}}(d;c,\omega) - C_{\text{post}}(d;c,\omega) \tag{9}
\]

Given the environmental variables \( u, c \) and \( \omega \), the politicians’ preferred degree of delegation is the (unique) value \( d^* \) such that the partial derivative of \( E \) with respect to \( d \) is zero. The presumable effects of changes in the environmental variables are the following: an increase in uncertainty, \( u \), raises \( d^* \); an increase in the degree of conflict between
politicians and bureaucrats, c, or in the degree of information asymmetry, \( \omega \), makes \( d^* \) go down. An increase of information asymmetry might also induce a substitution between the \textit{ex-ante} and \textit{ex-post} degree of control, making the former increase, unless the complexity of ex-ante regulation of agencies is very high.

6.2.2. Legislators as Strategic Agents

In the previous subsection, it was assumed that legislator and voter preferences are perfectly aligned, and that voter interests and policy preferences are reflected (via the electoral connection) in legislature preferences. However, in real-world politics the choice to delegate law-making authority (as an alternative to regulating directly via legislation) might be anchored in the legislator’s political calculus, rather than in his or her pursuit of agency expertise. In this subsection, I investigate the strategic use of delegation by legislators and its efficiency implications in terms of increased agency costs.

As emphasized in the introduction to this chapter, the essential feature of the bureaucratic agency is its \textit{insulation} from the electoral connection and (to a varying degree) from direct political influence. Legislators can strategically use delegation as an insulating device – i.e., \textit{they can delegate discretion to agencies so as to rid themselves of it}. I identify three cases in which delegation is used as an “insulating device”: (i) to blame-shift, (ii) to avoid legislative-drift, (iii) to enhance credibility. In all cases, \textit{insulating delegation does not lead to efficient law-making}: legislator use of bureaucratic \textit{insulation} often maximizes personal utility at the expense of law-making efficiency.

6.2.2.1. Blame-Shifting Delegation

The two major concerns for legislators in their relationship with constituencies are those of “shifting the blame” for the costs imposed by legislation and of “taking credit” for the benefits provided through legislative action. Fiorina has linked these two major concerns with a politician’s choice to delegate.\footnote{See Morris P. Fiorina, “Legislative Choice” \textit{supra} note 775; Morris p. Fiorina, “Group Concentration and the Delegation of Legislative Authority California” (1982) Institute of Technology Pasadena California Mimeo; Morris P. Fiorina, “Legislator Uncertainty, Legislative Control, and the Delegation of Legislative Power” (1986) 2 JL Econ. and Org., 33. On blame-shifting delegation, see also Aranson, et} He contends that delegation \textit{dilutes} the costs and
benefits attributed by constituents to legislative action; this way, legislative delegation enhances the ability of politicians to avoid blame but reduces their ability to claim credit. That is, when deciding whether to delegate, legislators face a trade-off between the gain of an enhanced ability to shift the blame and the loss of a reduced ability to claim credit. It follows that legislators will delegate law-making authority to agencies when this enables them to disguise the costs of regulation to a larger extent than delegation hides the benefits of legislative action. Fiorina’s insight has an important implication for the purposes of this study: legislators might delegate law-making authority to the bureaucratic process even in cases where delegation is not justified in terms of productive efficiency. Namely, it might be the case that significant portions of legislative authority are shifted to the bureaucratic process even if the agency costs associated with delegation are not justified by the expected benefits of agency expertise.813

Fiorina’s model, coupled with the Wilson-Hayes model814 enables us to identify the circumstances under which politicization is more likely to result in strategic delegation to agencies, with consequent unjustified increases in agency costs. First, legislators are expected to make use of delegation when the costs imposed by legislation are concentrated on narrow, well-identifiable groups while the benefits are widely distributed. When this circumstance occurs (concentrated costs, distributed benefits) the legislator is likely to remain inert (i.e., neither delegating nor legislating); that is, when the per capita benefits of the proposed regulation are so thin as to generate insufficient political advantages compared to the costs of delegating.815 Conversely, legislators will delegate authority to agencies when the political benefits of the proposed legislation more than offset the costs of delegation (despite the fact that the economic benefits are more

813 In this sense, the bureaucratization of the law-making process can be seen, at least in part, as a predictable consequence of the tendency toward politicizing law-making.

814 Wilson, Political Organizations, supra note 657 at 332-337; Hayes, Lobbyists and Legislators, supra note 657. For a textbook discussion of the Wilson-Hayes matrix, see Stearns and Zywicki, Public Choice, supra note 500 at 69.

815 Here I am assuming that the total economic benefit from the proposed legislation is not inferior to the total costs. Notice that the political benefit for legislators of a proposed legislation is a function of the per capita economic benefit.
widely distributed than the costs). It must also be recognised that, when weighting the relative magnitude of the “blame-shifting gain” with the “credit-taking loss” legislators may consider that voters – due to the loss-aversion phenomena – are more likely to remember (and to take into greater account) legislative actions that cost them than those providing them with benefits.\footnote{816 Morris P. Fiorina, Representatives, Roll Calls, and Constituencies (Lexington, MA: Lexington Books, 1974) at 38-39.}

A second circumstance that makes legislative delegation an attractive option for legislators is the emergence of a highly conflictual policy issue; stated differently, delegation becomes attractive when “[t]he affected constituencies on both sides of an issue are roughly evenly balanced” and “one constituent groups benefits directly at the expense of another.”\footnote{817 Stearns and Zywicki, Public Choice, supra note 500 at 71} In such cases, delegation enables legislators “to claim credit for facilitating a potentially responsive solution to the issues that the constituents raise, while at the same time also shifting blame to the agency when the results are other than what one side desires.”\footnote{818 Ibid. at 257. This point is also clarified in Aranson et al., “Legislative Delegation,” supra note 805, at 56-62.} Stearns and Zwicky identify the National Labor Relations Act passed by the US Congress in 1935 to establish the National Labor Relations Board as an example of blame-shifting delegation in a context of conflictual regulatory issues.\footnote{819 Stearns and Zywicki, Public Choice, supra note 500 at 74.} To avoid the political costs associated with the highly conflictual distributive issues in the area of industrial relations politicians found it convenient to delegate a significant portion of law-making authority to the National Labor Relations Board.

6.2.2.2. Uncertainty-Reducing Delegation

The impermanence of property rights is an important feature of the political process: those who are in power today cannot ensure that the regulatory policies they create will not be subverted or completely destroyed by tomorrow’s power-holders. This transitoriness generates political uncertainty. In this context, delegation becomes attractive for politicians as an uncertainty-reducing device. Politicians delegate legislative authority to bureaucrats as a way of diminishing political uncertainty and of minimizing
the risk that group opponents and future authorities will reverse the policies they create. Put shortly, politicians use delegation as an isolationist device to protect their regulatory policies from the threat of future transformations in regulatory policies as a result of changes in the composition of legislatures or oversight committees (i.e., “legislative drifts”).

Election-seeking politicians have strong incentives to protect current political deals with their constituents against future legislative drift, thereby increasing the actual value of their services and maximizing political support. In fact, political uncertainty lowers the expected value for constituents of favourable regulatory treatment, thereby diminishing the willingness of interest groups to lobby for favourable legislation; by contrast, when legislation has a longer life expectancy interest groups are more disposed to investing in political influence. The delegation of legislative authority to agencies insulated from direct political influence is a way for enacting coalitions to lock in the regulatory bias generating benefits for favoured constituents.

It is worth noting that this explanation of legislative delegation is partially at odds with the assumption that legislators seek to minimize the risk of bureaucratic drift. Indeed, because of the incentives to reduce political uncertainty, politicians tend in some cases to design bureaucracies that they cannot control; that is, they favour bureaucratic structures that minimize the opportunity for external control rather than procedures favouring political influence. A legislator’s main incentive is to create autonomous bureaucracies insulated from political control to protect the exchange relationship with their constituents against future legislative drift. Thus, there is a trade-off between legislative and bureaucratic drift. To diminish the risk of bureaucratic drift legislators must enhance the agency’s responsiveness to congressional demands, thereby increasing the risk of future coalitional drift. Conversely, to reduce the likelihood of legislative drift they have to further insulate bureaucracies from political control, thereby increasing the

---

820 Terry M. Moe, “Political Institutions: The Neglected Side of the Story” (1990) 6 J. L. Ec. and Org, 213, at 228 (“Political Institutions”) (“In so doing, of course, they will not only be reducing their enemies’ opportunities for future control; they will be reducing their own opportunities as well. But this is often a reasonable price to pay, given the alternative”).
risk of bureaucratic drift. In essence politicians must optimize the trade-off between *insulation* from and *responsiveness* to political influence.

Process efficiency analysis requires identifying the conditions of regulated environments that determine the optimal degree of insulation. De Figueiredo has shown that the solution to the responsiveness/insulation trade-off depends on (i) the degree of political *uncertainty*, and (ii) the *costs of insulation*.\(^{821}\) Political uncertainty is a negative function of the probability that the enacting coalition will be re-elected. The costs of insulating bureaucracies are: (i) the increased risk of bureaucratic drift, and (ii) the decreased opportunity to exert political influence in the event of re-election (the “self-binding” effect). That is, greater insulation entails higher risk of bureaucratic drift and more intense self-binding for politicians. The *actual* costs associated with the self-binding effect decrease as the political uncertainty increases; this means that as the level of political uncertainty\(^{822}\) rises, the *relative* cost of insulation versus responsiveness decreases.

The main result of De Figuereido’s model is that at moderate levels of political uncertainty, politicians will choose *not* to insulate unless the costs of doing so are low. Moderate political uncertainty increases the opportunity costs of insulation, as the re-elected coalition will have fewer instruments for minimizing bureaucratic drift and for influencing regulatory policies. Thus, if the costs of insulation are significant legislators will not choose it. A political coalition that expects to remain in power will value the ability to influence regulatory policy in the future more highly and the benefits of the insulation strategy less highly. By contrast, *under conditions of political uncertainty, politicians normally choose to insulate unless the costs of insulation are prohibitive*. By opting for insulation current power-holders accept losing control over the law-making process (i.e., the ability to adjust regulatory policy in a low cost fashion) so as to ensure greater durability for their current deals with favoured constituencies. To put it shortly, under high political uncertainty first-mover politicians prefer rendering tomorrow’s

---


\(^{822}\) The 0.5 probability of being re-elected for each party.
Legislators increasingly impotent even at the cost of losing actual control over the regulatory process. As political uncertainty increases, the costs of bureaucratic drift are discounted more than the costs of legislative drift.

The uncertainty-reducing argument provides one possible explanation for the increasing importance assumed by independent regulatory agencies in many European countries since the 1980s. Another example is the creation of independent central banks, which might be explained, at least in part, by the desire of politicians to limit the discretion of their successors in monetary policy-making. In addition, Volden has provided empirical corroboration of the uncertainty-reducing argument in his study of the Aid to Families with Dependent Children program across the American states from 1935 to 1966. He found that declining coalitions (those losing political strength within state legislatures) are more than four times as likely to empower bureaucracies and to isolate them from the political process. Further supporting evidence comes from Holburn and Vanderbergh who found that in the period 1970-90 pro-consumer political coalitions were more likely to create new consumer advocacy institutions independent of existing regulatory agencies when they where uncertain about remaining in power at the next election. Finally, De Figueiredo has provided supporting evidence for the uncertainty-reducing hypothesis through the study of the line-item veto in the budgetary process. He advances the hypothesis that many states in the U.S. have recently adopted the line-item veto because it provides fiscal conservative parties fearing the loss of power with a mechanism for insulating budgetary decisions from the influence of future majorities.

6.2.2.3. Credibility-Enhancing Delegation

As discussed in chapter 3, *ex ante* political law-making poses the problem of time-inconsistent behaviour, which undermines the credibility of the lawmaker. Scholars have argued that politicians can use delegation as an insulationist-device to increase the credibility of the lawmaker.\(^{828}\) In essence, legislators decide to give discretion to technocrats to remove it from politicians who lack sufficient credibility with the electorate.\(^{829}\) The creation of a politically insulated agency might restore the credibility of the law-making process, thereby making regulatory commitments believable.

The standard example of credibility-enhancing delegation is the creation of independent central banks as a means to ensure that monetary policy is credibly committed to the pursuit of price stability. This appears to be the case with the institutional independence of the European Central Bank (ECB), which was designed as a key element in the euro area’s governance framework, as it ensures that price stability is effectively achieved and maintained. The Maastricht Treaty established the independent status of the European Central Bank (ECB) as a bulwark for guaranteeing that the future Eurozone would benefit from price stability.

Likewise, delegation to independent regulatory agencies can be a means of making regulatory policies for the public utilities sector more credible. As Gilardi\(^{830}\) has pointed out, in the utilities sector the credibility problem is exacerbated by high asset-specificity – which make investment highly sensitive to the credibility of the regulatory authority – and by high political-salience – which make investors fear the risk of collusion between regulators and the powerful incumbent. The need for assuring investors about the credibility of the regulatory institutional framework explains the centrality of independent regulatory agencies in market regulation in many European countries.

It is important to recognize that credibility-enhancing delegation has implications that are at odds with the principal-agent perspective discussed in subsection 8.2.1. First, instead

\(^{828}\) See Gilardi, Delegation, *supra* note 823 at 38-46.

\(^{829}\) Moe, “Political Institutions,” *supra* note 820 at 228.

\(^{830}\) See Gilardi, Delegation, *supra* note 823 at 43.
of delegating to “allies,” legislators delegate to agents who have other preferences. Second, instead of increasing \textit{ex post} control mechanisms to compensate for the broad discretion delegated to agencies, legislators remove opportunities for external control by making agents independent from politics.

It is therefore difficult to formulate a generalizable, \textit{a priori} statement about when credibility-enhancing delegation is efficient. However, it seems that because credibility-enhancing delegations dramatically increase the autonomy of technocratic law-making they substantially augment the problems resulting from agency centralization (e.g., the incapacity of the lawmaker to take into account the local preferences of people subject to law). This appears to be the case with the institutional law-making design of the European Union. The members of the European Union delegated\textsuperscript{831} the legislative power to relatively insulated institutions, thereby setting the conditions for the growing difficulty of ensuring that the law-making process reflects the preferences of the people subject to the European law.\textsuperscript{832}

\section*{Conclusions}

\textit{Information Efficiency}

Bureaucracies enjoy the advantages of specialized knowledge. They \textit{possess} greater technical information and can \textit{acquire} relevant information at lower costs than legislatures and other law-making processes. However, these comparative advantages should be weighted against the costs of bureaucratic insulation.\textsuperscript{833} The insulation of the lawmaker is the chief structural feature of technocratic law-making. Indeed, unlike bureaucracy, courts are confronted by pressure from private litigants; legislators face political competition generated by the electoral connection; and private orderings face

\textsuperscript{831} To be sure, here I am using the word “delegate” in economic terms, independent of the fact that legally speaking the creation of European institutions cannot be qualified as an act of delegation.

\textsuperscript{832} The maladaptation costs generated by European law are so significant in some cases as to give the impression to part of the population that they exceed the costs of exiting the Union.

\textsuperscript{833} To be clear, the efficiency and effectiveness of bureaucratic processes vary significantly according to the characteristics of the institutional setting within which agencies and bureaucracies operate. However, the “insulation” of bureaucracy (i.e., the disconnection from the electorate, and the insulation from market pressure) tends to generate the three tendencies analyzed in the text.
incentives similar to those of a competitive environment. The insulation of agencies and bureaucracies generates three core tendencies in bureaucratic processes: (i) over-supply, (ii) X-inefficiency, and (iii) output-ineffectiveness.

First, bureaucracies tend to produce a global volume of activity that exceeds the Pareto-optimal level. This likely explains, at least in part, the increasing level of complexity that regulation exhibits in many areas of law and why the high degree of complexity in regulation is not always related to the level of output effectiveness. Second, bureaucrats exhibit a tendency toward low productivity. This is determined in part by their input preferences, and in part by their risk aversion. In general terms, the bureaucratization of the law-making process is likely to result in an excess of regulation and in excessive use of input factors. Third, technocratic law-making can fail on the grounds of output ineffectiveness. The bureaucrat’s risk aversion generates a systematic bias in favour of avoiding Type II errors, despite the fact that Type I errors might entail equal or even higher social costs. This further explains why bureaucrats tend to generate an inefficiently high level of regulation. Another cause of output-ineffectiveness is related to the tendency of ex ante centralized lawmakers (such as bureaucrats) to externalize the costs of their decisions. This tendency is only partially mitigated by legislators’ control strategies, and is exacerbated by the severe difficulties of identifying, measuring, and estimating the costs and benefits of regulation. It is important to emphasize that the difficulties associated with implementing cost-benefit analysis increase with the complexity of the regulatory environment. Finally, agencies and bureaucracies exhibit a tendency toward systematic biases, such as “tunnel vision,” “random agenda selection” and output “inconsistency.”

The foregoing considerations suggest that the call for technocratic law-making as a remedy for the informational inefficiency of political law-making should always be weighted against the three above-mentioned sources of productive inefficiency: overregulation, cost-ineffectiveness and output ineffectiveness.

834 Complexity of the regulated environment is one of the major reasons provided to justify the delegation of law-making authority to agencies.
Agency Efficiency

The delegation of law-making power to agencies and bureaucracies dramatically increases the level of the agency costs. In particular, there are two sources of agency costs associated with delegation: bureaucratic drift and the strategic use of delegation by legislators.

(i) Bureaucratic Drift.

The risk of bureaucratic drift rises with the increasing insulation of the bureaucracy from the political process. The costs of bureaucratic drifts are (i) outcome inefficiency costs generated by agency slack, and (ii) control costs associated with ex ante and ex post mechanisms of legislative oversight. Three characteristics of the regulatory environment largely affect the structure of the agency costs of delegation: regulatory uncertainty, conflict of preferences, and asymmetry of information between politicians and bureaucracies.

First, the technical complexity of the regulatory issue increases the marginal utility of delegating additional discretion to agencies, while simultaneously augmenting the demand for legislative control of bureaucratic actions. Holding constant all other conditions, the relative efficiency of delegation increases with the rising degree of technical complexity. As complexity increases delegation is likely to be characterized by a rising level of procedural requirements and other ex ante control mechanisms to counterbalance the growing regulatory scope. In addition, the expected benefit of delegation increases with the risk-aversion of the lawmaker. Risk-aversion is related to the type of interests regulation aims to protect. For example, the lawmaker acting as a faithful agent is likely to be more risk-averse when regulating matters concerning the protection of human life and personal safety on a mass scale than when regulating tariffs on imported goods, public pensions, or immigration.

The benefits of delegation decrease when legislators expect to have conflicting preferences with regulatory agencies, or when there is large information asymmetry
between legislators and bureaucrats. Because of policy conflicts and informational asymmetries, the marginal cost of *ex ante* control increases. In the case of information asymmetry the marginal cost of *ex post* oversight also increases. Thus, in regulatory environments characterized by either policy conflict or information asymmetry (or both) the *total* costs of delegation to technocratic law-making increase.

In the frequent scenario where the regulatory environment is characterized by high regulatory complexity and high lawmaker risk-aversion the expected benefits of technocratic law-making increase. However, along with increased benefits, the agency costs of delegation also rise in terms of both control costs and expected agency slack. In fact, regulatory complexity and outcome uncertainty are likely to generate informational asymmetries between bureaucrats and legislators, thereby increasing the *ex ante* proceduralization of delegation and, more generally, a higher level of total control costs. At the same time, information asymmetry increases the expected costs of agency slack, which further augments agency costs.

(ii) **Strategic Delegation**

The second source of agency costs is the strategic use of delegation by legislators who may choose to grant law-making authority to the bureaucracy – *not* to make use of the agency’s superior knowledge but to pursue personal political goals. These types of delegation can be referred to as “strategic” to emphasize that legislators seek insulation for reasons independent of efficiency in the law-making process.

In the case of *blame-shifting* delegation and *uncertainty-reducing* delegation legislators grant authority to the bureaucratic process *independently* of the need for the agency’s expertise. That is, legislative delegation is not a device to enhance the informational capacity of the law-making process, but rather an instrument to minimize the “political” costs of legislative action. This implies that there is no comparison between the increasing agency costs associated with law-making and the informational advantages of delegation. *Blame-shifting* delegation is motivated by the increased ability of the lawmaker to shift the political costs of legislative action. *Uncertainty-reducing* delegation
is a strategy to lock-in the regulatory bias to the advantage of favoured constituencies and to ensure more durability of the rents provided by politicians to favoured interest groups. To put it bluntly, strategic delegations entail increased agency costs without increased productive efficiency in law-making. In these cases, technocratic law-making is a “by-product” of the political struggle that exacerbates the agency costs of political law-making without enhancing its ability to meet the preferences of the people subject to law.

Unlike the two preceding types of strategic delegation, credibility-enhancing delegation can provide some advantages in terms of agency efficiency when the time inconsistency problem related to the lack of credibility of politicians significantly undermines the effectiveness of political law-making. More precisely, when the relative costs of the time inconsistency problem exceed those of insulation credibility-enhancing delegation might be beneficial for the overall efficiency of the law-making process. It is not possible to formulate a generalizable a priori statement on when credibility-enhancing delegation is efficient. However, it might be the case that significant portions of law-making authority are shifted to the bureaucratic process even if the agency costs associated with delegation is not justified when compared to the expected benefits of enhanced credibility. This is because credibility-enhancing delegation dramatically increases the degree of agency insulation, and removes opportunities for external control of bureaucratic behaviour.

The foregoing discussion suggests the following summarizing conclusions.

(i) Technocratic law-making is well suited to remedying the informational inefficiency of political law-making – which is rooted in the generalist nature of political representation – in cases of high regulatory environment complexity and associated outcome uncertainty.

(ii) The informational advantages of technocratic law-making do not necessarily entail a relative advantage in terms of productive efficiency. Informational superiority of agencies and bureaucracies should be weighted against the bureaucratic tendency toward overregulation, cost-ineffectiveness and output-ineffectiveness.

835 In this sense, the bureaucratization of the law-making process can be seen, at least in part, as a predictable consequence of the tendency toward the politicization of law-making that I have discussed in the introduction to this chapter.
(iii) In most cases technocratic law-making is ill-suited to remedying those inefficiencies of political law-making that depend on the *ex ante* centralized structure of the law-making process. Stripped to its essence, the “technocratization” of law-making does not resolve the problem of the misalignment of incentives between the *ex ante* centralized lawmaker and the people subject to law. More bureaucracy entails more expertise, but not *less* “*ex ante* centralization.” Hence, technocratization shifts the agency problem from legislative to bureaucratic drift. Which one is less costly is hard to establish through an *a priori* analysis of the lawmaker’s incentive structures.

(iv) In the case of *strategic* delegations, technocratic law-making may increase the level of agency costs, as compared with the political law-making, without providing any improvement in terms of productive efficiency, thereby resulting in a net utility loss for society.
In this chapter, the principles of a process efficiency analysis of the adjudication process are developed. The discussion focuses on identifying the relative advantages and disadvantages of adjudication versus politics as a method of producing efficient legal rules.

Introduction Adjudication as a Triadic Law-making Process

The adjudication process can be usefully conceptualized as a triadic mode of law-making, in which a triad—two disputants and a judge—is the protagonist of the production of law. The triadic model of adjudication is defined by three elements: emergence, mechanism, and dimensions.

Emergence. In its most elementary form, adjudication is activated by the decisions of two individuals in a dyadic relationship to delegate the solution of a controversy between them to a third impartial adjudicating authority. This act of delegation follows an earlier failure to solve the dispute dyadically (e.g., through settlement). From an economic standpoint, the adjudicating triad is likely to emerge when the dispute delegated to a third party is less costly, or more likely to yield a better outcome, than negotiating to settle the

---

836 In this chapter, I use the terms “adjudication” and “judicial” processes synonymously. I refer to (i) the outcome of the judicial process as “judge-made law”, (ii) the actors on the supply side as “judges”, and (iii) the actors on the demand side as “litigants”.

452
controversy or exiting the dyadic relationship would be. Stated differently, the adjudication process emerges when: (i) the spontaneous law is inefficient or ineffective, (ii) the value attributed by the parties to the dyadic exchange exceeds the costs of litigation; (iii) the expected net benefit of litigation exceeds that of settling the dispute out of court.

Mechanism. The doctrine of precedent is essential to the law-making function of adjudication. Broadly, the doctrine of precedent is a legal doctrine in force of which past case decisions become the source (persuasive or binding) of legal authority for future judges in similar cases. In resolving disputes arising from failed dyadic contacts, judges produce rules for subsequent rounds of litigation.\(^837\) That is, from an economic standpoint, the doctrine of precedent is an institutional mechanism transforming private litigation in a regulatory mechanism.\(^838\)

Dimensions. By virtue of the doctrine of precedent, judge-made law operates both retrospectively (i.e., affects the dyadic relationship between actual litigating parties) and prospectively (i.e., affects future dyadic exchanges). In our terminology, adjudication brings together the ex post and the ex ante dimensions of law-making: on the one hand, it produces particular retrospective rules that resolve pending disputes; on the other hand, it produces general prospective rules that provide a normative structure for future dyadic relations.

To summarize, three elements are crucial in providing an economic understanding of the adjudication process:
(i) the adjudication process is activated upon the decision of private parties;
(ii) because of the doctrine of precedent, individual decisions to litigate generate “regulatory externalities” in future cases;
(iii) judges produce both retrospective solutions and prospective rules to provide a normative structure for future dyadic exchanges.


\(^{838}\) Private parties’ decisions to bring their dispute before the court activate an adjudication process whose output generates a regulatory externality upon future decisions of similar cases.
The triadic structure of adjudication produces a legal evolutionary response to the demand for law: litigation leads to adjudication, adjudication produces legal rules providing a normative structure for future dyadic contacts, future dyadic contacts influence future litigation, and so on. Figure 26 summarizes this evolutionary dynamic.

Figure 26 Judicial Evolutionary Law-making
7.1. Productive Efficiency

This section provides an economic analysis of the following four elements that affect the productive efficiency of adjudication: (i) the individuals’ “settlement-trial decisions”; (ii) the doctrine of precedent; (iii) the judges’ informational constraints; (iv) the degree of polycentricism in the organization of the sources of law. The two central questions are (a) *how* do these factors affect the productive efficiency of adjudication, and (b) *what* are the conditions (on both demand and supply sides) that facilitate the judicial production of efficient judge-made law?

7.1.1. The Demand Side

Judges only rule in *litigated* cases.\(^{839}\) Hence, the settlement-trial decisions of private parties *select* the flow of legal matters adjudicated by judges. Based on this simple observation, the economic theory of adjudication has traditionally focused on the determinants of settlement-trial decisions.\(^{840}\) Adjudication has been assumed a *demand*-driven phenomenon or a “private good”.\(^{841}\) From this perspective, the efficiency of judge-made law is a “by-product” of the type of disputes brought by private parties before courts, instead of the result of judges’ deliberate efforts to promote legal efficiency. The only inputs into adjudication that count are *private*; judges play a passive role. In essence, judge-made law is the result of an “evolutionary” mechanism triggered by settlement-trial decisions, independent of judges’ attitudes towards efficiency.

In order to identify the strengths and limitations of the demand-side approach, it is worth identifying its underlying simplifying assumptions: (a1) Individuals are assumed rational and utility maximizing. (a2) Settlement-trial decisions are determined by (i) the magnitude of the *stakes* to the dispute, (ii) the relative *costs* of settlement and litigation,

---

\(^{839}\) For clarity, I use the term litigation to refer to the litigation process before judges. I do not consider alternative dispute-resolution mechanisms.

\(^{840}\) This literature has focused mostly on common law. It aims to find arguments supporting the “hypothesis of common law efficiency. This has determined a reaction amongst scholars from civil law jurisdictions, which are convinced that non-compelling arguments support the thesis of the superiority of the common law system over the civil law system. Here, I do not address this comparative issue. I think that given the high level of abstraction, many of the models developed in the literature on the “efficiency of common law hypothesis” illuminate the evolution of judge-made law as such, independent of the specific characteristics of different jurisdictions.

and (iii) the parties’ estimates of the outcome of litigation. Precisely, litigation occurs when the minimum settlement demand of the plaintiff exceeds the maximum settlement offer of the defendant by a substantial amount; in the opposite case, parties will find it convenient to settle the dispute out of trial, thus saving the costs of litigation. (a3) It is generally recognized that the likelihood of litigation is a function of private parties’ expectations of success. The economic content of settlement offers and demands is dependent upon the magnitude of the expected benefit from success at the trial, multiplied by the private parties’ estimated probability of victory. (a4) It is generally assumed that litigation costs exceed settlement costs, so litigation occurs only when the expected benefit from trial of the plaintiff exceeds the expected loss of the defendant by an amount exceeding the cost differential between litigation costs and settlement costs.

Based on the foregoing assumptions, two alternative approaches have been developed to explain the evolution of judge-made law: the “stakes” approach and the “information” approach. The following discussion examines these two alternative explanatory models, and emphasizes the importance of the “public good” aspect of legal precedent, which is a major challenge to demand-driven theories.

7.1.1.1. The Stakes

Priest, Rubin, and Goodman developed three explanatory models of common law evolution that laid the conceptual basis of the prevailing adjudication economic theory. Although they do not demonstrate a clear tendency toward efficiency, these models provide useful insights into identifying the conditions under which judge-made law is

---

842 In other words, individuals opt for litigation when the expected return on trial exceeds the costs imposed by substantively inefficient legal rules.


more likely to evolve toward efficiency. These three explanations can be summarized as follows: (i) inefficient rules are more frequently litigated than efficient rules are; (ii) inefficient rules are more frequently litigated by parties having future stakes in the dispute; (iii) inefficient rules are more intensively litigated than efficient rules are. Let me briefly elucidate these three summary statements.

7.1.1.1.1. Inefficient rules are more frequently litigated than efficient rules.

Priest maintains that common law evolves toward efficiency because inefficient rules are more likely to be litigated than efficient rules are.845 The differential rate of litigation between efficient and inefficient rules leads the adjudication process to generate a larger proportion of efficient rules gradually over time. Essentially, the reasoning is as follows. Rules that are more inefficient impose greater costs on individuals, thereby increasing the expected return from litigation. As the costs of substantive inefficiency rise, individuals have increasing incentives to invest in litigation to change the rules in their favour. Conversely, rules that are more efficient impose smaller costs, thereby reducing the expected return from litigating rules before courts. Crucially, in this model, the rate of litigation is an increasing function of the degree of substantive inefficiency: greater inefficiency triggers a higher litigation rate, thereby increasing the likelihood that the inefficient rule is re-examined.

According to the foregoing explanation, the mechanism generating efficient judge-made law is evolutionary in nature; namely, it is independent of the actors’ (judges and parties) attitudes toward legal efficiency. Judges, by assumption, decide cases randomly; and parties seek exclusively to maximize their individual utility function, without any concern for the aggregate efficiency of judge-made law. The tendency toward efficiency results from private parties’ selection of the inputs into adjudication, based on the “efficiency-differential” between alternative legal rules. At each round of litigation, the input selection process increases the ratio between the portion of efficient and inefficient outcomes established by judges. This is not because judges value efficiency more than

---

inefficiency, but because they *adjudicate* a larger number of *inefficient* rules than they do efficient ones.  

The logic of this explanation resembles that of the “invisible hand”: rational individuals pursuing their own particular interests *unintentionally* produce an efficient collective outcome. The underlying theory is the *commodification* of legal rules, which allows the equation of individuals’ decisions to litigate to individuals’ decisions to buy goods or services in private markets. The idea that inefficient rules are subject to a higher rate of private litigation is based on the logic of *Coaseian* bargaining: individuals in the adjudication arena litigate around the judicial allocations of legal entitlements, similar to Coaseian private markets, in which individuals *bargain* around the inefficient allocation of legal rights. In some ways, Priest’s explanatory model can be conceptualized as an “Adjudication Coase Theorem”. It demonstrates that in an ideal world with zero transaction costs, private parties’ decisions conduce the adjudication process to allocate legal entitlements efficiently. However, despite its formal coherence, this line of argument excessively simplifies the functioning of the adjudication process. Priest’s model disregards too much of the essence of the adjudicatory process.

---

846 An example might help clarify this point. Imagine a set of 200 disputes to be litigated in period 1. Assume that judges decide randomly with respect to efficiency, that is, there is 0.5 probability that the output of litigation is either efficient or inefficient. Thus, litigation generates 100 efficient rules and 100 inefficient rules. The ratio between efficient and inefficient rules is 1. If we assume that (out of this set of half-efficient half-inefficient 200 rules) 100 rules are re-litigated, at the beginning of period 2, the set of rules is composed of 100 contested rules and 100 not contested rules. Since, for the reason explained above, the rate of litigation of the inefficient rules is higher than that of the efficient rules, the proportion of inefficient rules that are re-litigated will be higher than the proportion of re-litigated efficient rules. Hence, for example, of the contested 100 rules, 60 will be inefficient rules while 40 will be efficient rules. The output of the second round of litigation—given the assumed attitude by judges toward efficiency—will be of 50 inefficient rules and 50 efficient rules. Now at the end of period 2, the set of uncontested rules will be composed of the sum of the initial 100 uncontested plus the output of the second round of litigation. The initial set of 100 uncontested rules was composed of 40 inefficient rules and 60 efficient rules; the output of the second round of litigation is given by 50 inefficient rule and 50 efficient rules. Thus, at the end of period 2, we will have 90 inefficient rules and 110 efficient rules. The ratio between the inefficient and the efficient rules is approximately 1.22. The composition of the set of rules has changed: the number of efficient rules had increased and the number of inefficient rules has decreased.

847 In this sense, Priest’s model provides a useful starting point for the economic analysis of judicial lawmaking.

848 Other implicit assumptions in Priest’s model are worth emphasizing because they are crucial in verifying the hypothesis of a demand-driven thrust toward efficiency: (a1) settlement is cheaper than litigation; (a2) judges are subject to the principle of stare decisis (i.e., judges should conform to past legal precedents); (a3) judges’ preferences for efficiency are exogenous to the model, as are their incentives and informational constraints; (a4) other characteristics of the legal order are exogenous to the model. These simplified assumptions allow us to focus on the demand-driven forces that affect the characteristics
Furthermore, Priest’s explanation lacks compelling reasons for the direction of the evolutionary mechanism toward efficiency. The results of Priest’s model can be easily reversed when a realistic account of the settlement-trial decision is introduced in the model. For example, within the same evolutionary framework, it might be convincingly argued that efficient rules are more likely to be litigated than inefficient rules. Parties’ settlement/trial decisions are influenced by the relative expected benefit from settlement and trial. In a system based on the adherence to past decisions, the settlement/trial expected benefit ratio is influenced by the estimated probability of winning the case. If the existing rule is inefficient, it means that cases litigated under this rule are more likely to be decided inefficiently. Parties that would benefit from an efficient decision will anticipate the probability of a negative outcome and will be less prone to go to trial, thereby generating a tendency toward less litigation of inefficient rules. Less litigation would entail a lower probability of the inefficient rule being changed, thereby strengthening inefficient legal precedents and discouraging efficiency-seeking litigation. In summary, the initial balance of precedents in favour of inefficiency generates a tendency that reinforces inefficiency. This does not demonstrate that Priest is wrong, but shows that a degree of the substantive inefficiency of legal rules may generate opposite effects on the frequency of litigation and on the evolution of judge-made law.

In conclusion, whether there is a causal connection between the size of the surplus and the likelihood of litigation remains an empirical question. The aprioristic analysis of the logic underlying individuals’ choices (on which Priest’s explanation rests) does not provide decisive arguments demonstrating a clear tendency toward efficiency.

---

of judge-made law. In particular, the analysis aims to isolate the independent thrust toward efficiency determined by the settlement-trial decisions of individuals.

849 This point is made by Landes and Posner, “Adjudication” supra note 841 at 70-74 and Cooter and Kornhauser, “Can Litigation Improve” supra note 844 at 155.
7.1.1.1.2. Inefficient rules are more frequently litigated when parties have future stakes in the dispute.

Priest’s explanation focuses entirely on present stakes. However, in many cases, parties who expect to be involved in future, similar disputes have an interest in establishing precedents that ensure them of a stream of future benefits. In other words, parties have “future stakes” in the dispute and place value on the “precedential effect” of the adjudication outcome. According to Rubin, when both litigators have future stakes in the dispute, the precedent will evolve toward efficiency.

Rubin discusses the example of a dispute between the injurer and the victim of an accident, where the injurer is the efficient bearer of risk. In this case, the efficient rule places the accident liability on the injurer. In the case of the inefficient allocation of liability, the degree of substantive inefficiency corresponds to the avoidance-cost differential between parties. Rubin shows that the inefficient allocation of the avoidance cost-differential provides incentives to both parties to litigate until an efficient precedent is established. In fact, as the cost-differential increases, the distance between a plaintiff’s expected return from litigation and a defendant’s expected loss become larger. In contrast, when the precedent allocates liability to the party whose costs are lower, the cost differential will incentivize parties to settle the dispute out of court. In the latter case, he assumes that the precedential effect of decisions does not enter in measuring parties’ benefits from litigation.

To clarify, both Priest and Rubin contend that (i) inefficient rules will be litigated more frequently, and (ii) efficient rules will persist uncontested. Priest and Rubin depart in the source of litigation-frequency differential. While Priest traces the differential to the impact that inefficient rules exert on private parties’ cost functions, Rubin focuses on the desire of repeated players to establish a favourable precedent governing their future transactions. Thus, while in Priest’s model the substantive inefficiency of legal rules determines alone a higher rate of litigation, in Rubin’s model it is the interest of both parties to establish a favourable precedent to generate the incentive to litigate inefficient rules before the courts.

According to standard law and economics, efficient rules are those that place liability on the person who can prevent future accidents at minimal cost, while the rules placing liability on the person whose avoidance costs are higher are the inefficient ones.

Assume two parties A and B, respectively, the tort-feasor and the victim of an occurred accident. Assume that B is the cheaper cost-avoidance so that efficiency would require him or her to be liable. The costs of the accident include the amount of damage X to be paid if the counterpart wins; the total cost of avoidance of future accidents (the “future stakes”), respectively, $T_a$ or $T_b$, and the cost of litigation C. The probability that B wins is equal to R, and the probability that A wins is equal to (1-R). Assume that B is the party whose avoidance cost is lower, and assume that the existing rule is inefficient because it places liability on A. In the assumed situation, the avoidance-cost differential ($T_a - T_b$) measures the substantive inefficiency of the legal rule.
the plaintiff’s minimum settlement offer does not exceed the defendant’s maximum settlement offer; consequently, litigation will not occur.

Based on this premise, Rubin examines three alternative cases. (i) When both parties have futures stakes in the dispute, the litigation process will generate a tendency toward the production of efficient legal rules. (ii) If only one party is interested in establishing a favourable precedent, then the party with future stakes in the dispute will have an incentive to litigate until the inefficient rule or efficient rule is overturned in his or her favour. In this case, the precedent will evolve in a direction favourable to the latter party, whether or not it corresponds to the efficient rule. (iii) If neither party is interested in precedent, then the existing rule will not be litigated, and it will remain in force whether it is efficient or inefficient.854

854 (1) Both parties have future stakes. A and B will decide to bring the dispute before the court based on their expected return from litigation. The expected return of A is equal to $R(-X) - (1-R)T_a - C$. This is because if B wins, A will pay damages; if A wins, he or she will save the avoidance costs; in any case, A will pay the cost of litigation C. The expected value of B is equal to $R(x) - (1-R)T_b - C$. This is because if B wins, A will pay him the damages; if A wins, B will pay the avoidance costs; in any case, B will pay C. Parties with a dispute will litigate when the expected gain of the plaintiff exceeds the expected loss of the defendant854. Thus, in this particular example, litigation will occur when $(1-R)(T_a - T_b) > 2C$. This means that litigation will occur when the avoidance-cost differential weighted by the probability that A wins is more than twice as large as the costs of litigation. Three variables determine the probability that litigation will occur. The higher R (i.e. the more the precedent is clearly in favour of B) the less litigation is likely. The higher the cost of litigation, the less litigation is likely. Finally, and more importantly for our purposes, the higher the avoidance-cost differential, the more litigation is likely. As the degree of inefficiency increases, the distance between the minimum B’s settlement offer and the maximum A’s settlement offer will increase, thereby increasing the benefits from litigation854. Put simply, the more the rule is inefficient the more it is likely that parties will go to court. On the contrary, that if $T_b > T_a$, then the rule is efficient, as it places liability on the party whose cost is lower. Litigation will not occur because the avoidance-cost differential is equal to zero. In essence, when both parties have future stakes in the dispute, substantive inefficiency increases the likelihood of litigation, while efficient rules will not be litigated.

(ii) Only one party has future stakes. Let us now assume that A has no future interest in the dispute. The expected value of A is $R(-X) - C$. If B wins, A must pay damages; in any case, he will pay the costs of litigation. It can be assumed that he has no interest in the future. The expected value of B is $R(X) + R(T_b) - C$. This is because if B wins he will receive damages from A, and will save the avoidance-costs for the future; in any case, he will pay the costs of litigation. Thus, in this situation there is not a positive $(T_a - T_b)$. The distance between B’s minimum settlement-offer and A’s maximum settlement-offer is now very large; as it equals the amount of avoidance-costs that B would have to pay if a wins. B’s expected value exceeds A’s expected loss when $R(T_b) > 2C$. Thus, B will have incentive to litigate unless R is very small (precedents is clear in disfavour of B), or unless litigation is very costly. In this situation, B will continue to file suit until the courts overturn the precedent and rule in his favour. Under the new rule that holds A liable, when the accident will occur A will find worthwhile to pay damages rather than going to trial. In fact, under the new rule it is likely $R(-X) - C$ will be less than X.
For the purposes of the process efficiency analysis, Rubin’s model is useful in demonstrating that the balance of interests between the two sides of litigation affect the evolution of the common law. *Judge-made law is likely to evolve toward substantive efficiency when classes of repeat players have symmetric (i.e., opposite, but equally sized) stakes in the adjudication outcomes.*

The formative period of the common law provides a historical example of the evolution toward efficiency under conditions of symmetry and repeated games. During the nineteenth century, the cost of organizations were high enough to preclude the formation of interest groups: the primary actors in the adjudication process were individual interests acting independently, rather than organized interest groups. \(^{855}\) This reflected a structure of the economy that was not yet characterized by pervasive economies of scale, and where legal-economic interactions were between individuals acting under conditions of reciprocity and role reversibility. The structural features of the regulatory environments (stake-symmetry and reciprocity) favoured the evolution of judge-made law toward efficiency. In the subsequent century—with the advent of an economic system characterized by large-scale industrial entrepreneurs, and where the organization of interests became much less expensive—the nature of litigants changed significantly, and stake symmetry and role reversibility broke down. In this new contest, in many areas of common law, the legal change was driven by the increased activity of organized interest groups attempting to obtain rules favouring systematically one side at the expense of the other.

Rubin’s explanation suffers from the same limitations as Priest’s model does. First, the causal connection between the substantive inefficiency and the frequency of litigation is *assumed* rather than explained. Rubin seems to neglect the fact that inefficient rules are likely to discourage litigation because they are expected to generate negative adjudication

outcomes. Second, basic law and economic theory teaches that the scope of Coasean bargaining increases as the available cooperative surplus increases. Now when accident liability is allocated inefficiently, the size of the avoidance-cost differential between parties provides room for Coasean bargaining. The less frequently that rules are efficient, the higher the surplus from out-of-court bargaining and, consequently, the higher the incentive is to bargain. Therefore, to the extent that transaction costs are low enough, when rules are inefficient, parties will settle the dispute rather than go to trial. On the contrary, Rubin assumes that the costs of bargaining around the inefficient law are prohibitive—so parties opt for litigation—but that the costs of reaching a settlement agreement under the efficient rule are low enough (lower than the litigation costs) to induce parties to settle the dispute.

The Nirvana fallacy is easily discernable here. In Rubin’s assumption, Coasean bargaining is prohibitively expensive when parties seek to negotiate around the inefficient rule (so they choose to litigate) but perfectly workable under current inefficient rules (so they settle the dispute and decide not to go to court in order to save litigation costs). The magnitude of transaction costs changes “by assumption” depending on the efficiency of the current legal regime. However, there is no explanation for the prohibitive costs of bargaining around the law. Indeed, the factors generating the prohibitive transaction costs of negotiating around the law (e.g., number of parties, asymmetric information, and so on) are likely to affect similarly the settlement of negotiating under the law. As Cooter, Kornhauser and Lane observe:

[...] neither Rubin nor Priest offers any evidence as to why bargaining breaks down and the parties land in court when the rule of law is inefficient, and why the parties cooperate successfully and reach an out-of-court settlement when the rule of law is efficient.

---

857 This has been illustrated by Cooter and Kornhauser, “Can Litigation Improve” supra note 844 at 152.
858 See Cooter et al., “Liability Rules” supra note 856 at 372.
7.1.1.1.3. Inefficient Rules Are More Intensively Litigated than Efficient Rules Are.

Goodman has provided a third explanation for the evolutionary development of judge-made law. Litigants who benefit from efficient legal rules invest more resources in litigation than do those who benefit less; consequently, litigants who enjoy greater benefits from legal efficiency are more likely to influence the outcome of adjudication in their favour. As in Rubin’s model, future stakes in the dispute play a fundamental role in affecting parties’ incentives to litigate inefficient rules. However, unlike the preceding two models, Goodman does not assume any positive relationship between (current or future) stakes in the dispute and frequency of litigation. Parties’ economic stakes in the dispute influence the intensity, rather than the frequency, of litigation.

Goodman conceptualizes litigation as a capital investment. Individuals are willing to invest in litigation to the point when the expected return equals the costs. The greater the size of the costs deriving from substantive inefficiency, the greater the amount of resources that parties are willing to invest in order to win the case or to establish an efficient rule. The assumptions of the model are as follows: (a1) judges are influenced by litigants’ effort to influence the decision, where litigants’ efforts are measured in terms of litigation expenditures;\(^{859}\) (a2) the probability of winning the case is an increasing concave function of the amount of resources invested in litigation; (a3) the adjudication outcome is a function of the ratio between the amount of legal expenses incurred by the two litigants.\(^{860}\)

Based on these assumptions, Goodman demonstrates that if (i) courts are unbiased (i.e., do not have any attitude neither toward nor against efficiency), and (ii) the degree of inefficiency is high enough to provide incentives to invest resources in litigation, then the party with the greatest economic stake in the decision will always have a higher

---

\(^{859}\) The underlying assumption is that the level of resources that parties invest in litigation reflect only their own stakes in the adjudication outcome.

\(^{860}\) To clarify, assume that the plaintiff has economic stakes \(B_1\) and the defendant has economic stakes \(B_2\). If \(\frac{B_1}{B_2} = 1\) parties have the same probability of winning; if \(\frac{B_1}{B_2} > 1\) the plaintiff has more probability of winning; if \(\frac{B_1}{B_2} < 1\) the defendant has more probability of winning.
probability of winning. The fundamental supporting argument is that a *litigant’s willingness to invest in litigation is greater for litigants that are negatively affected by inefficient law compared to individuals who seek to retain the inefficient law*. This follows from these assumptions: since individual willingness to invest is a positive function of the degree of substantive inefficiency, litigators seeking to overturn inefficient law have a stronger will to invest than those seeking to maintain the inefficient law do. This has an evolutionary implication: *it will be more likely that inefficient rules will be overturned because of litigation, than for efficient rules to be preserved*. Crucially, if the ratio of *private* costs and benefits from litigation is equal to the ratio of *social* costs and benefits, an efficient precedent will be established.

Goodman’s model usefully suggests a causal sequence between (i) the stakes in the dispute, (ii) the amount of resources invested in litigation, and (iii) the efficiency of the adjudication outcome. However, similar to Priest’s and Rubin’s models, Goodman’s explanation does not provide compelling arguments to demonstrate the *direction* toward the efficiency of the evolutionary mechanism. The litigation process is supposed to reach a Pareto optimal equilibrium when the ratio of private and social benefits coincides. Nevertheless, the model does not identify the circumstances under which this fundamental condition occurs. Therefore, there is no clue regarding *when* the causal sequence identified by the model is likely to lead to efficiency. Goodman is aware of this point: He emphasizes that the free rider problem may significantly affect the incentives of individuals to invest socially optimal amounts of resources to overturn inefficient legal precedents.

### 7.1.1.2. The Information Theory

Let me now turn to the information theory of judge made law. The expected value attributed by parties to a trial is influenced by the set of available information. In particular, the availability of private information on the adjudication outcome affects the

---

861 Recall that by assumption, judges’ decision-making is influenced by the level of litigation expenses, which in turn depends on the degree of legal inefficiency.

862 Goodman is aware of this point. A section of the paper devoted to illustrating the “Public Good Problem” emphasizes the importance of the free rider problem, which may significantly affect the structure of incentives.
subjectively perceived probabilities of victory in the litigation of plaintiff and defendant, thereby influencing an individuals’ decision to bring suit.

7.1.1.2.1. Symmetric Information

Because primary determinants of the relevant set of information are the *content* and the *clarity* of the law, the information approach to litigation identifies “uncertainty” and “information” as the two main determinants of an individuals’ decision to bring suit. Parties will bring suit when a high level of uncertainty prevents them from estimating with sufficient precision the costs and benefits of settlement and trial. What counts is the level of uncertainty relative to the size of the other variables affecting the litigation rate, including (i) the relative cost of trial and settlement, and (ii) the size of the stakes in the dispute. Let me clarify this important point by considering two opposite cases.

(i) *Stakes are constant.* Let me assume that the expected value of judgment is constant and exogenous. A *high* trial-settlement cost differential encourages parties to settle the dispute and save on litigation costs. In these situations, parties will decide to litigate, despite the high costs, only when the difference between their estimates of success is large enough that the differential between the plaintiff’s settlement demand and the defendant settlement’s offer exceeds the difference between the costs of settlement and the costs of litigation. Thus, holding stakes constant, *as litigation costs increase, the degree of legal uncertainty should increase for litigation to occur.* Disputes that can be decided with some certainty will be settled, not litigated, because the divergence between parties’ expectations will not be sufficiently large to generate litigation.

(ii) *Trial-settlement cost differential is constant.* Let me now focus on the variation in the size of the stakes and conversely assume that the difference between litigation and settlement costs is constant. *As the size of the stakes in a dispute increases, the difference between parties’ estimates of victory, which is sufficient to encourage litigation, decreases.* When the stakes in a dispute are *high*, a small divergence in the difference between the subjective probability of success is sufficient to generate a significant differential between the plaintiff’s expected gain and the defendant’s expected loss from trial. Conversely, when the interests at stake are *small*, the difference between parties’
estimates of success, which is necessary to generate litigation, will be large. The table in Figure 27 summarizes the two situations

**Figure 27 Settlement-Trial Decision**

<table>
<thead>
<tr>
<th>(1) Settlement-Trial Cost Differential</th>
<th>(2) Size of Stakes</th>
<th>(3) Difference in parties’ estimate of success</th>
<th>Settlement-Trial Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>Large</td>
<td>Small</td>
<td>Litigation</td>
</tr>
<tr>
<td>Large</td>
<td>Small</td>
<td>Large</td>
<td>Litigation</td>
</tr>
</tbody>
</table>

It is noteworthy that from this perspective, the degree of inefficiency of existing legal precedents explains neither the rate of litigation nor the evolutionary direction of judge-made law. *The differential between parties’ expectations of victory as influenced by the degree of determinacy of the law generates litigation.*

7.1.1.2.2. Asymmetric Information

The preceding considerations are based on the implicit assumption that parties have symmetric information about litigation outcomes. However, the presence of information asymmetries among litigants plays an important role in shaping the evolution of judge-made law.\(^{863}\) Information asymmetries produce two separate effects. First, they affect parties’ settlement-trial decisions\(^{864}\) because private information about the adjudication outcome influences the estimate of the probability of victory. In this respect, asymmetry of information can either facilitate or hamper the reaching of a settlement agreement, depending on the circumstance of the case.\(^{865}\) Second, the informational advantage of one

---

864 See Bebchuk, “Litigation” supra note 863 at 409.
865 Imagine a dispute where one party (the “informed” party) has an informational advantage over his or her opponent (the “uninformed” party). Let me assume that the defendant has a low probability of winning the case. In this case, the minimum settlement offer that the informed defendant is willing to
of the litigants exerts a significant influence on the *evolution* of legal precedent. The latter effect warrants some brief clarification.

Information theory predicts that in the presence of permanent asymmetric information between plaintiffs and defendants, the evolution of legal precedents embodies the information of the *informed* parties with the more *meritorious* claim.\(^{866}\) This is explained because the informed party, knowing it has a meritorious claim, has the highest perceived probability of winning the case, and thus the highest propensity to litigate. This enables us to predict, for example, that “innocent” doctors are likely to have more influence in the evolution of medical practices. This can be explained by the fact that negligence standards in medical malpractice are defined by medical customs, which gives a large informational advantage to medical defendants over injured plaintiffs. Among doctors, those having a more meritorious claim are likely to have the highest win rate because of their greater propensity to litigate. In fact, in the area of medical malpractice, the plaintiff win rates are significantly below 50%.\(^{867}\)

It is worth emphasizing that while both “stake theory” and “information theory” emphasize the role of informational asymmetries in shaping the evolution of judge-made law, the two approaches may lead to contrasting conclusions. In fact, stake asymmetry between parties does not necessarily push the evolution of legal precedent in the same direction as informational asymmetry. For example, in product liability litigation, in general, defendants have greater stakes in the dispute than the plaintiffs do. Stake theory would predict greater win rates for defendants. However, according to empirical observation, in the category of product liability, there is a significant differential between the plaintiffs’ win rates in *tort* versus *contract* actions.\(^{868}\) As Hilton observes, this differential is little explained by stake theory, but it is consistent with information

---


\(^{867}\) *Ibid.*

The win rate differential does not reflect the stake differential, as stake theory would predict. Instead, the observed pattern of win rates appears to be consistent with the information theory. In product liability, each side has private information concerning his or her own conduct, which is relevant in the trial. Now in *contract* actions the stake differential in favour of the defendant is partially offset by the private information of the plaintiffs. This is because in contract actions, the conduct of both parties (with respect to both offer and acceptance) is relevant in establishing contractual liability. By contrast, in *tort* actions, the focus of judicial determinations is on the defendants’ conduct. Thus, in tort actions the information asymmetry reinforces (rather than offsets) stake asymmetry, thereby providing a possible explanation of the lower plaintiff’s win rate in tort product liability actions.

### 7.1.1.3. The Public Good Problem

As previously emphasized, demand-side models of litigation conceptualize the adjudication outcome as a private good. However, although there is undoubtedly a degree of economic rivalry between the parties to a legal dispute, it must be recognized that judicial decisions exhibit strong public goods aspects, which are related to the joint-consumption interaction among the beneficiaries of legal precedents. When a legal precedent has been established, individuals can hardly be excluded from enjoying the benefits of a favourable legal rule. From an evolutionary perspective, the non-excludability of legal precedents may cause free-rider problems resulting in the under-production of legal precedents. In essence, the “publicness” of legal precedents seriously undermines the efficiency of the evolutionary mechanism.

Consider, for example, accidents caused by defective products. In a specific lawsuit, the injured consumer seeks compensation for the damages suffered in the immediate accident. In this context, litigation can generate a positive regulatory externality in favour

---

870 One could say that adjudication is a private good in its retrospective dimension (i.e., the solution of a pending dispute), while it is a public good in its prospective dimension (i.e., the creation of legal precedent). When using the term “retrospective”, I have in mind the effect of the decision on the case before the courts; I do not refer to the “retroactive” application of the effects of a judicial decision to past similar cases. This is an important aspect, which I do not consider here.
871 See Landes and Posner, “Adjudication” supra note 841 at 41 and 50.
of all future consumers to the extent that consumers’ *aggregate* interests enter into the litigation process. To consider consumers in general as the party adverse to the manufacturer, the plaintiff should be able to bind other consumers to an agreement around the law. Clearly, in the absence of suitable institutional devices (e.g., class actions), the transaction costs of such negotiation among consumers would be prohibitive. Furthermore, the injured consumers’ willingness to invest in litigation would not reflect the willingness to invest in consumers as a “class”. The resulting collective action problem is likely to prevent litigation from achieving efficiency.

Historically, a clear example of how the public good problem might undermine the evolution of judge-made law toward efficiency is provided by the history of tobacco litigation. Until the early 1990s, for more than forty years, the tobacco industry “had proudly proclaimed its invincibility from product liability”. 872 The tobacco industry had been able to resist two “waves” of litigation without being substantially threatened by litigation: “it appeared that cigarette companies, unlike most product manufacturers, were effectively immune from regulation by tort law”. 873 This situation changed dramatically in the 1990s, when a series of lawsuits against tobacco manufacturers posed a substantial threat to the cigarette industry. A crucial element in strengthening the effectiveness of the litigation process against the tobacco industry was the increased coordination among injured consumers. 874 First, the lawsuits (seeking to recoup the Medicaid-related costs associated with cigarettes) were launched by *state* attorney generals. The *states*—not just individual consumers—were seeking to retrieve through the litigation process the value of an externality imposed on the collectivity by the tobacco industry. Second, the litigation launched by the states led to the production of relevant legal documentation available to others private litigants: plaintiffs’ attorneys could now base litigation on the same public documents. As Rabin observes: “By the late 1990s, a tobacco litigator could build a case against the industry on the voluminous document discovery in the state

---

874 Rabin, “Tobacco Litigation” see *supra* note 872 at 345-346.
health care cost recovery suits and the class action litigation”.875 A network externality to the benefit of tobacco litigators had been produced. Third, consumers were able to improve their efforts in coordinating litigation. For example, in 1999, The Tobacco Trial Lawyers Association was formed with the aim of facilitating efforts in networking and coordinating litigation for plaintiffs involved in tobacco lawsuits.

Another problem closely related to the publicness of legal precedents is generated by the divergence between private and social incentives to bring suit. In fact, the private costs and benefits of litigation do not necessarily reflect the social costs and benefits associated with the production of judge-made law. The difference between private and social costs and benefits may result in either an excessive or an inadequate (from a social standpoint) level of litigation.876 An inadequate level of litigation affects not only the evolution of judge-made law but also the effectiveness of the liability mechanism in inducing the desired behavioural effects. Namely, even if the existing rule is socially efficient (thus, there is no need to establish a legal precedent), the misalignment between private incentives and the social consequences of litigation might prevent efficient legal rules from generating the intended behavioural response because private parties do not adequately activate the litigation mechanism. For example, the liability regime for defective products becomes effective in increasing manufacturers’ level of precautions only above a certain critical threshold of per-capita damages that correspond to the incentive for individuals to bring suit. Hence, below that critical level of damages, manufacturers will not have incentive to invest in socially optimal levels of precaution. That is, manufacturers can afford to inflict small per capita stakes injuries on a large number of victims, since no individual has the private incentive to bring suit.

In conclusion, adjudication is plagued by two externality problems. The first is related to the joint-consumption interaction among the beneficiaries of legal precedents, which

875 Ibid. at 345
undermines the evolutionary mechanism of judge-made law. The second is related to the misalignment of private and public incentives to bring suit, which might result in inadequate or excessive litigation, thereby undermining the behavioural effectiveness of judge-made law.

7.1.2. Conclusions on Productive Efficiency

No a priori line of reasoning demonstrates the existence of an evolutionary tendency of judge-made law to the efficient production of legal rules. Demand-side explanations do not identify a mechanism that, given parties’ preferences, incentives and constraints, leads unambiguously to legal efficiency. Furthermore, many explanatory hypotheses that are based on the idea of a selection process determined by parties’ settlement-trial decisions need to be empirically tested.

Despite these limitations, the foregoing discussion enables us to identify the following characteristics, which a sound process-efficiency analysis should consider when assessing on a case-by-case basis the efficiency of judicial law-making: (i) the degree of the substantive inefficiency of existing legal rules; (ii) parties’ incentives in establishing favourable legal precedents; (iii) parties’ willingness to invest in litigation; (iv) the structure of parties’ available information; (v) the relative costs of settlement and litigation; (vi) the structure of the relative costs of organizing collective action for the different interests at stake. Combined, these factors are likely to influence the evolution of judge-made law.

7.1.2. The Supply Side

The major criticism of demand-side theories is that they conceive adjudication as a two-party game in which outcomes are primarily determined by individuals’ trial-settlement decisions. Judges play a passive role. In contrast, the idea of adjudication as a triadic form of law-making provides a better explanation of law-making through adjudication.
From this perspective, *supply-side explanations appear better suited to illuminating the institutional mechanism generating the efficiency advantages of adjudication.*

The following subsections examine the following: (i) the judges’ information problem; (ii) the doctrine of precedent as a law-creating mechanism; (iii) the degree of centralization of the adjudication system.

### 7.1.2.1. The Information Problem

Judges are confronted by severe information problems. They have limited *information* and are affected by bounded *rationality*, which makes it difficult for them to identify the efficient rule governing the case. The most important limitation of judicial decision-making is the impossibility for judges to compute relative prices and estimate marginal utilities. Austrian economic theorists have long identified this problem as “the economic calculation problem”, which unavoidably confronts centralized law-making bodies.

This section briefly summarizes some lines of critique developed by the Austrian school of law and economics, which, compared to other schools of scholarship, has better illuminated the nature of the economic calculation problem faced by judges. Furthermore, it briefly identifies the characteristics of the law-making process that help judges to deal with the information problem.

#### 7.1.2.1.1. The Austrian Critiques

From an economic standpoint, when the parties to a dyadic relationship decide to delegate the resolution of a controversy to a third party, they create the structural premises of the “economic calculation problem”. Essentially, by mandating a judge to adjudicate the controversy, they “*supplant a price system with centralized, non-price direction*”. In the absence of a price system, the decision regarding alternative competing uses of the contended resource is delegated to the judge, who must estimate...

---

877 As the discussion proceeds, it will become clearer that supply-side models include all the sequences of the triadic model, that is, (i) the selection process by private parties, (ii) the structure of judges’ decision-making, and (iii) the feedback effect of judicial decisions on private parties’ behaviour.

878 See *supra* notes 366 and 369

marginal utilities and compute relative prices to identify the efficient solution of the controversy. That is, *centralized decision-making is involved to mimic the market prices*. The Austrian school of economics has long emphasized that this computation problem is far beyond human capacity.

First, prices are very difficult to discern. Buchanan clearly points out that opportunity costs are *subjective* and therefore not susceptible to *ex post* estimate by an external observer. Individual decisions are based on opportunity cost as much as they are subjectively perceived at the moment immediately preceding the decision. The “choice-influencing costs” correspond to the individual’s subjective evaluation of the best available opportunity at the moment of choice. Once the choice is made, the cost as a subjective experience “vanishes.” Being inherently subjective, costs are inextricably tied to the moment of choice. This consideration has an extraordinarily important implication. If costs exist only as a subjective experience before the choice, it follows that after the choice, costs cannot be ascertained and, even more importantly, cannot be replicated. The *ex post* assessment of costs through objectively ascertained measures cannot capture the costs that really influenced the choice (i.e., the choice-influencing costs). The cost objectively measured *ex post* is a consequence of the choice (i.e., the “choice influenced cost”): the subjectively perceived costs that really influenced the choice remain unknown. It follows that any decision-making agency (e.g., judges, legislators, bureaucrats, and so on) called to measure the costs factor that influenced an individual choice is condemned to failure. Choice-influencing costs can be neither measured nor replicated.

Second, judges cannot reliably predict how people will react to changing relative prices, which actions they will undertake in response to these changes, and what the consequences of these actions will be. The economic process is inherently dynamic:

---

880 This problem is exacerbated by the increased centralization within the adjudication system.
884 *Ibid.* at 44.
when the ratio of marginal utility to price changes (because of unpredictable exogenous factors), the relative price changes; consequently, demand reacts through competitive adjustments so that prices change again. The same happens on the supply and demand side. When the ratio of marginal product to marginal costs changes, the relative costs change. Hence, supply reacts through competitive adjustments, thereby changing both production costs and prices. No centralized decision-making body is capable of reliably mimicking the complex dynamics of market prices on the supply side.

Third, the market is a highly decentralized process. In private markets, the knowledge of the circumstances of \textit{time} and \textit{place} is dispersed in the minds of actors throughout the economy. This decentralized information exacerbates the difficulties associated with the economic calculation problem and makes it impossible for judges to render legal judgments that efficiently direct market allocation. As the following discussion will clarify, the decentralization of adjudication mitigates, at least in part, the problem of dispersed information.

\subsection{Imperfect Decision-Making}

The economic calculation problem suggests that it is impossible for the courts to achieve efficiency because the amount of information needed to mimic market prices is excessive. Four structural characteristics of the adjudication process help judges deal with the information problems, as identified by the Austrian school: (i) the information provided through the litigation process; (ii) the coordination with other sources of law; (iii) the doctrine of precedent; (iv) the degree of decentralization. The central question is \textit{whether these informational mechanisms are likely to conduce judge-made law toward efficiency}. My answer is that although these mechanisms certainly help judges to deal with otherwise prohibitive information costs, they are unlikely to conduce the production of efficient legal rules. Obviously, efficiency varies significantly across different regulatory environments. However, some core tendencies in how these mechanisms affect productive efficiency can be identified.
Litigation. First, the litigation pressure on judges is “selective”, and judges learn only from cases brought before them by private parties. Thus, as Hadfield demonstrates, judges do not learn about the regulatory objects from a sample of cases that are representative of the full set of possible cases (i.e., a random sample). Instead, they see only a biased sample, which results from private parties’ settlement-trial decisions. It follows that judges do not have access to the information needed to develop a rule that is efficient in the full set of cases. In short, an “activity selection bias” hampers judges’ ability to move toward efficiency. Furthermore, it is worth noting that the litigants’ decisions in selecting the sample of cases available to judges are influenced by the existing rules that govern access to litigation. Thus, the structure of the adjudication process is decisive in determining the flow of information to judges.

Other Sources of Law. Certainly, under appropriate circumstances, political and technocratic processes might perform better than adjudication does from the point of view of the efficiency and the effectiveness of gathering the decision-relevant information. In these cases, legislatures and agencies might exempt judges from the burden of creating legal content by providing high levels of \textit{ex ante} specificity. This strategy of coordination among sources of law might significantly increase the process efficiency of the overall system. Consider, for example, the law governing judicial deference to agency statutory construction. Consider also the example of constitutional provisions to establish centralized bodies having the authority to produce legal rules in environments characterized by highly specialized technical knowledge.

Doctrine of Precedent and Polycentrism (Reminder). The remainder of this section discusses at length the informational advantages of the doctrine of precedent and the increasing polycentrism of the adjudication system.

\textsuperscript{885} Hadfield, “Bias” \textit{supra} note 6 at 585.
\textsuperscript{887} For example, as already known, when the demand for regulation exhibits a high degree of homogeneity, centralized \textit{ex ante} mechanisms might enable the law-making process to realize consistent cost-saving effects. In addition, in regulatory environments plagued with high uncertainty (e.g., regulation of unknown risk), centralized bureaucratic agencies might perform better than courts do in force of the higher specialist knowledge available to technocrats.
7.1.2.2. The Doctrine of Precedent

This subsection examines the doctrine of precedent as a law-producing mechanism and identifies its relative advantages in terms of productive efficiency. The doctrine of precedent plays a two-fold role in adjudication: it is both an informational constraint and a source of knowledge for judges. The two most important doctrines of precedents are the doctrines of *stare decisis* and *jurisprudentia constante*. Here the focus is on the practice of *stare decisis*. However, the analysis can be extended to the other doctrines of precedent.

7.1.2.2.1 The Reason-Based Decision-Making

In chapter 3, I briefly illustrated the antecedent-consequent structure of the legal norm. Here it is essential to understand how “factual predicate” and “consequent” are formed in the adjudication process, as opposed to the legislative process. In this respect, Grant Lamond proposes a reason-based model of precedent that illuminates some important features of the process that generates legal precedents. The model emphasizes that the principle of *stare decisis* requires judges to treat earlier cases “as correctly decided on their facts”. Namely, judges should decide cases before them “in a way that is consistent with the correctness of the earlier decision”. The criterion of correctness does

---


889 The Latin expression stare decisis is literally translated as “standing by previously undertaken decisions”. Under the stare decisis doctrine (in its latest historical development) previous decisions have a binding authority on future decisions; hence, past case precedents become a primary source of law, and judges should adhere to them when deciding current cases before them. In comparison, the doctrine of jurisprudence constante holds that past case precedent exerts a persuasive legal authority on future decisions to the extent that they have reached a certain level of uniformity. Legal precedence is not binding per se; instead, the authoritative force of past decisions is directly related to the degree of uniformity of previous case law. Simply put, the doctrine of jurisprudence constante holds that the source of authoritative force of previous legal decisions lies in the ability of legal principles “to mature into a prevailing line of precedent” (Parisi and Fon, “The Economics”, supra note 12 at 78). As the discussion proceeds, I will emphasize some differences between the doctrine of stare decisis and jurisprudence constante.


not simply refer to the “ratio” or the “rule” established in previous cases (as the
traditional view of precedent suggests). Correctness requires adherence to the balance of
reasons assessed by previous courts with respect to the facts of the precedent cases. I will
expand this crucial point:

Consider a precedent case $P_1$. This is defined by the set of decision-relevant specific
factual circumstances:

$$F_1 = \{a_1, b_1, c_1, d_1, e_1, f_1, g_1\}$$

and by the reasons grounding the decision resolving the case:

$$R_1 = \text{if } \{A, B, C\} \text{ then } K$$

where $R_1$ is the ratio grounding the decision; $A$, $B$ and $C$ are the factual categories that the
judge considers sufficient to apply the ratio to the case;\(^{892}\) $K$ is the conclusion following
the logical reasoning.

Now, it is crucial to identify with precision the content of the binding effect of the
precedent. With the decision of $P_1$ the judge establishes that in the context of $F_1$

(i) $A$, $B$ and $C$ provide a sufficient ground to conclude $K$, and

(ii) $D$, $E$, $F$ and $G$ do not defeat the ratio provided by $R_1$ as sufficient ground to

conclude $K$.

Simply stated, the judge proclaims that among the set of decision-relevant facts,\(^{893}\) (i)
some factual elements are sufficient to conclude $K$, and (ii) other elements are
insufficient to defeat the grounds provided by $R_1$. To understand the implications of this
point, it is useful to introduce the concepts of “first-order reasons” and “exclusionary
reasons”. First-order reasons are the grounds sufficient to conclude $K$. Exclusionary
reasons are the grounds sufficient to exclude other first-order reasons from defeating the
first order reasons that support $K$ in the decision. In our example, first-order reasons are

\(^{892}\) A, B and C are abstract descriptions including $a_1$, $b_1$, and $c_1$.

\(^{893}\) The relevance of factual elements is established on the basis of R.
the reasons binding the future judge to conclude K on the ground of R, when the factual categories \{A, B and C\} are present; exclusionary reasons are the reasons binding future judges not to defeat K if additional categories \{D, E, F, G\} are present in future cases.

Crucially, both first-order reasons and exclusionary reasons constrain future decisions, even if they do so in different ways. Future judges will be bound (i) to conclude K on the ground of \(R_1\) in the presence of the first-order conditions supported by \{A, B and C\}, and (ii) not to defeat \(R_1\) in the presence of the exclusionary reasons considered in the case \(P_1\) and associated with \{D, E, F, and G\}.

To understand the implications of the reason-based model, consider the alternative between following and distinguishing precedent cases. Judges distinguish a new case from a precedent similar case if they identify some decision-relevant factual elements that are not present in previous cases, which justify not following the established precedent. Distinguishing innovates the common law, in the sense that the judge produces a new rule based on a different ratio that is supported by different circumstantial elements. It is generally thought that the judge innovates when distinguishing a new case, whereas he or she merely confirms the existing law when following previous precedents. Unlike this traditional view of legal precedent, the reason-based model sheds lights on the functional symmetry between “following” and “distinguishing” in the creation of legal rules: judges produce rules not only when innovating, but also when confirming precedents under partially new circumstances.

To illustrate, suppose that a new case N is brought before the judge. The set of decision-relevant circumstances of the case is

\[ F_2 = \{a2, b2, c2, d2, e2, h2, l2\} \]

where \(F_2\) contains the first-order reasons A, B and C, which are sufficient to conclude K based on the precedent \(P_1\). It also contains D and E, which are insufficient to defeat A, B and C based on \(P_1\). In addition, \(F_2\) includes \(h2\) and \(l2\), which are not considered in \(P_1\). This implies that the judge must decide whether (i) \(h2\) and \(l2\) support first-order reasons \{H and L\} sufficient to defeat first-order reasons \{A, B and C\} or whether (2) \(h2\) and \(l2\)
are not sufficient to provide first-order reasons defeating \{A, B and C\}. In the former case, the judge distinguishes \(P_2\) from \(P_1\), and \(P_2\) is grounded on \(R_2 = \{H \text{ and } L\}\). In the latter case, \(\{H \text{ and } L\}\) will amount to judicial consideration as exclusionary reasons that are not sufficient to defeat \(R_1\) in the presence of \(\{A, B \text{ and } C\}\). It is crucial that in both cases the judge’s decision “changes” the law. If the judge distinguishes and creates a new precedent \(P_2\), then future judges will be bound in the presence of \(F_2\) to conclude \(K_2\) on the ground of \(R_2\). If the judge follows \(P_1\), the content of the law changes because \(\{H \text{ and } L\}\) are added to the set of factors the presence of which will not in future cases provide reasons sufficient to defeat the first-order reason provided by \(\{A, B \text{ and } C\}\).

To conclude, as the precedents accumulate, the number of factual elements covered by exclusionary reasons increases, which enriches the constraint on future judges. Every decision on similar cases determines a change in the existing law, no matter whether the judge distinguishes or follows existing precedents.

7.1.2.2.2 The Exclusionary Reasons of Precedents versus Statutes

The reason-based view of legal precedent illuminates important structural differences between precedents and statutory rules. The legal precedent is “a decision that is sufficient in the context of the case to reach a [legal] conclusion”;\(^{894}\) it produces consequences in contexts other than the instant case only to the extent that the facts in the context in which it was generated are present. The ratio contained in a statutory rule applies to the whole set of cases falling within its scope of application. In the language of the reason-based model, the scope of the exclusionary reasons covered by a legal precedent differs greatly from the scope of the exclusionary reasons embedded in a statutory rule.

Statutory rules “pre-empt” the decision on the entire set of cases falling within the generalization embedded in the factual predicate. That is, the exclusionary reasons provided by a statutory law are presumptive and general: it is presumed\(^ {895}\) that the legal

\(^{894}\) Grant, “Precedents”, supra note 890 at 18 (emphasis mine).

\(^{895}\) To be clear, in this context the qualification of the exclusionary rule provided by statutes as presumptive is opposed to the qualification of the exclusionary rule provided by precedents as absolute. That is, in the absence of legal precedents that identify and establish the relevant exclusionary rules, statutes are based on the presumption that the first-order reasons not included in the scope of application...
effect prescribed by the consequent applies to the whole set of cases falling within the class of cases identified by the factual predicate. In contrast, the exclusionary reason provided by a precedent is absolute and specific: it applies only to the specific factors that have been considered in precedent cases. In short, the difference in the scope of the exclusionary reasons between precedents and statutory rules explains much of the costs and benefits associated with the doctrine of stare decisis. The two following subsections elucidate this point.

7.1.2.2.3 Advantages of the Doctrine of Precedent

(i) Decision-Making Costs

Let me now examine the advantages (in terms of productive efficiency) associated with the doctrine of stare decisis. First, under the doctrine of stare decisis, the accumulation of legal precedents increases the stock of information available to future judges, thereby enabling them to reduce decision-making costs. The reason-based model of precedent illuminates the mechanism of the formation of the stock of information. The accumulation of cases increases the number of exclusionary reasons. This means that as case law develops, the set of potential first-order reasons that are excluded from the reason-balancing assessment becomes increasingly large. As later judges follow, the list of decision-relevant facts that are insufficient to defeat the ratio of the precedent increases; analogously, as later judges distinguish, the list of factual circumstances that provide new first-order reasons sufficient to support a legal decision increases. In either circumstance, the development of case law expands the scope of application of legal

---

896 I will not consider the internal foundations of stare decisis, that is, the question of how the regime of precedent benefits the judges, and why judges provide decisions that are consistent with legal precedents, even when this is not consistent with their legal policy preferences. On this point see Erin O’hara, “Social Constraint or Implicit Collusion-Toward a Game Theoretic Analysis of Stare Decisis” (1993) 24 Seton Hall L. Rev., 736 [O’hara, “Social Constraint”]; Eric Rasmusen, “Judicial Legitimacy as a Repeated Game” (1994) JL Econ and Org, 63 [Rasmusen, “Judicial Legitimacy”].
precedents, thereby adding further constitutive elements to substantive legal doctrines. It is above all an *ex post* production of legal content.

The economizing effect associated with reason-based decision-making is understood by comparing the practice of *stare decisis* with a hypothetical regime without precedents. In the absence of precedents, judges must articulate *ex novo* substantive legal doctrine in order to decide cases. This requires them to create the *ratio* that grounds the decision by identifying first order and exclusionary reasons. In our example, judges confronting for the first time a case such as $P_1$ must balance the reasons provided by the factual elements \{A, B, C, D, E, F, G\}, and they must identify (i) which of them can be regarded as first order reasons sufficient to conclude K, and (ii) which of them provide exclusionary reasons sufficient to block the practice of distinguishing. This balancing assessment can be extremely difficult for judges, which in many cases, are not in a position to identify the substantively correct legal rule.\(^{898}\)

The doctrine of precedent enable judges to rely on both first-order and exclusionary reasons established by precedents, thereby exempting them from assessing substantive legal issues and from balancing the reasons underlying the cases. Bluntly put, the principle of *stare decisis* enables judges to replace a substantial judgement with a less demanding *similarity* judgement: “[I]instead of rethinking [the substantive issues of the case] judges can ‘free-ride’ on the opinion of previous judges”, \(^{899}\) thereby saving a considerable amount of decision-making costs.

(ii) Specialization

The second informational advantage of *stare decisis* is that it allows for the specialization of judicial knowledge. As Macey emphasizes, the doctrine of precedent “permits judges to ‘trade’ information among one another, thereby enabling them to develop competence areas of comparative advantages”. \(^{900}\) Since the principle of *stare decisis* protects the precedent from the practice of future distinguishing, judges are authorized to free-ride on

---

\(^{898}\) See *supra* sub-section 7.1.2.1.


\(^{900}\) *Ibid.* at 95
the decisions of other judges in areas in which they have relative less expertise, and they can specialize in the areas where they feel more competent. Higher specialization of judicial knowledge tends to reduce the number of judicial errors and enhance the probability that judges will identify the efficient legal rules. In fact, when judges are uncertain about the correct result of a case, the use of *stare decisis* permits them to rely on the application of substantive legal doctrines made by “more specialized” judges in similar cases.

(iii) Outcome Correctness

A third informational advantage associated with the doctrine of precedent is usefully explained through the intuitions underlying the Condorcet Jury Theorem.\textsuperscript{901} Imagine a hypothetical jury faced with a decision that admits two answers, one of which is correct. Assume also that no member of the jury is an expert, but that a non-identifiable minority of its members knows the correct answer. Under these assumptions, the Jury Theorem posits that (if each member of the jury has at least a 50% chance of selecting the correct answer) the likelihood of the jury selecting the right answer increases as the size of the jury increases. This is easily explained by considering the following: if the response of those who ignore the right answer is evenly distributed between the correct and the incorrect answer, then the correct answer will be the one selected by the jury. Now if members of the jury decide independently without influencing each other (i.e., there are no information cascades), then the larger the size of the group, the more likely it is that the responses of those who do not know the correct answer will be *evenly* distributed between the correct and the incorrect answer. In other words, the larger the size of the group, the less likely it is that the responses of those who do not know the correct answer will coalesce in a single incorrect answer.

The intuition underlying the Jury Theorem enables us to identify an impersonal mechanism underlying the judicial supply of legal rules, which leads toward efficiency. Assume for a moment that judges decided *independently* of each other on cases based on similar facts. The Jury Theorem suggests that reliance on the output resulting from the

\textsuperscript{901} The theorem was first expressed by Marquis de Condorcet in his “Essay on the Application of Analysis to the Probability of Majority Decisions” (1785).
body of legal precedent provides valuable signals to judges regarding the efficiency of the adjudicating outcome in later cases. It might be easily objected that, as Stearns demonstrates, the practice of *stare decisis* introduces a mechanism of path dependence that contradicts the assumptions on which the Jury Theorem is based. Nonetheless, the intuition underlying the Jury Theorem proves useful in evaluating the confidence level of the adjudication outcome in contexts where judges do not expect their decisions to be treated as precedents (e.g., lower court decisions and, under certain conditions, civil law court decisions).

*(iv) Positive Regulatory Externalities*

The fourth advantage associated with the doctrine of *stare decisis* is that “it maximizes the public good aspect of judicial decision-making”. The doctrine of precedent maximizes the beneficial effects of judges’ decisions that are external to the pending dispute, which lowers the levels of legal uncertainty, provides predictability, and allows for better foreseeability of the legal consequences of people’s behaviour. It must be recognized that from the perspective of individual litigants, the principle of *stare decisis* raises the costs for private litigants that seek legal change (i.e., changing the law is more difficult when the law is embedded in binding legal precedent). However, these costs are likely to be offset by the economic advantages associated with increased legal predictability.

7.1.2.2.4. The Disadvantages of the Doctrine of Precedent

*(i) Replicability*

The application of a statutory rule requires cases to be *subsumable* within the scope of application of the rule as it is defined by the factual predicate. The subsumption of the case within the scope of the statute is enough to produce the legal effect prescribed by the consequent. By contrast, as noted earlier, the adjudication of a case under the doctrine of precedent entails a reason-*balancing* assessment. “Balancing reasons” is a more

---

903 Macey, “Stare Decisis” *supra* note 899 at 106
904 I will not consider the internal foundations of stare decisis, that is, the question of how the regime of precedent benefits the judges, and why judges provide decisions that are consistent with legal precedents, even when this is not consistent with their legal policy preferences. On this point see O’hara, “Social Constraint” *supra* note 896 and Rasmusen, “Judicial Legitimacy” *supra* note 896.
sophisticated mental process than “subsuming a concrete single case in a class of cases” is. This entails that decisions based on statues are more easily (i.e., at less cost) replicable than those based on a balancing assessment.

In addition to the foregoing consideration, the relative replicability of statutory versus jurisprudential rules depends upon various factors, including (a) the relative clarity of statutes versus precedents; (b) the relative difficulty of statutes and precedents of solving the conflicts with, respectively, previous enacted rules versus previous legal decisions; (c) the relative degree of “maturity” reached through the application over time of the enacted law versus case law, that is, the number and quality of past cases that have been resolved through subsumption in the statute or through balancing of reasons; (d) the degree of heterogeneity of the cases to be regulated. The first three elements (clarity, consistency with existing law, and maturity) require empirical assessment on a case-by-case basis. The degree of heterogeneity of cases poses the problematic economic trade-off between decision-making costs and maladaptation costs already identified in Chapter 4.

(ii) “Average” Efficiency

The doctrine of precedent introduces an ex ante component in judicial law-making that is detrimental to substantive outcome efficiency (although beneficial in many other respects). Namely, under the doctrine of precedent, the decisional perspective of judges

---

905 Grant, “Precedents”, supra note 890 at 7.

906 For example, a recently enacted statutory rule might pose more complicated interpretative issues than would a consolidated long-lasting jurisprudential principle applied over time in a long sequence of cases. In this case, the application process of the newly enacted statutory rule would likely be less replicable than the consolidated jurisprudential principle would be.

907 Heterogeneity entails new facts. New facts are those not covered by either first-order or exclusionary reasons embedded in previous decisions. The higher the number of new decision-relevant facts is, the higher the level of the decision-making costs of adjudication. This is because new factual elements entail new reasons to be balanced with the established binding reasons covered by precedents. In comparison, statutory rules are based on a general presumptive exclusionary rule that applies to all new circumstances not falling within the scope of the statute. Hence, statutory rules allow for significant cost-saving effects. However, the general presumption of reason-indefeasibility embedded in statutes entails high maladaptation costs. The higher the substantial inefficiency of the statutory rule with respect to the first-order reasons supported by heterogeneous factual elements, the higher is the maladaptation-costs level generated by the indefeasibility presumption.

In essence, the efficiency comparison of statue versus precedent (all else being equal) entails a trade-off between decision-making costs and maladaptation costs, which depends on the degree of heterogeneity of the legal demand. Case-by-case decision making entails higher decision-making costs, while statutorification allows for a considerable economizing effect. However, as the degree of heterogeneity increases, the costs-saving effect associated with the statue tends to be offset by the increase in maladaptation costs.
is that of the “average” case. The economic analysis of adjudication usually assumes that every judge attempts to identify the rule that is efficient with respect to the single case brought before her or him. However, as Kornhousen observes, judges are aware of the unavoidable uncertainty in the determination of some decision-relevant facts. Furthermore, judges know that this uncertainty is a potential source of judicial errors. For example, in a case of injuries caused by the use of hazardous substances, the judicial determination of the “riskiness” associated with the use of the substance in the specific context of the instant case might be subject to a wider error than the determination of other facts relevant to decision would be. Consequently, a judge who was aware of the complexity of the deliberation and of the probability of judicial error “might think it wiser to adopt the legal rule on the basis of the “average” [case]”. That is, given the uncertainty associated with ascertaining some complex factual elements, the judge might find it efficient “to decide the class of cases to which the single case belongs correctly at the cost of deciding the instant case wrongly”. Since the decision is binding upon future judges, which is the ex ante element of judicial decision-making, the “average-case-perspective” is an efficient error-minimizing strategy. Judges aware of the imperfections of judicial decision-making find it rational to spread the error costs across cases. In conclusion—putting aside all the other information problems identified above—since rules are applied to classes of similar cases, efficiency will only be achieved on average.

7.1.2.3. The Degree of Polycentrism

Institutional judicial-law-making designs vary significantly over time and across jurisdictions. This subsection aims to demonstrate that a high degree of polycentricism of the legal order accentuates the informational advantages of adjudication. That is, the beneficial effects of polycentrism are facilitated by decentralization.

---

908 Kornhousen, “An Economic Perspective”, supra note 886 at 73
909 In general, the assessment of risk is plagued with uncertainty.
911 See Hadfield, “Bias”, supra note 6 at 585.
7.1.2.3.1. “Grown Orders” versus “Made Orders”\textsuperscript{912}

For the limited purposes of this study and with strong simplification, two opposite organizational models can be identified. In the first model, which has long prevailed in the history of law, adjudication is a highly decentralized set of institutions and is part of a \textit{bottom-up} process of law-production.\textsuperscript{913} In this context, customs and private contracting practices are the primary sources of law, and adjudication is primarily seen as an enforcing mechanism of private legal orderings. The role of the judge is “to discover and make explicit the rule that is implicit in the practices, customs, and institutions of the people. His job is not to create the rule, but to discover it […] and apply it to the specific case before him”.\textsuperscript{914} In the second model, which prevails today but is relatively recent in the history of law, adjudication is carried out by a highly hierarchical and centralized set of institutions and is part of a \textit{top-down} process of law-production. In this context, statutes and other forms of enacted law are the primary sources of law, and adjudication is an enforcement mechanism connected to a centralized production of law designed for the pursuit of collective goals. Using Hayek’s well-known terminology, the first model describes the adjudicatory process as a “grown order” and the second as a “made order”\textsuperscript{915}.

Two structural features of grown orders are relevant to our discussion. First, grown orders are “polycentric”: private litigants are free to choose among a plurality of adjudicating authorities that compete with each other to provide legal rules that better satisfy the demand for fairness, consistency and predictability. In polycentric systems, as explained in Chapter 4, the greater facility of exit options for people subject to law triggers an inter-jurisdictional competition dynamic, which makes it more difficult for the lawmaker to impose inefficient legal rules. In essence, \textit{competition among jurisdictions limits the possibility of using the court system as a redistributive mechanism}. Second, in a

\begin{thebibliography}{9}
\bibitem{Hayek} See Hayek, Law Legislation and Liberty, \textit{supra} note 201 at 35-54. See also Grady, “Positive Theories”, \textit{supra} note 912.
\end{thebibliography}
polycentric system, past decisions do not produce binding effects on future decisions. The body of precedents has a persuasive authority: namely, it provides evidence supporting the existence of legal doctrines. Precedent is by no means a source of law. *Law is generated by customs and private legal orderings; judges reinforce the objectives of private law.*

Zywicki explains the historical transition of the common law system from a grown type to a made type of order.\(^{916}\) This analysis is relevant here because it provides a useful supply-side model of the evolution of the common law. Zywicki identifies four institutional variables affecting the efficiency of adjudication: (1) the doctrine of precedent; (2) the degree of competition on the supply-side of the adjudicatory system; (3) the judicial tendency to produce either default or mandatory rules; (4) the judges’ reliance on customs. The behaviour of these four institutional variables is crucial in determining the evolution of judge-made law toward efficiency.\(^{917}\)

### 7.1.2.3.2. Judge-Made Law as a Grown Order

In its formative era, the development of common law largely corresponded to the description of the development of a grown order. A plurality of national and local, royal and ecclesiastic, public and private courts competed against each other to expand their jurisdiction and maximize the number of cases heard\(^{918}\). Important factors ensuring competition among courts were that the fees paid by litigants mainly funded the salaries of judges, and the formal boundaries among jurisdictions were easily evaded by using fictions.\(^{919}\) The polycentric nature of the legal order generated a competitive process that provided strong incentives to the adjudicating authorities to be responsive to the demand for legal efficiency. In addition, in order to respond to their “customers”, judges had the

---


917 Importantly, while Zywicki’s focus is on the evolution of the common law, its analytical framework allows for generalizable conclusions on the adjudicatory process across jurisdictions.

918 These courts included three different common law courts: the King’s Bench, the Court of Common Pleas and the Exchequer Courts, as well as four types of courts external to the common law system: the ecclesiastical courts, the law merchant courts, the local courts and the Chancery Courts.

919 For example, Chancery Courts developed a number of fictions to evade the restrictions of common law. This aspect of the historical development of English law resembles that of Roman Law, where the Ius Praetorium developed a set of jurisdictional remedies to correct the rigid formalism of the Ius Civile.
incentive to rely largely on private customs and to provide default rules instead of mandatory rules.

These structural features (private funding, reliance on custom, and default rules) are coherent with the role of the adjudicatory process as a mechanism for private legal ordering. First, private financing ensures responsiveness to the demand for efficient legal rules. Second, judges’ reliance on custom rests on the underlying conception of law as the result of the judicial discovery process of the rules generated through individual voluntary interactions. Third, the default nature of the legal rules allows people the freedom of contracting around inefficient rules. In conclusion, a complex system of rival courts, with overlapping and competing jurisdictions, providing default legal rules prevalently relying on customs are the crucial institutional features that generated the tendency toward the efficiency of common law prior to the nineteenth century. Zywicki observes that a similar form of judicial competition occurred in the history of US common law system under the regime of Swift v. Tyson.920 This decision allowed a common law case to be brought either under state common law courts or under federal common law. The possibility of choosing between two alternative jurisdictional fora triggered a competition among courts with partially overlapping jurisdictions, which was in some respects similar to the English competitive system.

7.1.2.3.3. Judge-Made Law as a Made Order

In the nineteenth century, the institutional framework within which the common law had traditionally developed changed dramatically in both English and US systems. The non-hierarchical and decentralized organization of the sources of law was progressively replaced by a more centralized system, which changed profoundly the nature of the adjudicatory process, and gradually decreased the competition among judicial jurisdictions. To illustrate the change in the direction of constrained judicial competition, Zywicki discusses the importance (in the history of the US legal system) of the overruling of Swift made by Erie Railroad v. Tompkins. In the latter decision, the Supreme Court held that federal courts did not have the judicial power to create general federal common law when hearing state law claims under diversity jurisdiction. Hence, the adjudication

920 Swift v. Tyson 41 U.S. 1 (1842).
process gradually became a law-creating mechanism more and more inserted in a *centralized* institutional law-making framework.

Several observations on this shift toward centralization are in order. First, *the increasing centralization of the law-making system provided the premises for the development of the doctrine of precedent*. As a first consequence of the decision, judges acquired much more discretion over the production of legal rules. The demise of judicial-competition thwarted the constraining effect on judicial power, which had been ensured up to that time by the possibility of an exit option on the demand side of the law-making process. The resulting lack of constraints on the increasing power of judges generated the need for an alternative constraining mechanism, which explains the development of the practice of *stare decisis*. Second, concurrent to the centralization of legal orders, *the role of custom and private ordering has been progressively reduced*. This process culminated in the explosion of the “statutorification” of law in the last decades of the twentieth century.921 Third, *mandatory legal rules progressively supplanted default rules*. Mandatory legislative rules have increasingly reduced the possibility of contracting around inefficient rules by private parties. Similarly, judge-made law was progressively transformed into a set of mandatory rules limiting the freedom of contract. The history of the European civil law systems, which differs in many other respects, followed a similar evolution in the direction of constantly increasing *centralization* and * politicization* in the production of law.922

7.1.2.3.4. Efficiency and Polycentrism

The preceding historical considerations support the conclusion that the progressive centralization of the legal order entails a change in the structure and nature of the adjudication process from a “grown” order to a “made” order. In conducting a process efficiency analysis, it is important to identify the informational comparative advantages

---

922 In the continental European tradition, in the eighteenth century, the idea prevailed that judicial power could be limited by complete subjection to the political process. Two fundamental institutional arrangements were designed for this purpose: (i) the doctrine of the separation of power, and (ii) the “codification” of private law.
and disadvantages of these two alternative models.\textsuperscript{923} A useful way to address the issue is to focus on two features that characterize adjudication systems: (i) the sources of judicial information, (ii) the institutional constraints on judges’ power to create legal rules, and (iii) the underlying notion of legal efficiency. First, in decentralized grown orders, the main constraint on judicial power derives from the preferences of litigants: private parties react to judicial decisions by choosing what they perceive as the most efficient judicial authority.\textsuperscript{924} Second, in grown orders, the principal source of judicial information is the accumulated stock of non-binding legal precedents, which provides legal solutions generated by customs and private orderings. Legal solutions are constantly selected through a complex and unpredictable selective process at the societal level. Prevailing norms are those accepted more often than their alternatives by private actors. Third, efficiency is a matter of the predictability and coordination of individual expectations to enable individuals to make the maximum use of their local knowledge in the pursuit of their goals. Grown orders are based on the recognition of the impossibility for any third party to measure and predict the impact of human behaviour according to a set of external social functions (i.e., the economic calculation problem). Efficiency can only result from the actual consensus of individuals. The role of judges is to ensure predictability, thereby facilitating the interpersonal coordination and compatibility of choices made by separate individuals. Predictability is achieved through the provision of internally consistent legal rules and through the constant adaptation of the rules to changing regulatory environments and individual preferences. Rules are theoretical constructs resulting from a non-deliberate evolutionary mechanism generated at the societal level.

\textsuperscript{923} The discussion on centralized vs. decentralized law-making has already provided many useful insights. Here I do not repeat previous conclusions, but I specify them in the context of adjudication.

\textsuperscript{924} To clarify, the preferences of people play an important role in made orders. The difference between made orders and grown orders is the mechanism through which the law-making process is influenced and constrained. In grown orders, the preferences of people are signaled to lawmakers mainly through the “exit” from inefficient legal regimes. In contrast, in made orders, people’s preferences are transmitted to the lawmaker through the intermediation of centralized institutional mechanisms (e.g., the legislative process). Moreover, constraints to judicial law-making are provided by centralized mechanisms (e.g., doctrine of precedent, and judicial review) not provided by the structure of individuals’ preferences expressed through exit-options at the local level that generated inter-jurisdictional competition.
The made-order type of adjudication differs in all three respects.\textsuperscript{925} First, the constraint on judicial power comes from centralized institutional arrangements, including the following: the practice of \textit{stare decisis}; the monopolistic and hierarchical structure of the court system; and the subordination of legal precedents to enacted law.\textsuperscript{926} Second, the flow of \textit{information} follows a top-down direction. Law does not have the function of enabling individuals to make the maximum use of their local knowledge. Instead, law is understood as a mechanism \textit{deliberately} designed to orient the allocation of resources in the direction established through the definition of some collective criteria. \textit{Rules are conceived as “acts of will” rather than theoretical constructs.} From this perspective, the judges’ role is to \textit{reinforce} the objectives of the political order. The selection process of the efficient legal rules is not the result of a complex and unpredictable dynamic at the societal level. Instead, it is the result of a cumulative process involving the deliberate maximizing choices of judges. Third, legal efficiency is not the result of the actual consensus of individuals. Instead, it is approached through the maximization framework, that is, the maximization of some collective outcome. Judges are assumed capable of measuring individuals’ cost and utility functions and predicting individuals’ behavioural responses to legal rules.

Grown orders reflect greater \textit{consensus} and \textit{wisdom} than made orders do because judges are constrained by people’s preferences rather than by hierarchical institutional constraints. They reflect greater wisdom because legal rules have been tested over time in several different factual contexts and have been found to fulfill parties’ expectations. The absence of the binding effect of previous decisions, coupled with a competitive decentralized court system, results in the greater ability of adjudication to experiment and explore innovative legal solutions. Judges in a made order have more judicial power than judges in a grown order do because they are entrusted with the power to issue rulings and compel their enforcement on both current and future parties. However, this greater power does not translate in greater judicial \textit{knowledge} or in a greater ability to tailor decisions to

\textsuperscript{925} Made orders are based on the idea of legal centrism, that is, the idea that law is provided through centralized and highly formalized institutional structures generating a top-down creation of legal content. On legal centrism see, Ellickson, \textit{Order Without Law, supra} note 356 and Robert D. Cooter, “Against Legal Centrism” (1993) 81 Cal L Rev., 1, 417.

\textsuperscript{926} Which reflects the subjection of judges to the centralized political power.
local circumstances. In conclusion, *made orders exacerbate the informational problems of centralized collective decision-making processes*, as Austrian economic theorists identify them.927

### Figure 28 Grow Orders versus Made Orders

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Information</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grown Orders</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) Exit-Option By Litigants</td>
<td>(c) Individual Preferences</td>
<td>(e) Predictability</td>
</tr>
<tr>
<td>(b) Inter-Jurisdictional</td>
<td>(d) Non-Binding Precedents</td>
<td></td>
</tr>
<tr>
<td>Competition</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Made Orders</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Binding Precedents</td>
<td>4. Rules as act of will enforced</td>
<td>5. Outcome</td>
</tr>
<tr>
<td>2. Hierarchical Organization</td>
<td>by judges</td>
<td>Maximization</td>
</tr>
<tr>
<td>3. Enacted Law (e.g. Statutes, Constitutions)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 7.2. Agency Efficiency

A major attribute traditionally associated with the adjudication process is to make it difficult for legislatures to effectuate the redistribution of wealth in favour of organized interests groups at the expense of less organized, low per-capita stakes interest groups.928 This idea is based on assumptions that do not withstand scrutiny. As Einer R. Elhauge demonstrates,929 the imperfections of the political process identified in interest-group theory do not necessarily imply that the adjudication process produces better results.

---

927 The essence of these critiques is that collective decision-makers, including judges, do not have the ability to collect and assess the amount of information required to channel individuals' behaviour consciously according to any definition of desirable social values.


Simply put, the Nirvana fallacy underlies the common idea that political pressure from interests groups justifies an expanded judicial review.

Legal and economic analysis has provided convincing arguments to support the claim that significant rent-seeking pressures may plague the adjudication process as much as they affect the political process. However, the analysis of the supply and demand sides of adjudication provides some useful insights on the differences between the rent-seeking mechanisms plaguing adjudication and politics. This section examines these differences and identifies the structural features of adjudication, providing comparative advantages over politics in terms of agency efficiency.

7.2.1. The Demand Side

On the demand side, comparative advantages must be sought in the mechanisms regulating access to litigation and politics. The access to justice and politics entails overcoming a cost threshold in order to participate in the production process and influence the outcome. Litigants have to bear expenses to gain access to trial. Voters and interest groups have to expend great amounts of resources to elect their favourite

---


Recent (Following Rubin’s original insights) articles have offered a “bidding theory” of the evolution of the common law. Under the bidding model, common law moves in a direction that favours the parties that are best able to devote resources to litigate in favour of their preferred rules. Thus, even if Rule A is inefficient, courts may be driven to adopt it if its beneficiaries have an advantage relative to others in organizing and devoting resources to litigation; see, Jack Hirshleifer and Paul H. Rubin, Evolutionary Models In Economics And Law (Greenwich, Connecticut: Jai Press, 1982); M. J. Bailey and P. H. Rubin, “A Positive Theory Of Legal Change” (1994) 14 Int’l Rev L and Econ, 467; Paul H. Rubin, Christopher Currain and John F. Currain, “Litigation versus Legislation: Forum Shopping by Rent-Seekers” (2001) 107 Public Choice, 295 [“Rubin et al., “Litigation versus Legislation”]. For a related model, see Francesco Parisi, “Rent-Seeking Through Litigation: Adversarial And Inquisitorial Systems Compared” (2002) 22 Int’l Rev L and Econ, 2, 193. See also A. Farmer and P. Pecorino, “Legal Expenditure as a Rent-seeking Game” (1999) 100 Public Choice 271.


932 Demand-side theories explain, as previously noted, that judge-made law is largely influenced by the structure of stakes involved in litigation. Rubin and Goodman’s model identifies an evolutionary pressure in favour of litigants with larger future stakes, and greater willingness to invest resources in the litigation process. This model does not provide any argument supporting the idea that adjudication enjoys some comparative advantages in terms of agency efficiency compared to politics. The model instead proves that adjudication, as well as the political process, is susceptible to the influence of interest groups.
candidates to representative assemblies. The point is that the cost threshold of gaining access to the adjudication process is likely to be lower than the cost threshold of gaining access to the political process. This is, as Merrill suggests,\footnote{Merrill, “Does Public Choice Justify”, supra note 931.} a potential source of the comparative advantages of adjudication over politics.

It is useful to distinguish three alternative sets of interests: (i) Some interests might not be strong enough to overcome the threshold costs for accessing either the political or the adjudication process. For these groups, there is no comparative advantage from of adjudication versus politics. (ii) Other groups might be strong enough to gain access to both processes. In this case, as in the previous one, the cost differential between access to adjudication and politics does not entail any comparative advantage for adjudication. (iii) Finally, and more importantly, some sets of interests are not strong enough to gain access to the political process, but they can nonetheless sustain the costs required to gain access to the litigation process. For this latter set of interests, the adjudication process has an undoubtedly comparative advantage over the political process. The lower threshold for accessing the adjudication process enables these interests to have access to the law-making process, influence the adjudication outcome and foster legal change. Figure 29 provides a graphical illustration of this point.

The different level of the costs of access to adjudication and politics explains, at least to some extent, the reason that in the US the legal change in the context of civil rights was initially spurred by courts’ decisions rather than by statutory reforms. That is, the collective action problem plaguing civil rights advocacy groups did not allow them access to the political arena; however, these groups had enough organizational strength to engage in litigation and obtain legal change. A related example is the participation of NGOs in judicial proceedings before international courts. These groups are not strong enough to influence the political process at the international level. However, the lower cost of access to international courts gives them the possibility of exerting influence on the adjudication outcome and spurring the process of legal change.
As the foregoing considerations suggest, a key source of the comparative advantages of adjudication versus politics is the *relative* ability of interest groups to influence the judicial decision-making process.⁹³⁴ From this perspective, a careful empirical analysis of the distribution of influential power among interest groups with stakes in the regulatory environment is essential in conducting an efficiency analysis.

⁹³⁴ It is the differential ability of opposing interest groups to influence the judicial decision-making process that counts.
Here I use simplified version of the diagram developed in Merrill, “Does Public Choice Justify”, supra note 931 at 223.
7.2.2. The Supply Side

Law-making supply functions summarize the relationship between the input and output of law-making. Influence functions identify the relationship between the amount of resources invested in the attempt to influence the law-making process and the outcomes of law-making. Supply functions and influence functions are closely related. An analysis of their relationship illuminates the differential in rent-seeking opportunities between adjudication and politics. In the following three subsections, I demonstrate that because of the institutional features of adjudication, judges are more limited in rent-offering power than politicians are.

7.2.2.1. Rent-Offering Power and Influence Functions

Law-making supply functions are characterized by a positive first derivative coupled with a negative second derivative. This implies that there is a point beyond which both politics and adjudication become insensitive to the pressures of influence. They respond to external pressures only as far as a certain amount of influence expenditures; beyond that critical threshold, law-making supply functions become inelastic with respect to attempts to influence the outcome through additional expenditures.

Merrill identifies an important difference between adjudication and politics: “judicial supply curves become inelastic at much lower prices than legislative supply curves do”. This can be explained in two related but distinct ways. First, there is a limit to the possibility that both politicians and judges could affect the quality of the legal outcome in a direction favourable to influential interest groups. Politicians and judges are subject to institutional and technical constraints that limit their ability to supply legal change. However, the adjudication process reaches the limit to the capacity of supplying legal change long before the political process, where “before” means “at lower levels of influencing expenditures”. The concept of “rent-offering” power is useful to illustrate this point. Rent-offering power is the lawmaker’s ability to satisfy rent-seekers’ demands by

---

936 The diminishing return of influence expenditures in the adjudication process is a reasonable assumption. Goodman also assumes this, see Goodman, “An Economic Theory of Evolution”, supra note 844.
937 Merrill, “Does Public Choice Justify”, supra note 931 at 228 (Emphasis mine).
938 This depends on a number of institutional constraints that are both exogenous and endogenous to the law-making arena.
affecting the quality of the legal outcome in their preferred direction. An approximate measure of rent-offering power might be the monetary amount of rents that the lawmaker can offer to rent-seekers. For reasons that I will explain below, it is realistic to assume that the rent-offering power of politicians is likely to be much higher than the rent-offering power of courts (with the significant exception of centralized Constitutional Courts).\textsuperscript{939} From this perspective, the higher level of expenditures at which the supply curve of political law-making becomes inelastic reflects the greater rent-selling power of politicians versus judges.\textsuperscript{940}

The second explanation for the elasticity differential between political and judicial law-making is related to the influence function of interest groups. It is generally assumed in the literature that influence functions are characterized by a positive first derivative and a negative second derivative,\textsuperscript{941} that is, there is a level of expenditure beyond which the influence power of interest groups becomes inelastic. Because of the different institutional features of politics and adjudication, the limit beyond which influence expenditures become ineffective is reached in the judicial arena at much lower levels of expenditures than in the political arena. Consider, for example, the attempt by interest groups to influence either a judicial decision or legislative outcome. The amount of money that a powerful interest group can spend in a single trial to try to influence the decision is limited to the legal expenses, including litigation fees, lawyers’ compensation, the costs of private investigations, and so on. The maximum amount of legal expenses is likely in the vast majority of cases to be far lower than the maximum amount of resources that could be invested to exert influence on the political process through elected representatives. The difference between the amount of money that the adjudication process allows to be expended on influencing the outcome and the amount allowed by the political process explains the elasticity differential in the influence functions of politics and adjudication. To summarize, the elasticity differential between the judicial and


\textsuperscript{940} As I will emphasize below, the difference in selling power between courts and legislatures decreases as the relative level of centralization of the court increases or the relative level of centralization of the legislatures increases.

\textsuperscript{941} Goodman, “An Economic Theory of Evolution”, \textit{supra} note 844.
supply curves depends on (i) the different rent-offering powers of the lawmaker, and (ii) the different effectiveness of influence expenditures by interest groups.

This difference between politics and adjudication has two important implications in terms of agency costs. First, when the supply curves reach the point of inelasticity, the influence curves of opposing interest groups become identical. This means that when the law-making supply curve becomes inelastic, adjudication is unresponsive to any increase in the level of influence expenditures; no group is capable of enhancing its influence power.\(^\text{942}\) As explained in Chapter 4, when the influencing curves of two opposing interest groups are identical, the two influence functions neutralize each other. In essence, the inelasticity of the supply curves levels the playing field of interests groups, thereby neutralizing, or at least reducing, their relative advantage in terms of influence power.

Now the fact that the judicial supply curve reaches the point of inelasticity at lower expenditure levels represents a significant advantage of adjudication versus politics because the relative rent-seeking power of interest groups is neutralized at a lower level of influence expenditures. The adjudication process provides weaker interest groups the opportunity to enjoy the benefits of the playing-field-leveling effect at a much lower level of expenditures. This significantly reduces the difference in the influence power between stronger, more organized, interest groups and less strong, less organized, interest groups. Crucially, this leveling effect on the playing field enhances outcome efficiency because the increased number of interest groups playing on a level playing field reduces the size of the agency-outcome costs (i.e., the cost of the inefficient outcome generated by external pressures). In general, competition among interest groups reduces the joint influence power of groups on the law-making process.

Second, from the aggregate standpoint, when the relative influence power of pressure groups is neutralized, the amount of resources invested in influence expenditures constitutes a deadweight loss for society.\(^\text{943}\) Since the neutralizing effect (triggered by supply inelasticity) occurs at a much higher level of expenditures in the political arena,

\(^{942}\) I assume here that all groups have reached the level of influence expenditures at which the supply curve has become inelastic.

\(^{943}\) I will discuss this important aspect in greater detail in the next chapter, when I analyze the agency costs of political law-making.
the political process has the potential to generate higher deadweight losses in the form of influence expenditures for no extracted rents.

In conclusion, two institutional features of the adjudication process provide comparative advantages of adjudication over politics: (i) the rent-offering power of the lawmaker is generally lower in adjudication than in politics; (ii) courts become insensitive to external influence at lower levels of influence expenditures. The more limited scope for rent-seeking and rent offering allows for (a) reduced outcome-agency costs and (b) less rent-seeking-agency costs. To clarify, I do not argue that adjudication always neutralizes pressures from interest groups. In fact, some groups are strong enough to gain access to the adjudication arena but are unable to invest resources as far as the inelasticity point. These groups will be obviously disadvantaged compared to powerful groups that are able to spend as far as the inelasticity point. Figure 30 provides a graphical representation of the elasticity differential between judicial and political supply curves.

These comparative agency-efficiency advantages of adjudication might explain, at the least to some extent, cases where courts have had a leading role in promoting legal change, despite the fact that the regulatory environment was characterized by strong, well-organized, high per-capita stakes interest groups. These cases have been sufficiently prominent to be designated the academic nomenclature “regulation through litigation”. Consider for example the wave of tobacco litigation that emerged in the US during the 1990s, which successfully led the tobacco industry—for the first time in its history—to accept a settlement proposed by the states. In this regulatory area, litigation has led to a transformation of the liability regime of the tobacco industry, which has had profound implications for the overall character of the production of law concerning the industry. In

944 See, for example, W Kip Viscusi, ed., Regulation Through Litigation (Washington: Brookings Institution Press, 2002) (Viscusi, “Regulation Through Litigation”). The expression “regulation through litigation” should not be abused. Tort law has always been a form regulation. Basic law and economics teaches that the definition of liability rules by judges affects the ex ante incentives of economic actors. For example, manufacturers know that when they design products, they will be held liable under tort law if they choose an unreasonably dangerous design. The expected liability costs enter into the manufacturers’ costs function as much as the expected costs of future sanctions for violation of command-and-control rules do. What is new in the phenomenon (“regulation through litigation”) is not that litigation has exerted regulatory effects; instead, to some extent, it is (i) the use of “class action” by individuals as a mechanism to promote legal change; and (ii) that litigation has in various ways provided incentives to the other law-making actors to bring about legal change.
particular, it is important to emphasize here that in the case of tobacco industry, litigation has spurred legal change in an area that had long been the target of regulation by the federal government and by the Congress. That is, *litigation succeeded in bringing about legal change in an area where political and bureaucratic processes failed*. Other examples of regulation through litigation are gun litigation, lead paint litigation, and breast implant litigation. 945 These cases share common characteristics: (i) litigation has operated as a mechanism forcing regulatory changes (ii) in an area previously regulated by politics or bureaucracies (iii) in industry sectors characterized by the presence of concentrated, well-organized interests. I do not address the merit of the adjudication outcome in these cases; I simply emphasize that the agency-cost differential between adjudication and other law-making processes has enabled courts to be decisive in promoting legal change.

---

945 Viscusi, “Regulation Through Litigation”, supra note 944
Figure 30 Demand and Supply Curves in the Political and Adjudication Arena with Output Thresholds

946 Here I use a modified version of the diagram developed in Merrill, “Does Public Choice Justify”, supra note 931 at 227.
7.2.2.2. Precedent

In the previous subsection, I assumed that (i) the rent-offering power of the lawmaker is generally lower in adjudication than in politics, and (ii) judges are likely to become insensitive to external influence at lower levels of expenditures. The next step is to identify the institutional features of adjudication that justify these assumptions.

The different degrees of the (1) generality and (2) stability of the law-making outcome are a source of important, relative advantages and disadvantages of adjudication versus politics in terms of agency efficiency. In this section, I maintain that while the greater generality of legal precedent is a source of comparative advantages of adjudication over politics, the greater stability of legal precedents might counterbalance this advantage by providing incentives for greater rent-seeking.

7.2.2.2.1. Generality

First, I clarify my use of the concept of generality here. I define generality as the amplitude of the spectrum of interests that the lawmaker has in mind at the moment of the creation of the rule. The generality of the relevant interests is strictly related to the possibility that the people subject to the law will foresee *ex ante*—with relative certainty—their personal position with respect to the future application of the law. The focus on this aspect of generality is useful to distinguish between “narrow-interest-group rules” and “general rules”. Since interest group rules are focused on a specifically narrow set of interests, they allow for the *ex ante* identification of winner and losers. In contrast, general rules, which address the generality of people subject to law, make it more difficult to predict *ex ante* the winners and losers. As Crew and Twight assert, “[r]ules that are more general in nature engender greater uncertainty regarding their personal impact”. General rules tend to be more efficient because they reduce the

---

947 For clarity, in this subpart I will refer to the prospective content of the judicial outcome, that is, to the part of the decision that is binding for future judges, and not to the decision of the specific dispute solved by the decision of the judge.

948 One might think that judicial decisions, which are by nature context-specific, provide rules that are more particularistic than are legislative rules that are by nature addressed to a generality of a community’s members. This is certainly correct. However, here I look at such generality from a slightly different perspective.


incentives for interest groups to engage in rent-seeking activity. Rent-seekers prefer narrow interest group rules over general rules.

Having clarified the notion of generality, I demonstrate that “narrowly focused interest-groups rules are more characteristic of statutory law than of common law rule-making”.\textsuperscript{951} This is not because—as widely argued—adjudication is less subject to external rent-seeking than politics is. The reason that adjudication is likely to produce general rules is related to the difference in the nature of adjudication and legislative outcomes.\textsuperscript{952} As repeatedly emphasized, statutory rules are acts of will. Constraints on legislators’ will are provided by constitutional principles, which are broad in nature and allow for relatively large margins of interpretations. Apart from constitutional rules and principles, in general, the constraints on the legislator’s will are political in nature.\textsuperscript{953} For reasons discussed at length in chapter 5, the political process is naturally inclined toward the production of private divisible goods (at collective cost). Now given the political nature of constraints on the legislator’s will, legislators are likely to be pushed toward the production of narrowly focused interest-groups rules, that is, rules that provide divisible private benefits.

Unlike statutory rules, judicial rules are theoretical constructs.\textsuperscript{954} The constraint confronted by the judge when fashioning the content of the rule is logical rather than political. Logical constraints tend to be more severe than political constraints are. As illustrated above, the reasoning to support judicial decisions is constrained by the consistency requirement with respect to earlier cases. In addition, the whole set of procedural rules regulating the adjudication process (e.g., civil procedure and criminal procedure) are specifically designed to conduce logically consistent legal decisions instead of to produce acts of will. Furthermore, the binding effect of precedent, as previously noted, stems from judges’ balancing of reasons in the context of the facts of

\textsuperscript{951} Ibid. at 26 (Emphasis is mine).

\textsuperscript{952} In the same direction are the arguments provided by professor by Epstein, “Social Consequences” supra note 939 at 1718 (“My argument is that structural features limit what the manipulation of common law rules can achieve. The more focused and sustained methods of legislation and regulation are apt to have more dramatic effects than does the alteration of common law rules and thus will attract the primary efforts of those trying to use the law to promote their own interests” at 1718 [Emphasis is mine]).

\textsuperscript{953} (or other legal limitations deriving from international treaties).

\textsuperscript{954} To clarify, this does not obviously mean that political considerations do not inform judicial reasoning.
the case. Somewhat paradoxically, *the fact-specific nature of the case-by-case decision-making process induces courts to a higher degree of generality in the production of law than broader constitutional rules do to the will of politicians.*

Furthermore, logical constraints impose stricter limitations on the power of judges than political constraints do against the power of politicians. As Lamond emphasizes, “statutes and precedents differ […] because legislators enjoy far wider discretion in what they can do (and what considerations they can act on)” 955 This is all that judges can do in deciding which cases go “against previous cases”. In contrast, legislators are empowered by the majority rule (or other decision-making rules based on some measuring stick of consensus) to produce legislative rules within the large limits of constitutional principles. Epstein clarifies this point:

[...] most common law rules are not cast in class form; there is no easy one-to-one correspondence between a given rule and the advancement of a particular social class. The ability to work substantial transfers of wealth between social classes is [...] severely hampered by the demand for public justification by written opinion that lies at the heart of the common law process. [...] The systematic demand for formal generality and neutrality in the common law may not guarantee that the proper substantive result will be reached in any individual case, but it does provide one bulwark against the invidious application of legal rules. The easiest way to oppress the poor, or to confiscate from the rich, is by laws directed at the rich and poor as such; the clandestine use of formally neutral principles is a poor second choice, to be used only when political or constitutional obstacles block the direct route. 956

Furthermore, it must be emphasized that while legislative action can (and often does) trump judicial outcomes, judges (with the exception of Constitutional Courts) cannot trump, but instead are subject to, legislative outcomes. Epstein observes:

A second means to counter the redistributive effect of common law rules is to seek modification of the legal standard by legislation. The greater are the stakes, the more likely it is that the distributive gains will be challenged. The original gain must be discounted by the probable collective response to the common law decision. 957

In conclusion, the doctrine of precedent constrains judges’ decision-making, ensuring a greater generality of judicial rules compared to legislative rules. Although this seems

---

955 Grant, “Precedents” *supra* note 890 at 6.
957 *Ibid.* at 1720
counterintuitive, it follows from the different types of institutional constraints to which judges are subject, compared to politicians.

7.2.2.2.2. Stability

The foregoing considerations should be weighed against their potentially negative effects in terms of the agency costs of the greater stability that the doctrine of precedent confers on judge-made law. Zywicki\textsuperscript{958} clarifies this aspect with respect to the doctrine of \textit{stare decisis}. The expected value of the rent extracted from a favourable legal outcome by an interest group is influenced by the following: (i) the size of the rent in a specific period; and (ii) the number of periods the favourable outcome is likely to remain in force. An interest group is willing to invest in rent-seeking up to the amount equal to the expected value of the extractable rent. The sum invested in rent-seeking will increase as one of the two variables increases. Thus, the practice of \textit{stare decisis}, which increases the stability of the legal precedent, has the effect of increasing the expected value of the rent extractable from a favourable legal precedent. This generates three distinct cost effects.

First, the adherence to \textit{stare decisis} increases rent-seeking agency costs because interest groups are willing to invest more in rent-seeking as the expected value of future rent increases because of the greater stability of precedents. Second, because of the doctrine of precedent, both efficient and inefficient precedents become more stable. This might exacerbate the costs imposed by substantively inefficient judicial rules. Third, stronger incentives to rent-seeking might result in stronger pressures on judicial law-making, which might result in the greater production of inefficient precedents to the extent that judges are susceptible to influence by rent-seeking pressures. In summary, a doctrine of “strong” precedent might result in (i) greater investment in rent-seeking activities, (ii) exacerbated costs of substantively inefficient rules and (iii) increased production of inefficient legal rules. These agency costs might counterbalance the comparative advantage of law-making through adjudication deriving from the greater generality of legal precedents.

\textsuperscript{958} Zywicki, “The Rise and Fall” \textit{supra} note 916.
7.2.2.3. Polycentricity

Judges’ rent-offering power tends to increase as the level of centralization of the court system increases\(^959\). In general, the less polycentric a judicial system is, the more it tends to share the flaws in terms of the agency inefficiency of centralized political processes. Obviously, as already emphasized in Chapter 4, the competition among different jurisdictions does not necessarily result in a “race to the top”. Competition among courts might result in increased incentives for judges to provide highly inefficient legal precedents in order to attract litigation and expand their jurisdiction at the expense of competing judicial actors. The efficiency enhancing nature of inter-jurisdictional competition depends upon both (i) the characteristics of the institutional framework (e.g., degree of independence from political power, availability of class actions, and so on) and (ii) the type of interest groups that are active in the regulatory environment (e.g., the differential ability of groups to engage in effective litigation).

7.3. Social Choice Efficiency

This section illustrates the mechanisms through which adjudication ensures the stability and rationality of the collective outcome. Furthermore, it emphasizes that the problem of social choice takes a different form in adjudication than it does in other law-making processes. This is related to the different sources of legitimacy that characterize judicial law-making compared to other law-making processes.

7.3.1. Non-Cyclical Path-Dependence

Path-independence and non-cyclicality are desirable properties of an adjudication system because the judicial allocation of rights should not depend on the order in which cases are presented to, or decided by, judges, but on the intrinsic merit of alternative allocations of rights.\(^960\) Despite its desirability as a condition for aggregate rationality, the idea of path independence is in sharp contrast with the structural features of the adjudication process, which is by nature “path-dependent” and subject to “agenda influence”. Moreover, as I


will explain below, it must be considered that path-dependence produces great advantages in terms of cost-saving effects and outcome stability. Thus, the problem arises of how to reconcile the problems of avoiding agenda manipulation and outcome instability with the structural path-dependency of the adjudication process.

In understanding this point, it should be considered that the adjudication process poses the social choice problem in a very different manner than political decision-making does. In politics, path-independency is urged in order to avoid agenda manipulation and ensure the consistency of the outcome. In politics, path-independence is one condition for the consistency of the outcome. In contrast, in adjudication, path-dependency is justified by the need for the coherency of the decision-making process, and it is a condition for ensuring the reasonableness of the outcome.\(^{961}\) The point here is to distinguish between cyclical and non-cyclical path dependence.\(^{962}\) While the former impairs the rationality of the decision-making process, the latter generates economizing effects on both the demand side and the supply side of the law-making process, without undermining aggregate rationality.\(^{963}\) To understand this crucial point, I shall briefly illustrate how rationality, as a legitimizing criterion of the decision-making process, differs among adjudication, politics and private orderings.

The legitimacy of judicial-law-making is derived from giving reasons and being coherent with previous decisions. The quality of the legal reasoning supporting judicial decisions grounds the rationality and legitimacy of the adjudication process. In politics and private orderings,\(^ {964}\) the quality of reasoning does not play a crucial role. Chapman clarifies this point:


\(^{963}\) Kornhauser, “Modeling”, supra note 962 at 182 (“A process […] may depend on history without ever leading back to previously rejected alternatives”.)

\(^{964}\) See Chapman, “The Rational and the Reasonable”, supra note 961 at 43 (“[…] the quality of the reasoning provided for a legal decision is decisive for its authority in a way that is largely irrelevant to the authority of market and political choices”).
In politics [...] while the final vote may be preceded by much discussion and exchange of reasons, the authority of the vote is not affected by the quality of this discussion. Such discussion need not even have occurred at all. And in market transacting, so long as one has the property right in question, one need not justify, or defend with reasons, one’s willingness or reluctance to trade. What makes a political vote or market transaction ultimately authoritative is simply the fact that those different acts of choosing have occurred, not the cogency of the reasoning involved.965

Therefore, unlike in politics and private orderings, the demand for rationality in the adjudication process stems from conflicting judgements rather than from conflicting wills. Of course, there are contrasting wills underneath any judicial conflict, but they manifest themselves in the form of opposing judgments (i.e., opposing theoretical constructs). In essence, the rationality in adjudication is a “judgement-based” rationality, whereas rationality in politics and private orders is a “preference-based” rationality.966

Judges provide reasons to support their decisions independent of individuals’ preferences for collective orderings.

The problem of cyclical decision-making is generated by sequential path dependence, that is, the dependence of the outcome from the sequential order of decisions. This must be distinguished from evolutionary path dependence, which does not generate the cyclicality problem. Evolutionary path dependence is generated by the increasing returns from following prior decisions and produces a stabilizing effect on the adjudication outcome.967 Adjudication, as we already know, performs two functions: resolving disputes and making law. Sequential path dependence arises more frequently in the adjudication process of single cases by collegial courts; evolutionary path-dependence arises over time across cases and characterizes the production and evolution of judge-made law. Although cyclical path-dependence is likely to beset the adjudication function, it is less likely to affect the law-making function. Stated differently, the cycling problem tends to arise with greater probability in the context of majority voting within collegial adjudicatory bodies. In contrast, the law-making function, which is performed through an incremental, evolutionary, decision-making process, is generally not affected by cyclical path-dependence. In conclusion, judicial law-making is a (i) non-cycling (ii) path-dependant process.

966 See Kornhauser “Aggregate Rationality” supra note 639 at 12-18.
967 That generates the tendency to lock-in the outcome of the decision-making process.
7.3.2. The Stabilizing Features of Adjudication

Let me now identify the structural features of the adjudication process that preserve judge-made law from the problem of agenda manipulation and cyclical path-dependence. First, unlike legislatures, courts do not set their own agenda. As repeatedly emphasized, judges rule only in litigated cases. That is, the flow of cases is not decided by judges but by litigants’ settlement-trial decisions. Thus, there is little room for agenda manipulation by judges. Agenda manipulation is also difficult for private parties. Litigants have conflicting interests at stake, which reduce the scope for monopoly agenda control. Furthermore, litigants must meet justiciability requirements to activate the adjudication process (e.g., standing, actual case in controversy, and so on), which significantly reduces the possibility for private parties to gain control of the agenda. Thus, it is reasonable to conclude that the problem of agenda control does not plague the adjudication system. As Epstein asserts, “Control over the judicial process is divided between courts and litigants; no one group can dictate the case agenda.”

Second, in the model of reason-based decision-making a judge is never confronted with the option of reconsidering a rule that has been excluded by a prior decision. If the adherence to a precedent is strict (as described in a previous section of this chapter), a judge can apply a new rule only when new factual elements render the previous ruling inadequate or incomplete. However, in the latter case, no new ruling that is inconsistent with some previous decision could defeat the rationality of the decision-making. Each new decision results from new factual elements and hence does not reconsider a previously rejected rule. In short, there is no cycle because no rejected alternatives are reconsidered. Stern has illuminated this mechanism by explaining that stare decisis works within the judicial law-making process as “a cycle breaking rule that works to ensure that an option defeated in a prior cases cannot be brought back to undermine a later judicial outcome”.

---

968 Epstein, “Social Consequences” supra note 939 at 1719
969 Kornhauser, “Modeling” supra note 962 at 183.
970 Stearns and Zywicki, Public Choice, supra note 500 at 462.
Third, judicial law-making rests on a reason-based rationality, which is a source of comparative advantages over politics. While preference-based rationality (as in the political process) requires the identification of some consistent collective ordering of individuals’ preferences (which poses the cyclicality problem when aggregate preferences are intransitive), the rationality of judicial law-making is reason based, that is, it requires the identification of the reason that supports the judge’s decision to follow or distinguish precedents. Briefly, rationality in political law-making poses the problem of “maintaining some consistent collective ordering of the preferences of individuals”, whereas rationality in judicial law-making requires “maintaining transitivity in the application of the law to new situations”. This means that transitivity in adjudication is ensured by the rules of legal reasoning, independent of the full transitivity of the social preference orderings. When law-making is based on the practice of giving reasons and on following similar cases, the problems of cycling are avoided.

Fourth, unlike statutory rules, common law rules can be negotiated by private parties. Individuals have the right, in many cases, to contract around common law rules. This significantly reduces the cyclicality problem. Obviously, this comparative advantage will be greater in decentralized adjudicatory systems, where judicial rules are “means for solving private disputes rather than outcomes of collective decisions”.

7.3.3. Cyclical Path-Dependence

In the previous subsection, I identified the features of the adjudication process that promote transitivity and enhance outcome stability. It now must be recognized that judicial decision-making is not immune from the problem of cycling. First, law, economics and public choice scholars have long identified conditions under which the presence of a cyclical majority within collective adjudicating bodies poses problems for the stability of the judicial outcome. This point goes beyond the purposes of this discussion. Second, “the problem of intransitive judicial preferences can arise not only

---

972 Ibid. at 121.
973 Ibid.
974 Ibid.
975 Ibid. at 116.
within individual cases [...] but also across groups’ cases over time”. My chief concern here is with this second type of situation because the intransitivity of judicial preferences across cases can potentially affect the evolution of judge-made law. In this respect, however, although cycling across cases can arise, the institutional features of judicial law-making help judges to ensure transitivity in the application of the law and to preserve judge-made law from the problem of cycling.

Finally, it is worth emphasizing that the same institutional features that prevent the cyclicity problem also make it more difficult to manipulate the judicial agenda. This is undoubtedly the source of a significant comparative advantage of judge-made law versus statutory law. As discussed in the following chapter, although many institutional arrangements prevent cycling and induce stability within the political process, the problem of agenda manipulation is much more difficult to resolve in the political process.

7.4. Adaptive Efficiency

The analysis conducted thus far emphasizes that judge made law is an evolutionary path-dependent process. From the perspective of adaptive efficiency, evolutionary path dependence entails the risk of inefficient lock-in effects. This section examines the extent to which adjudication enjoys comparative advantages over alternative sources of law in minimizing the risks of entrenching inefficient legal regimes.

7.4.1. The Advantages of Adjudication

I identify the following advantages of adjudication versus politics: (i) the levelling-playing-field-effect associated with adjudication mitigates the lock-in effect associated to the doctrine of precedent; (ii) the incremental nature of judicial legal change reduces the decision-making costs, and provides some strategic political advantages; (iii) adjudication

---

976 Stearns and Zywicki, Public choice, supra note 500 at 460 (Emphasis is mine).
977 Adjudication and politics share the advantages of centralized processes over decentralized ones. As explained in Chapter 4, centralization enables the lawmaker to provide focal points around which individuals can coordinate their compliance behaviours, thereby overcoming the collective action associated with the transition to superior alternative legal regimes. Judges’ decisions might prove useful in promoting legal change by facilitating individuals’ simultaneous shift to a new superior legal regime.
might be advantageous for change-seeking actors that are a minority in the political process.

7.4.1.1. Level-Playing-Field Effect and Lock-in Effect

The agency-costs minimizing features of the judicial process (i.e. judges’ limited rent-offering power, and limited effectiveness of influence expenditures) entails a further comparative advantage of adjudication versus politics in terms of adaptive efficiency. The agency-costs-reducing effects mitigates the risk of entrenching inferior legal regimes when political power is concentrated in a restricted number of interest groups having a vested interest in maintaining an inefficient legal regime. I have previously explained that the advantage of adjudication depends upon the structure of the costs of gaining access to litigation and influencing the legal outcome. In cases where adjudication allows for the access of interests groups that are not able to access the political process, adjudication operates as a vehicle of legal change. In the history of US, as emphasized earlier, tobacco reform, the progressive affirming of civil rights and the overcoming of the death penalty are examples of legal changes triggered by the use of the adjudication process.

As Gillette observes, “whether lock-in occurs more readily in courts or in legislatures will depend to a significant extent on which forum is more amenable to the legal entrepreneurs who seek to modify the status quo.” On the one hand, “parties that coalesce relatively easily might attempt to obtain legal change through legislation”. Thus, when legal entrepreneurs face low costs of conjoint action, lock-in is more likely to occur in judge-made law than in legislation. On the other hand, “individuals who have difficulty coalescing into groups because potential members are geographically diffuse or have insufficient individual stakes might find it costly to approach legislators relative to courts”. Thus, when legal entrepreneurs face high costs of conjoint action, lock-in is more likely to occur in the legislative arena than in the adjudication process. When well-entrenched interest groups preferences that are in favour of the status quo succeed in locking-in the legislative process, courts might provide a chance to obtain legal change for interest groups that advocate legal change.

---

979 Ibid at 828 (my emphasis).
The preceding consideration enables me to predict the following: *adjudication enjoys comparative advantage in terms of adaptive efficiency when the influencing power is concentrated in the hands of groups opposing the legal change*. On the contrary, when the dominating groups coincide with legal entrepreneurs, politics enjoys a comparative advantage over courts in terms of adaptive efficiency.980

### 7.4.1.2. Incremental Legal Change

A second significant advantage of adjudication derives from the reason-based nature of judicial decision-making. While legislation can be subject to sudden, dramatic changes, judges’ reason-based decision-making generates *incremental* legal change. Incremental decisional making is an efficient strategy in collective decision-making. As Shapiro observes, it “eases the path of collective decision-making”.981 Let me explain this point.

Incremental decision-making is anchored to the *status quo*. It proceeds (i) by experimenting with *marginal* changes, and (ii) by reacting to retroactive *feedbacks*. Judges are incremental decision-makers.982 They change the existing law only when it is incomplete or inadequate. They proceed by making incremental changes within the limitations imposed by the doctrine of precedent (i.e., marginal changes), and they act in response to the trail-settlement choices of private parties (i.e., they react to feedback). Crucially, the *doctrine of precedent ensures the incremental nature of judicial decision-making*.

---

980 The idea that different groups seek legal change in different institutional arenas is confirmed by Rubin et al., “Litigation versus Legislation”, *supra* note 930
982 Martin Shapiro illuminates this important point. Shapiro conceptualizes the practice of stare decisis as a strategy of incremental decision-making whereby judges can save on decision-making costs while concurrently keeping in step with social and economic legal changes. Shapiro explicitly draws on the incremental theory developed by Lindblom and March. The theory of incremental decision-making also has been developed in J G March and Herbert A Simon, *Organizations* (Oxford, England: Wiley, 1958) See the followings works by Martin Shapiro: “Toward a Theory of Stare Decisis” (1972) 1 J Legal Studies 125; *Supreme Court and Administrative Agencies* (New York, Free Press, 1968); and “Stability and Change” *supra* note 981 in which he supports the central claim of political jurisprudence that courts as much as legislatures and executives are political agencies. In fact, courts share the same incremental decision-making process of other political agencies. However, in subsequent articles (“Decentralized Decision Making in the Law of Torts” in S. Ulmer, ed., *Political Decision-Making* (New York: Van Nostrand, 1970), Shapiro emphasizes the advantages of incremental legal change generated by judicial law-making versus the non-incremental legal change produced by legislatures.
Judicial incremental legal change provides important advantages in terms of adaptive efficiency. First, it provides a strategy to economize in the prohibitive marginal information costs that would be entailed by the synoptic approach, which is otherwise postulated by rational choice theory. Second, “taking small steps and waiting for feedback before taking more steps reduces the risks generated by [the] uncertainties.” The strategy of proceeding through marginal changes and adjusting direction when negative feedback is received allows for an “error-correction-through-feedback” strategy, which significantly reduces the risk of errors. Third, incremental change protects individuals’ expectations and allows individuals to make an accurate estimation of the future costs of alternative courses of action. In this respect, the practice of stare decisis ensures the greater long-run predictability of judicial rules compared to legislative rules.

It is worth noting that judicial incrementalism entails a strategic political advantage for the lawmaker. As Shapiro observes, “incrementalism specifies stopping the decision-making process as soon as a solution is discovered that each of the decision-makers is willing to accept.” Similarly, the fact-specific and retrospective nature of judicial decisions enables judges to lower the political visibility of the legal change, thereby lowering the level of adaptive transaction costs. In fact, from an aggregate perspective, the immediate stakes involved in the adjudication process are low because the

---

983 Synoptic decision-making coincides in large part with the “rational decision-making” postulated by neoclassical economic theory. In synoptic decision-making, all the information on the facts relevant to the decision is gathered and evaluated in light of all the relevant goals. The goals are clearly identified and weighed according to the decision-maker’s values and preferences, which are clearly identified and prioritized. Every available, rationally conceivable, alternative policy is considered, and the consequences of any given alternative are weighed according to the identified values, preferences and goals. Finally, the “correct” policy alternative is identified and implemented. Incremental theory contends that synoptic decision-making entails an excessive informational burden for the decision-maker.

Incremental decision-making begins by considering the status quo, rather than assessing every policy alternative. Importantly, the status quo is generally maintained until the present policy fails, which triggers the need for some change and the search for alternatives. Only a restricted number of alternatives is examined, and only a limited range of consequences is weighed. In addition, the alternatives that are considered consist in marginal variations of the existing policy. Moreover, the alternatives that are considered are not necessarily better ones; instead, the considered alternatives are those available, keeping in mind the limited information and limited competence of the decision-maker. That is, the ends are adjusted to the available means, and therefore change as the available means change. A further important characteristic of incremental decision-making is that decisions are continuously adjusted to the feedback generated by the regulatory environment emanate after previous marginal changes.


985 Ibid.
adjudication outcome produces immediate effects only on the litigants.\(^{986}\) The real impact of the adjudication outcome becomes apparent only in the passage of time. By contrast, the forward-looking perspective of the legislature pushes decision-making in a non-incremental direction, which entails larger resistance by losers. The general, *prospective* (ex ante) nature of the enacted rules imposes a greater “dramatization” of the legal change. In politics, decision-making actors are usually constrained by much shorter times. Because the immediate stakes are higher, resistance by the losers is exacerbated, which is one reason that politicians usually confront higher resistance costs than judges do.

An example of the adaptive advantages of incremental legal change is the enactment of the Sherman Act in 1890. At that time, the political lawmaker was torn between the relatively small, but politically active, farmers who were seeking protection from larger, corporate competitors and powerful and prospective monopolistic industries. Under these conditions, it is generally recognized that the political lawmaker intended the courts to play the leading role in promoting competition and attacking monopolization. Politicians could not go too far in “levelling the playing-field” of the national economy against the interests of trusts and cartels. The immediate stakes were high, and the forward-looking perspective of the political process exacerbated the political conflict between the opposing interests. Clearly, the political lawmaker wanted the courts to make the difficult decisions by slowly developing a body of case law establishing and promoting free competition in the American market. Hence, at the time of approval of the law, politicians opted for the strategic allocation of the law-making power on the shoulder of the courts; they approved a “framework” statute in which the provisions were extremely sparse and important questions were not covered.

Despite its undoubtedly significant advantages, incremental judicial legal change might prove detrimental in terms of adaptive efficiency. In fact, *judicial law-making might excessively retard the production of law*. In this respect, many case studies have emphasized that reforms increasing the scope of judicial-review—with the aim of

\(^{986}\) Unless parties have strong future stakes in the dispute.
alleviating the shortcomings of the bureaucratic process—have produced unintended consequences in terms of the *ossification* of the rulemaking process. 987 For example, in the period from the late 1960s to the early 1980s, federal courts transformed the law of administrative procedure by expanding the scope of judicial oversight and control of agency action. 988 The shift of authority from agencies to courts resulted in many regulatory contexts that exacerbated the decision-making costs of the bureaucratic process to the point of policy paralysis. 989 For example, Pierce 990 provides strong arguments supporting the claim that the attitude of federal circuit courts toward agency rulemaking has played a significant role in leading the Federal Agency Regulatory Commission to abandon *systematic* approaches to policymaking in favour of *ad hoc* policymaking. In particular, in the electricity sector, the courts have hampered the creation of regulatory structures conducive to adequate investment. Similar conclusions have been reached with respect to the judicial review of the Occupational Safety and Health Administration 991 and National Highway Traffic Safety Administration regulations. 992

7.4.1.3. Inertia and Status Quo

Legislatures can decide *not* to act, thereby allowing the *status quo* to continue. In the legislative arena, groups favouring the *status quo* have the strategic option to exploit the *inertia* of the legislator and avoid political battles that entail the risk of loss. Unlike politics, adjudication is activated by the decision of interested parties, and unlike legislators, courts are institutionally obliged to decide the cases before them. Thus, when the groups favouring the *status quo* succeed in securing the *inertia* of the political

lawmaker, legal entrepreneurs can opt to engage in litigation in order to overcome the problem of lawmaker’s inertia.

7.4.2. The Disadvantages of Adjudication

The doctrine of precedent is explicitly designed to generate some degree of path-dependence on the adjudication process and to lock in legal rules. The lock-in effect can be considered either (i) as a source of outcome stability and reduction of decision-making costs or (ii) as an obstacle to legal change that exacerbates the costs of inefficient legal rules. Because the lock-in effect is endemic to any form of law-making, a comparative analysis must determine (i) whether the lock-in of inefficient legal rules is more likely to occur in adjudication than in other law-making processes and (ii) whether the costs of lock in might be outweighed by the benefits deriving from the doctrine of precedent (e.g., greater predictability, reduction of judicial uncertainty and so on). This subsection attempts to answer these questions.

7.4.2.1. The Doctrine of Precedent and the Lock-in Effect

The first issue begs the question of whether strict adherence to the doctrine of precedent is positive or negatively related to the amount of the adaptation costs. Let me again consider an example where courts react to a technological innovation in automobile security system by ruling that in cases of car accidents, the drivers of “non-adapted” cars are presumed negligent and thus liable. Non-adapted cars are those that do not adopt the most efficient technology. Recall the distinction between different categories of adaptation costs:

(i) adaptive transaction and information costs are the costs for individuals to appreciate the efficiency advantages of a superior pattern of behaviour and to coordinate a simultaneous, mutually advantageous shift to the superior alternative behaviour. In the example, these are the costs for people to appreciate the advantage of switching to the new technology and to coordinate a simultaneous switch;

(ii) resistance costs are the increased costs for people to maintain the old behaviour in violation of the law. The negligence rule established by judges increases the costs of maintaining the old technology;
(iii) *maladaptation* costs are costs generated by the substantive inefficiency of legal rules. In the example, these are the costs of a legal rule that does not prescribe the switch to a more socially efficient technology in order to minimize the costs of accidents; (iv) *adjustment* costs are costs individuals incur in transiting to the new legal regime and changing their behaviour as required by the law. In the example, these are the costs incurred by car drivers in buying and installing the new security system.

We now can compare the economic impact of *stare decisis* versus *non-stare decisis*. During the discussion, bear in mind that the socially desirable outcome is that people transit to the new technology in order to minimize the social cost of accidents (i.e., minimize the maladaptation costs of the old technology).

Let me first consider a world in which adaptive transaction costs are low. Let $t_0$ denote the time in which a change in the regulatory environment occurs, that is, the new, more efficient technology becomes available. In $t_0$, if the individual decides to adopt the new technology, he or she incurs adjustment costs; if he or she decides to maintain the old inefficient technology, he or she faces resistance costs. Standard microeconomic theory predicts that the individual chooses to change his or her technology if the present value of adjustment costs does not exceed the present value of resistance costs.

Under *no stare decisis* and low transaction costs, parties can identify the new socially efficient technology and decide to coordinate spontaneously to transit to the new technology. That is, legal rules may not have to change formally before the parties alter their behaviour. People recognize that is more efficient for them to change their behaviour and adopt the new technology without waiting for courts to establish, through the announcement of a new binding legal decision, a new legal regime. Under similar circumstances, individuals might incur adjustment costs spontaneously in period $t_0$. In contrast, under a regime of *stare decisis*, drivers might find it advantageous to delay making the transition until a court in the jurisdiction formally rejects the pre-existing precedent. Parties might prefer to wait to incur adjustment costs until they are certain that the law will change accordingly. In this case, individuals incur adjustment costs in period
In other words, in a world with low adaptive transaction costs, under a no stare decisis regime, people are willing to face adjustment costs and change their behaviour more quickly than they would be in a stare decisis regime. A quicker adaptation to the superior technology reduces the maladaptation costs: if the change in technology occurs in $t_0$, the maladaptation costs associated to $t_1$ are saved. Thus, under these assumptions no stare decisis seems to be more efficient than stare decisis is. Furthermore, it seems plausible to state that a regime of no stare decisis is likely to allow for a faster legal change than a legislation-induced legal change would be.993

In a world where adaptive information and transaction costs are high, it is more difficult for individuals to appreciate the efficiency advantages of the new technology and to coordinate a simultaneous shift to the new car-security system. Furthermore, because simultaneous legal change is difficult to coordinate, expected adjustment costs also increase because individuals who decide to adopt the new technology face the risk of a solitary migration. Under these circumstances, individuals are more likely to maintain the status quo until the court formally announces the new legal rule. For reasons previously clarified, the adjudication outcome provides a focal point around which individuals can coordinate more efficiently the transition to a new superior alternative. Obviously, the regime of no stare decisis would be inefficient because parties might not be able to transition to the new technology. To summarize, if adaptive transaction costs are high, people change their behaviour in time $t_1$ under a regime of stare decisis, while they remain using the old technology under no stare decisis.

In conclusion, (i) if the level of adaptive information and transaction costs is low, no stare decisis might contribute to favouring efficient legal change, and (ii) if the level of adaptive information and transaction costs is high, a regime of stare decisis might entail less maladaptation costs then no stare decisis would. This suggests that efficiency might require a weaker doctrine of precedent in regulatory environments where adaptive

993 It is more difficult to establish a priori whether a legislation-induced legal change is quicker or slower than legal change under stare decisis. In fact, both legislation and adjudication under stare decisis are centralized systems that entail high law-making costs, as compared to spontaneous adaptation under no stare decisis.
transaction costs are low and a stricter doctrine of precedent where adaptive transaction costs are high.

7.4.2.2. The Rate of Change of the Regulatory Environment

The structure of adaptation costs depends on the differential between the “rate of adaption” of law-making and the “rate of change” of the regulatory environment. It is worth clarifying this point.

Rate of Adaption. The doctrine of precedent provides benefits in terms of the reduction of legal uncertainty. Increasing costs of lock-in counterbalance these benefits. That is, there is a trade-off between uncertainty reduction and adaptive costs. Assume that the costs and benefits associated with the doctrine of precedent can be measured, so for varying degrees of strictness of the doctrine of precedent we can calculate the benefits deriving from the uncertainty reduction and the adaptive costs of the lock-in effect. The optimal degree of “strictness” of the doctrine of precedent is reached at the point where the marginal cost of an increase in the strictness of precedent equals its marginal benefit. Figure 31 illustrates this point.

---

994 Both judicial uncertainty and uncertainty apply to people subject to the law.
**Rate of Change.** The rate of change of the regulatory environment makes the practice of *stare decisis* more costly. Frequent changes increase the opportunity costs of path dependence in terms of the maladaptation costs associated with the obsolescence of the legal regime. This means that the rate of change of the regulatory environment affects the optimal level of strictness of the doctrine of precedent. For example, an increase in the rate of change determines an upward shift in the curves of the marginal costs of severity. This implies that a greater rate of change corresponds with a *lower* optimal degree of severity. Figure 32 illustrates this point.
To be precise, it should be recognized that what counts is not the absolute rate of change, but the rate of change *relative* to the rate of law-making adaption. This entails that, all other conditions being equal, the optimal strictness of *stare decisis* decreases as the ratio between the rate of change and rate of adaptation increases. Figure 33 depicts a curve representing the degree of strictness of *stare decisis* as a function of the ratio between the rate of change in the technological environment and the rate of the adaptation of the law.
The preceding considerations suggest that the doctrine of precedent should be flexible depending on (i) the rate of change on the regulatory environment, (ii) the size of maladaptation costs, and (iii) the magnitude of the beneficial effects of the legal precedent.

The field of information technology is an example of a regulatory environment characterized by a high rate of change, which has necessitated continuous adaptive efforts by lawmakers engaged in the production of legal rules to protect intellectual property. The history of copyright law is a history of continuous, adaptive adjustments of law-making to the changes induced by technological innovation. For instance, the mass diffusion of software, the increased digitalization of copyrighted goods, and the ubiquity of network technologies have dramatically changed the structure of the costs of creating and disseminating information. In this area, the U.S. Congress has traditionally regarded the courts as an institutional actor that is better suited in balancing the important concerns
raised by the evolving technological innovation with the legal protection of intellectual property rights. However, although judge-made law probably enjoys significant advantages over centralized political law-making, it might be that in this area an attenuated practice of *stare decisis* (or, generally, a weaker doctrine of precedent) might provide a better balance of the offsetting costs and the benefits of the rule of binding precedent.

### 7.4.2.3. Dichotomous Nature

A second source of the comparative disadvantage of adjudication is the dichotomous nature of the adjudication outcome. The resolution of (complex) cases usually reduces to a sequence of dichotomous questions having a yes or no answer. According to Kornhauser, “cases may have only two results or outcomes 1 or -1 that might be thought of as ‘for plaintiff’ or ‘for defendant’.” The dichotomous nature of adjudication affects the ability of the adjudicatory process to facilitate legal change. This point can be understood by considering that—as discussed in Chapter 1—legal change usually entails the presence of losers. The compensation of losers is often a necessary condition in overcoming their political opposition. Once the lawmaker has identified a social-welfare maximizing move, it must provide losers with appropriate compensation. Otherwise, they might strongly oppose the legal change, thereby dramatically increasing the adaptive transaction costs.

Adjudication entails discrete rather than continuous choices, which results in the dichotomous nature of the adjudication outcome, which limits the capability of judges to

---

996 This is, of course, a simplification; yet, it fairly captures one significant structural difference between the adjudication process and the political process.
997 Kornhauser, “Modelling”, *supra* note 962
998 Trebilcock, *Dealing with Losers, supra* note 37.
999 A discrete choice model is one in which decision makers choose among a set of alternatives. To fit within a discrete choice framework, the set of alternatives—the choice set—needs to exhibit three characteristics: (i) alternatives need to be mutually exclusive, (ii) alternatives must be exhaustive, and (iii) the number of alternatives must be finite.
compensate losers. Judges have limited “offsetting capabilities”. Once they have identified the efficient solution to a dispute, if it entails a change of the existing rules, they can hardly compensate the losers with some offsetting measures. This determines a comparative disadvantage of adjudication over politics. As I will discuss subsequently, within the legislative arena opposing groups can engage in bargaining strategies (e.g., vote-trading and logrolling). In contrast, in the adjudication arena opposing groups that choose to engage in litigation must accept the risks of a winner-take-all game. Furthermore, the limits to the offsetting capabilities of judges are exacerbated by the dispositive principle, which further constrains the offsetting capabilities of judges.

7.4.2.4. Judges’ Incentives

A third source of the comparative disadvantages of adjudication versus politics is the judges’ incentive structure. It seems plausible that judges receive fewer expected returns from legal innovation than politicians do. This derives from the institutional framework within which judges and legislators operate. Judges are expected to decide legal disputes by following past legal decisions. Their institutional role is that of reinforcing the objectives of either the private legal ordering (in decentralized systems) or publicly enacted law (in centralized systems); in both cases, judges are expected to preserve the objectives that are established by other institutions. On the contrary, legislatures are institutions designed to organize collective action in pursuit of some collective goals, which entails statutory reforms or revisions of bodies of law.

---

1000 This point has been identified by Gillette, “Lock-in Effects”, supra note 351 at 831 (“A judge who shares the interests of a group that wishes to maintain the status quo can do little more than render decisions that maintain the existing rule. The zero-sum nature of litigation and the inability to decide issues not presented to them provide courts little opportunity to trade away social gains in favor of preventing those losses that would be imposed on the groups that prefer to lock in the existing regulation”).

1001 Ibid at 831.

1002 Let me be clear on this. I am referring to the prospective component of the judicial decision, that is, to the creation of the rule through decisions that are binding (in the sense discussed previously) on future decisions. The offsetting capability of judges is much greater when deciding prospectively future cases than when deciding retrospectively the pending disputes before them.

1003 According to the dispositive principle, courts are entitled to hear cases and decide issues presented before them by litigants, and cannot satisfy claims to a greater extent than requested.

1004 This is a slippery analytical terrain because, as previously noted, it is difficult to identify a set of maximanda that allow constructing a judicial utility function.

1005 See Gillette, “Lock-in Effects”, supra note 351 at 827 (“The institutional framework within which a legislature operates may facilitate transitions more easily than is the case with judicial efforts, simply
In brief, judges are expected to keep doing what they have done in the past unless new circumstances arise, whereas legislators are more likely to be expected to promote legal change and to improve the status quo. Consequently, by seeking legal change, politicians are more likely than judges are to enjoy higher expected returns in terms of their reputation.\(^{1006}\) It must be recognized that although plausible, the preceding considerations do not say anything on the direction and the quality of the legal change. In particular, they do not distinguish between incentives to favour legal change either (i) by accommodating rent-seeking pressures or (ii) by promoting efficiency. The analysis of politics in the next chapter will illuminate this crucial point.

**Conclusions**

Since the arguments in this chapter have been long and complex, a short summary might be helpful.

*Productive efficiency*

The economic calculation problem suggests that it is impossible for courts to achieve efficiency; the amount of information needed to mimic market prices exceeds judges’ abilities to calculate relative prices and estimate utility functions. If a tendency towards efficiency exists, it must be sought in some evolutionary dynamics that are mostly independent of judges’ attitudes toward efficiency. Thus attention must be placed on the *structural* features of the adjudication process that judges use to deal with limited information, bounded rationality, and the resulting uncertainty. I identify the following information-generating mechanisms: (i) the information provided through the litigation process; (ii) the doctrine of precedent; (iii) the degree of decentralization.

Regarding the demand side, various variables are likely to influence the efficiency of law-making through adjudication, including (i) the degree of substantive inefficiency of existing legal rules, (ii) the parties’ incentives to establish favourable legal precedents,
(iii) parties’ willingness to invest in litigation, (iv) the structure of parties’ available information (depending on the clarity of law, the presence of informational asymmetries, and so on), (v) the relative costs of settlement and litigation, (vi) the structure of the relative costs of organizing collective action for the different interests at stake. Combined, these factors are likely to influence the evolution of judge-made law.

The doctrine of precedent is a decision-making technique based on the assessment of the reasons determined by previous courts and relating these to the facts of the case. Reliance on precedents allows for the following advantages: (i) reduction of decision-making costs, (ii) reduction of uncertainty and errors, (iii) specialization of judges, and (iv) maximization of the regulatory dimension of the judicial decisions. These advantages do not entail a clear tendency of judge-made law toward efficiency. If any efficiency is achieved, it is efficiency “on the average” of cases rather than in single cases. Furthermore, judges see only a biased sample of possible cases as a result of private parties’ settlement-trial decisions. It follows that judges do not have access to the information needed to develop the rule that is efficient for the full set of cases. Finally, compared to statutory rules, precedents entail, in general, higher decision-making costs. This is because statutes are based on presumptive and general exclusionary reasons, whereas precedents are based on absolute and specific exclusionary reasons.

On the supply side, a high degree of polycentricism of the legal order accentuates the informational advantages of adjudication. In particular, four features of polycentric orders are relevant: (i) the degree of competition among alternative jurisdictional fora; (ii) the judges’ tendency to produce default rules; (iii) the judges’ reliance on customs; (iv) the decentralization of the sources of judicial information.

Agency Efficiency
From the standpoint of agency efficiency, two features of adjudication have comparative advantages over politics: (i) lower cost access thresholds; (ii) courts’ limited rent-offering power; (iii) courts insensitivity to external influence at lower levels of influence expenditures. These features allow for (a) reduced outcome-agency costs and (b) reduced rent-seeking-agency costs.
Related to the preceding considerations, the degree of generality and stability of the adjudication outcome have significant implications in terms of agency efficiency. On the one hand, the greater generality of legal precedents, as compared to statutory rules, is a source of comparative advantage of adjudication over politics because it reduces the scope for purely redistributive transfers of wealth. On the other hand, however, the stability of legal precedents might counterbalance this advantage by providing greater incentives for rent-seeking.

This chapter has emphasized that, unlike statutory rules, judicial rules are theoretical constructs. The constraints confronted by a judge when fashioning the content of the rule is logical rather than political. Logical constraints tend to be more severe than political constraints are. In fact, the fact-specific nature of case-by-case decision-making process requires courts to adopt a higher degree of generality in the production of law than rules formulated at the will of politicians.

Finally, the degree of centralization of the judicial system has significant implications for agency efficiency. Judges’ rent-offering power tends to increase as the level of centralization increases. In general, the less polycentric a judicial system is, the more it tends to share the flaws, in terms of agency inefficiency, of the centralized political process.

**Social Choice Efficiency**

The discussion in this chapter has identified four structural features of the adjudication process that mitigate the problem of agenda manipulation and cyclical path-dependence. First, unlike legislatures, courts do not set their own agenda because the flow of cases is decided by litigants’ settlement-trial decisions. In addition, agenda manipulation is also difficult for private parties. In fact, litigants have conflicting interests at stakes and must meet justiciability requirements, which reduce the scope of monopolizing agenda control. Second, in the model of reason-based decision-making, a judge is never confronted with the option of reconsidering a rule that has been excluded by a prior decision. If adherence to a precedent is strict, a judge can apply a new rule only when new factual elements render the previous ruling inadequate or incomplete. Third, the cyclicality problem is
generated by preference-based rationality, whereas judicial law-making rests on reason-based rationality. Transitivity in adjudication is ensured by the rules of legal reasoning, independent of the full transitivity of the social preference orderings. Fourth, unlike statutory rules, common law rules can be negotiated around by private parties. This significantly reduces the cyclicality problem, which provides a comparative advantage that is greater in decentralized adjudicatory systems.

Adaptive efficiency

From the standpoint of adaptive efficiency, adjudication enjoys the following advantages over politics: (a) the levelling-playing-field-effect associated with adjudication mitigates the lock-in effect associated with the doctrine of precedent, (b) the incremental nature of judicial legal change reduces decision-making costs and provides some strategic political advantages to courts in promoting legal change, and (c) adjudication might be advantageous to change-seeking actors that are minoritarian in the political process.

Two variables that are characteristic of regulatory environments should be considered when assessing the adaptive efficiency of adjudication. First, if the level of adaptive information and transaction costs is low, a less strict doctrine of precedent might better contribute to favouring an efficient legal change. If the level of adaptive information and transaction costs is high, a stricter regime of stare decisis might entail less maladaptation costs, thereby reducing outcome inefficiency. The second crucial variable is the rate of change in the regulatory environment. In particular, the optimal strictness of the doctrine of precedent decreases as the ratio between the rate of change and the rate of law-making adaptation increases.

As previously emphasized, although the incremental nature of the judicial process entails strategic advantages, it must be recognized that the dichotomous nature of the adjudication outcome affects judges’ capability to compensate losers adequately, thereby limiting the ability of the adjudicatory process to facilitate legal change.
Normative Considerations

Although the findings of the foregoing analysis are fundamentally positive, two normative implications emerge. First, a flexible doctrine of precedent might increase the efficiency of law-making through adjudication. That is, the degree of strictness of the doctrine of precedent might be adjusted to the varying features of regulatory environments. For example, when rent-seeking pressure by interest groups is strong, a greater flexibility of the doctrine of precedent would reduce the expected value of the extractable rents, thereby reducing the agency costs associated with judge-made law. In addition, where the rate of change of the regulatory environment is high and adaptive transaction costs are low, a weaker doctrine of precedent might improve efficiency. Conversely, in changing environments where adaptive transaction costs are high, a stricter doctrine of precedent might be preferable, from the standpoint of efficiency.

A second normative implication emerging from the foregoing discussion is that an efficient evolution of judge-made law is likely to be facilitated by institutional reforms that increase the degree of decentralization in the judicial system. The flexibility of the doctrine of precedent and the decentralization of the judicial law-making design are two sides of the same coin. Jointly applied to the adjudication system, they enhance the evolutionary nature of judge-made law and strengthen the market forces underlying the production of judicial precedents. This will not generate unambiguous tendencies toward efficiency but will probably reduce the degree of inefficiency.
Chapter 8

SPONTANEOUS LAW-MAKING

The issue of right or wrong arises only when cooperation is disturbed and each wants to shift the blame for the discord.

Rudolf Dreikurs, The Challenge of Marriage, 1946

This chapter focuses on the spontaneous emergence of norms in the absence of public, centralized law-making institutions. The analytical focus of the discussion is on (i) the process of norms emergence, (ii) the environmental conditions under which efficient norms are likely to emerge spontaneously, and (iii) the comparative advantages and disadvantages of spontaneous law-making relative to other sources of law.

Introduction: The Spontaneous Emergence of Norms

Assume the following environmental setting:

(i) the presence of social constraints on individual behaviour preventing people from engaging in violent appropriations of others’ property (such as fraud, robbery, theft, and so forth);

(ii) the absence of a centralized law-making authority that has the capacity to create and enforce legal rules rewarding cooperation and punishing opportunistic behaviour in transactional contexts;\(^\text{1007}\)

(iii) the availability of a potential Pareto superior equilibrium for the society that is achievable if and only if a minimum number of community members observe and enforce

---

\(^{1007}\) This scenario corresponds largely to the one described in Anthony T. Kronman, “Contract Law and the State of Nature”, (1985) 1 J L Econ and Org, 5 [Kronman, “Contract Law State of Nature”] (identifying a scenario characterized by “possessory security” and “transactional insecurity”). It also corresponds to the scenario analyzed in Ellickson, Order Without Law, supra note 356 at 144.
an efficient standard of behaviour by overcoming collective action problems and strategic opportunism.\textsuperscript{1008}

Based on these assumptions, the two questions that concern us are (i) whether—and, if so, under which conditions—a process of repeated dyadic contacts among community members can lead over time to the emergence of efficient norms, and (ii) whether the spontaneous emergence of norms (hereinafter “SEN”) has comparative advantages in terms of efficiency relative to alternative law-making institutions.

Before proceeding, two preliminary clarifications are needed. First, the analysis that follows includes both social norms and customary rules. I ignore the differences between these two categories of norms and collapse them into the notion of “spontaneous norms”. I am interested in understanding the conditions and the process of emergence of socially enforced standards of behaviour in the absence of centralized mechanisms, regardless of whether these are qualified as social norms or warrant legal consideration as customary rules. Second, a vast array of processes has the capacity to generate norms in the absence of a public centralized system. These are referred to in the literature as “private legal orderings” (hereinafter “PLOs”).\textsuperscript{1009} A comprehensive analysis of PLOs would carry us

\textsuperscript{1008} See the definition of social dilemma provided by Thomas Voss, Game Theoretical Perspectives on the Emergence of Social Norms [Voss, “Game Theoretical Perspectives”] in Michael Hechter and Karl-Dieter Opp, eds., Social Norms (New York: The Russel Sage Foundation, 2001) [Hechter and Opp, “Social Norms”].

In the literature, there is no one recognized common classification of PLOs, and alternative criteria are used to identify different typologies. The first classifying criterion distinguishes between two different functions of norm-producing processes: a) the creation, and b) the enforcement of the norms. Based on this distinguishing element, three types of PLOs can be identified: (i) norms created by private actors and enforced by governments; (ii) norms created and enforced by private actors pursuant to governmental delegation; (iii) norms originated and enforced by private actors without governmental sanction or enforcement. In each case, rules are private, where the differences concern the enforcement dimension.

A second classification (see David Charny, “Illusions of a Spontaneous Order: ‘Norms’ In Contractual Relationships” 144 (1996), U Penn L Rev, 1841) is based on the type of conflict of interest generating the demand for norms, and it identifies three types of PLO: (i) transactional systems; (ii) common-pool systems; and (iii) external-harms systems. In transactional systems, the situation generating the conflict of interests is the exchange of goods or services. In common pool systems, the conflict arises because the parties share a common resource that becomes quickly depleted if the participants do not regulate rival consumptions. In the third type of non-legal sanctioning systems, the conflict of interests is an occasional harm: one party occasionally harms the other or both parties occasionally harm each other.

A third classification identifies three types of nongovernmental regulation, based on the entity entrusted with the task of detecting and preventing norm violation. This key element is the proximity of the
too far afield. I focus here on the spontaneous emergence of norms through repeated interaction in the absence of a centralized public enforcement authority. However, because some insights offered by the discussion can be generalized and extended to other PLOs, in section 5 of this chapter I will discuss some illustrative examples including both SEN and other PLOs.

8.1. Productive Efficiency

8.1.1. The Relative Advantages of Spontaneous Norms

Customary law and social norms rest on the widespread consensus of people subject to the law. The crucial feature of SEN is that norm producers (i.e., the actors who participate in the norm-formation process) coincide with the target actors (i.e., the actors whose behaviour is regulated by the norm), and norm beneficiaries (i.e. those whose interests are protected by the norm). The fact that norm producers coincide with both targeted actors and norm beneficiaries is a source of significant comparative advantage relative to centralized processes. In particular, under certain conditions, SEN is more efficient and effective in capturing local individual preferences and identifying socially efficient standards of behaviour.

regulator entity to the norm violator. Three general types of regulators can be identified as follows: (i) first-party, (ii) second-party, and (iii) third-party. In first-party regulation, also referred to as “self-regulation”, the role of enforcing authority is attributed to the potential norm violator (e.g., self-industry regulation)\textsuperscript{1009}. In second-party regulation, also called transaction regulation, the role of the regulator is attributed to the parties engaged in transaction. Finally, third-party controllers can be governmental or nongovernmental organizations.

We can summarize the different classifications of private law-making mechanisms as shown in the following chart.

<table>
<thead>
<tr>
<th>(1) Public/private enforcing</th>
<th>Private enforcement</th>
<th>Private enforcement pursuant to public delegation</th>
<th>Public enforcement</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Conflict of interests</td>
<td>Transactional</td>
<td>Common pool</td>
<td>External harm</td>
</tr>
<tr>
<td>(2) Proximity of enforcing entity</td>
<td>Self-regulation</td>
<td>Transaction regulation</td>
<td>Third-party regulation</td>
</tr>
</tbody>
</table>
First, SEN has a comparative advantage over centralized mechanisms because people express their preferences through direct participation in norm formation process instead of delegating the norm-creating function to a third party. This increases the informational efficiency of the process of norm creation because, unlike centralized law-making bodies, individuals have a direct perception of their costs and benefits. As discussed in section 4 of this chapter, the relative advantages of SEN should always be weighed against the problems of adaptive efficiency under conditions of high information and transaction costs. Second, the fact the norm-producers and norm-beneficiaries are two identical sets of individuals raises the cost of irrational behaviours in the process of norm creation, and therefore generates incentives to have rational beliefs and behave accordingly. This is a significant difference between the incentive structure confronted by norm-producers in SEN and the incentives faced by voters and politicians in political processes. In essence, the absence of delegation of law-making authority is the source of the major comparative advantages of SEN over politics and bureaucracy.

Finally, SEN can have a relative advantage over bureaucracies in environments characterized by technical complexity. This is true, for example, in professional environments where the identification of efficient standards of conduct requires technical information and the specialized knowledge possessed by professionals. In general, greater economic specialization widens the informational disadvantages of centralized processes and calls for decentralized forms of norm-production. From this perspective, the identification of the environmental conditions facilitating SEN is a useful exercise in identifying cases where significant efficiency advantages can be obtained by shifting law-making power from politics and bureaucracies to customs and social norms. The following discussion aims to clarify this important issue.

---

8.1.2. Norms as Evolutionary Responses to Game Inefficiencies

Having clarified the potential advantages of SEN, it is now necessary to identify the conditions under which efficient norms are likely to emerge spontaneously in the absence of centralized law-making mechanisms.\footnote{1011}

Broadly, SEN is efficient to the point where the marginal net benefit of letting private parties devise efficient solutions to their coordination problems exceed the marginal benefit of doing so by means of increasing degrees of centralization of law-making and law enforcement.\footnote{1012} However, this general principle is devoid of meaningful practical content without previous careful reconstruction of the costs associated with the process by which norms emerge, persist, and change over time. For this reason, the next step is to identify the mechanism through which norms arise spontaneously. This subsection discusses norm-emerging mechanisms, and the following subsections will examine their relative efficiency implications.

For the purposes of the present study, the process of SEN can be usefully explained by using the analytical instruments offered by transaction-costs and game-theory approaches.\footnote{1013} Let me consider a basic scenario where the emergence of some coordination problems among private actors generates a demand for norms.\footnote{1014} In the ideal world of zero transaction costs,\footnote{1015} the evolutionary forces generated by the competitive adjustments of individuals to changing situations leads parties to adopt the norms that they would have chosen if they had been free to enter into binding agreements.\footnote{1016} That is, in a zero-transaction costs world, the spontaneous emergence of

\footnotesize
\begin{enumerate}
\item \footnote{1011} The two questions are closely related since the conditions facilitating the SEL are in many cases the same conditions that increase the net-benefits of cooperative behaviours, thereby generating the comparative advantages of SEL versus more centralized forms of law-making.
\item \footnote{1012} Kronman, “Contract Law State of Nature”, supra note 1007.
\item \footnote{1013} The costs associated with SEL are in large part related to the structure of the strategic interaction between individuals, and conversely, the outcome of the strategic interaction among individuals depends upon the structure of the costs associated with the adoption and enforcement of such a standard of behaviour.
\item \footnote{1014} The problem can be usefully conceptualized with that of the demand for “switching” to a new and more efficient norm. I will discuss this aspect in the section of this chapter devoted to the adaptive efficiency of SEL (i.e., that of the demand for “switching” to a new and more efficient norm). I will discuss this aspect in this section of this chapter devoted to the adaptive efficiency of SEL.
\item \footnote{1015} I.e., in a world without the costs associated with the process of adopting and enforcing norms.
\item \footnote{1016} Parisi, “Theory of Spontaneous Law” supra note 1010.
\end{enumerate}
law through repeated contractual practice would be the most efficient law-making process. There would be no obstacles to cooperation, and parties would be able to adopt efficient legal rules according to changing economic circumstances. From a repeated game-theoretic perspective, in an environment without transaction costs, every time a Pareto-superior alternative becomes available—thereby making the existing equilibrium suboptimal—evolutionary corrective mechanisms, spurred by repeated bargaining practice, would lead private actors to reach a Nash equilibrium that corresponds to the adoption of a superior norm that enables them to attain the maximum available (cooperative) surplus.\textsuperscript{1017}

In contrast, the presence of significant transaction costs of creating and enforcing norms explains the emergence and persistence over time of inefficient customary legal rules, as well as the emergence in society of alternative sources of law. In general, if the costs for the individual of contributing to the collective enforcement of a superior standard of behaviour exceed the expected efficiency losses generated by the coordination problem the rule is supposed to resolve, then there is no incentive for the individual to contribute to the process of law creation and enforcement. Consequently, either the community remains in a suboptimal equilibrium or the construction of a centralized law-making process becomes necessary for the society to adopt the superior norm.

Once the assumption of zero transaction costs is dropped, the existence of a Pareto-superior alternative is not a sufficient condition for the spontaneous emergence of the efficient legal rule.\textsuperscript{1018} A Pareto-superior rule tends to emerge and become consolidated over time only if the pattern of behaviour that is compatible with the efficient norm is supported by Nash strategies, by all parties.\textsuperscript{1019} If the strategy consistent with the emergence and persistence of a superior norm (e.g., a cooperative behaviour) does not correspond to the dominant strategy for one of the actors involved, he or she will always be tempted to switch to the dominant strategy, thereby preventing the community from consolidating a stable optimal Nash equilibrium. In essence, in environments with

\textsuperscript{1017} Ibid. at 214
\textsuperscript{1018} Ibid.
\textsuperscript{1019} Ibid. See also Voss, “Game Theoretical Perspectives” supra note 1008.
significant transaction costs, two conditions are necessary to support SEL: (i) the availability of a Pareto-superior norm prescribing a pattern of behaviour that enables society to attain the higher cooperative surplus; (ii) a Nash strategy equilibrium supporting the pattern of behaviour compatible with that superior alternative. Under these two conditions, norms emerge *spontaneously* over time, which enables the members of a community to overcome evolutionary stalls and achieve an efficient outcome instead. In essence, under appropriate circumstances, customary rules arise as “evolutionary responses to game inefficiencies” or “possible correctives to strategic stalls”.1020

The foregoing discussion leads to an important conclusion. When the behaviour conforming to a norm is supported by a Nash strategy for all actors involved, the norm is *self-enforcing*, that is, rational actors conform in response to incentives that make conformity to the norm convenient to them, despite the absence of a centralized enforcing entity administering sanctions to norm violators. Crucially, the self-enforcement mechanism allows for the spontaneous emergence and persistence of norms. Self-enforcing norms are also called “conjoint norms”,1021 emphasizing the fact that the beneficiaries of the norm and the target actors *conjoin* to create and enforce norms, which enables them to cooperate and attain the maximum available cooperative surplus. The sets of the beneficiaries and target actors are identical and, by force of the incentive structure underlying Nash strategies, they generate and enforce mutually binding norms of behaviour.

The next step is to identify the conditions under which an *efficient* Nash equilibrium can be reached in a social dilemma situation. To answer this fundamental question, it is useful to inquire into the *structure* of the situation that generates the demand for norms. In this respect, it is useful to distinguish between (i) prisoner-dilemma (hereinafter “PD”) situations; and (ii) coordination problems (hereinafter “CP”).1022 This distinction helps identify alternative patterns of emergence. In CP situations, players share a common

---

1020 Parisi, “Theory of Spontaneous Law”, *supra* note 1010 at 211.
interest in mutual cooperation, and they have incentives to coordinate their choices; when players have overcome the coordination problem, they have no incentive to move away from the cooperative equilibrium and engage in unilateral defection. In contrast, in PD situations, opportunistic behaviour and unilateral defection are the dominant strategies for the actors involved. The classic two-persons, PD situation can be summarized as follows: Consider two individuals \( a \) and \( b \); in a PD situation \( a \) faces the following an incentive structure: (1) if \( a \) unilaterally defects and \( b \) cooperates, then \( a \) obtains the highest possible payoff; (2) if \( a \) unilaterally cooperates while \( b \) defects, then \( a \) receives the worst possible payoff; (3) if \( a \) defects and \( b \) defects, then \( a \) is worse off than in (1) but better off than in (2); (4) if \( a \) and \( b \) cooperate, then \( a \) receives less than in (1) but is better off than in both (2) and (3). The same payoffs schedule holds for \( b \). It follows that in the absence of any assurance of the other player’s cooperation, each player is confronted \textit{ex ante} with the following payoff schedule:

\begin{align*}
(1) \text{ unilateral defection} &> (4) \text{ bilateral cooperation} > (3) \text{ bilateral defection} > (2) \text{ unilateral cooperation}.
\end{align*}

That is, both players have incentive not to cooperate in order to avoid the worst possible scenario of unilateral cooperation. In contrast, from an aggregate standpoint, the payoff schedule is as follows:

\begin{align*}
(4) \text{ bilateral cooperation} &> (2) \text{ unilateral cooperation or (1) unilateral defection of either actors} > (3) \text{ bilateral defection}
\end{align*}

That is, bilateral cooperation maximizes the social surplus. However, as clarified above, because both individuals have incentives to a unilateral defection in the absence of any coordinative device, the Nash equilibrium is a bilateral defection, which minimizes the social surplus. This means that from the standpoint of social efficiency, self-interested rational individuals choose the strategy that is conducive to the worse possible scenario of mutual defection.
The different incentive structure of CP and PD situations entails different conditions of norm emergence. In CP situations, the establishment of a *convention* is sufficient to enable a group to produce a cooperative surplus and benefit each individual. Because of the common interest in cooperation, the convention is a self-enforcing equilibrium; individuals do not have an incentive to defect unilaterally because the conformity to the convention benefits everyone. In essence, there is no incentive problem. The convention provides a focal point around which individuals coordinate their behaviour; therefore, the enforcement of sanctions is unnecessary to ensure cooperation. In contrast, the establishment of a convention is not sufficient to ensure cooperation in PD situations because once the convention is realized, parties face an incentive *not* to conform to the prescribed behaviour. Consequently, sanctions against violators must be administered to ensure mutually advantageous cooperation.

### 8.1.3. Self-Enforcing Norms in Prisoners’ Dilemma Situations

Let me now focus on the conditions favouring the spontaneous emergence of conjoint norms in PD situations. This is a puzzling issue. As repeatedly emphasized, in PD situations because rational actors do not have incentives to cooperate, they fail to achieve an efficient outcome that would benefit them. The failure to achieve a cooperative outcome generates the demand for a norm ensuring cooperation. This is also called the “first order” collective action problem. However, because creating and enforcing norms are costly for the individual (as is cooperating to achieve the efficient outcome), a “second order” collective action problem arises.\(^\text{1023}\) If parties do not have incentives to achieve the efficient cooperative outcome, *how do they have the incentive to cooperate to produce a conjoint social norm, which requires cooperation in the creation and enforcement process?* In other words, if individuals fail to overcome the first order collective action problem, how can they overcome the second order collective action problem? This subsection provides an answer to this question and identifies the mechanisms that induce people to create social norms in the absence of a centralized lawmaker.

\(^{1023}\text{Oliver, “Rewards and Punishments”, supra note 236 at 210.}\)
Repeated game theory has long demonstrated that under appropriate circumstances, long-term repeated interactions generate the conditions for the decentralized self-enforcement of conjoint social norms. To understand this point, it is useful to focus on the core parameters of game-theory models: (i) the structure of parties’ payoffs, that is, the relative costs of conflict and cooperation, (ii) the discount factor, that is, the “shadow of the future”, (iii) the number of decision-making actors, that is, the size of the group that benefits from the norm, and (iv) actors’ ability to promptly punish defection. Based on these four variables, repeated game theory suggests that self-enforcing norms of cooperation are likely to emerge within (relatively) small communities of culturally and economically homogenous farsighted individuals engaging in repeated mutual interactions. Consistent with these conclusions (but from a different methodological perspective) Ellickson has suggested that social welfare-enhancing norms tend to emerge spontaneously if the population is a close-knit community. The notion of close-knittedness summarizes the essential conditions facilitating the spontaneous emergence of efficient norms in environments lacking a centralized law-making authority. The next two subsections identify the constitutive elements of “close-knittedness” and inquire whether institutional mechanisms exist that induce cooperation by mimicking the conditions of close-knittedness.

8.1.3.1. Close-Knittedness

Close-knit environments reduce the transaction and information costs associated with informal enforcement, thereby fostering cooperative behaviour in PD situations. By drawing on Ellickson’s work and other more recent contributions, it is possible to identify

---


1025 See Ellickson, *Order Without Law, supra* note 356 at 167: “members of a close-knit group develop and maintain norms whose content serves to maximize the aggregate welfare that members obtain in their work-day affairs with one another”. The hypothesis that norms emerge in society as a social-welfare enhancing mechanism is also put forward by Ullmann-Margalit, “The Emergence of Norms” supra note 1022 and Opp, Karl-Dieter, “The Emergence And Effects Of Social Norms: A Confrontation Of Some Hypotheses Of Sociology And Economics” (1979) 32 Kyklos, 4, 775.

1026 Ellickson’s contribution has sparked a growing body of legal-economic scholarship focusing on the study of social norms. I inquire into the notion of close-knittedness by also drawing from other scholars’ contributions that have advanced our understanding of the emergence of efficient self-enforcing norms.
the following conditions under which repeated human interactions provide individuals with enough incentive to choose cooperative social welfare maximizing strategies.

(i) Non-Zero Sum Situations
Zero-sum (or constant-sum) games are purely distributive situations: what one player gains, the other loses. These are purely conflictual games with no room for mutually advantageous cooperation. Because there is no available aggregate surplus from reciprocal cooperation or coordination, the conflictual nature of distributive games entails no incentive for mutually advantageous cooperation or coordination. Thus, in zero-sum (or constant-sum) situations, norms will not emerge spontaneously without the presence of a centralized third-party institution. ¹⁰²⁷ Unlike purely distributive games, close-knit environments are characterized by the availability of a surplus gained from mutual cooperation. That is, norms ensuring cooperation can emerge spontaneously in positive sum games.

(ii) Reciprocal Enforcing Power.
In strategic interactions, credible threats and promises affect actors’ strategies.¹⁰²⁸ This principle applies to the context of a close-knit community, with the presence of a promise (or threat) to promptly enforce sanctions in the case of norm violation. The condition of credibility is because the informal power to punish norms’ violations is widely distributed among the members of the community. Differently stated, the power reciprocity among group’s members constitutes the structural condition ensuring the credibility of the enforcing mechanisms. I will discuss reciprocity in greater detail in the next section.

(iii) Continuing Relationships.
In the absence of a centralized enforcing mechanism, continuing relationships among community members provide the opportunity for a credible enforcement mechanism. Close-knit communities are stable social networks in which the probabilities of future

¹⁰²⁷ See Gordon Tullock, The Social Dilemma: The Economics of War and Revolution (Blacksburg: University publications, 1974).
¹⁰²⁸ To be effective, threats and promises should also be consistent with a Nash equilibrium.
encounters are high enough to make the enforcement of sanctions credible. The prospect of future repeated interactions has a twofold incentive effect. On the one hand, it increases the expected costs of defection versus cooperation for norm violators (thereby increasing the net benefits from abiding by the rule). On the other hand, it increases the net benefit gained by participating in collective enforcement because the expected benefit of the compliance of a target actor increases with the likelihood of future encounters. In short, the probability of future interactions among members of a close-knit community modifies current individual behaviours by strengthening the enforcement mechanism and weakening the attitude toward norm violation. In turn, this enables the community to prevent or limit opportunistic behaviours that defeat cooperation.

The possibility of future interaction plays an important role in situations where individual incentives are misaligned.\textsuperscript{1029} In this respect, it has long been demonstrated that the effect of future interactions on the levels of cooperation in a prisoner’s dilemma game is largely influenced by the players’ discount rate. This line of reasoning culminates in the folk theorem according to which any (individually rational) payoff can be supported by a Nash equilibrium in infinitely repeated games, if players have sufficiently high discount factors. Although a discussion of the folk theorem is beyond the scope of this study, it is important to underline the existence of a threshold value of the discount factor, above which rational private parties decide to cooperate.

The discount factor is a function of both (i) the time preference of players, and (ii) the expected present value of future interactions.\textsuperscript{1030} First, the time preference is negatively related to the discount factor. As time preference increases, the discount factor decreases (i.e., becomes closer to zero), so the present discounted value of future payoffs decreases. This means that the higher the time preference is, the lower the players’ willingness to give up a present value for an expected future value. High time preference determines a lower payoff from future cooperation and thus less willingness to cooperate. Second, the expected present value of cooperation is positively related to the probability of future cooperation.


interaction. As the probability of future interaction increases, the discount factor increases (i.e., becomes closer to one), so the present discounted value of future payoffs increases, which means that the higher the probability of future interaction is, the higher the players’ willingness to give up a present value for an expected future value. *A higher probability of future interactions entails higher benefits from future cooperation and thus more willingness to cooperate.* In summary, environments promoting the *high* probability of future interactions and *low* time preference are favourable for the spontaneous emergence of stable evolutionary responses to game inefficiencies.

(iv) *Adequate Knowledge and Information*

The conditions of power reciprocity and continuing relationships are not sufficient to ensure cooperation. People have incentives to cooperate only if they have adequate information about the structure of *payoffs* associated with the alternative available course of action in each period. Second, cooperation requires historical knowledge about the behaviour of opposing players in prior games. In both respects, the existence of close-knit groups reduces the costs of obtaining the knowledge and information necessary for cooperation. Close-knittedness helps group members to develop a common “objective valuation system, which they need in order to assess the welfare enhancing tendencies of various norms”.*^{1031} This is a crucial factor: strategies that evolve into social norms are those that maximize the expected payoff for each group member and that are enforced against case-by-case opportunism. In addition, gossip networks and reputational mechanisms, which normally emerge in close-knit communities, enable people to gain information about how group members acted in the past and how reliable and trustworthy they are as potential trading partners.

Before proceeding, the following clarification regarding close-knittedness is needed. The four structural conditions identified above do *not* necessarily require that close-knit groups are *small.**^{1032} Although a small group size certainly facilitates cooperation, the structural requisites of reciprocity, long-term relationships, adequate knowledge and information may occur *independently* of the group size. Ultimately, targeted and

---


1032 *Ibid* at 182.
beneficiary actors may be geographically distant and operate in different industries. This point emphasizes the high potential of SEN as an effective source of regulation in various regulatory environments.

An example of SEN in a close-knit environment is the system of trade customs and dispute resolution regulating the translation commercial activity in Medieval Europe: the so-called *lex mercatoria*. This private legal order developed within the merchants’ community in response to the increasing development of *long-distance* trade relationships in which merchants had to entrust payments or the delivery of goods to persons removed in time and space. The community of merchants extended to Europe, North Africa and Asia Minor. In this emerging transnational contest, the creation of norms by small highly localized communities was inadequate to provide transactional security. Local authorities did not have the knowledge to devise efficient rules to support transnational commerce. Instead, an international community of medieval merchants possessed the capacity—and had the appropriate incentives—to develop rules establishing efficient standards of behaviour. This enabled a great expansion of the scope of international trade. Institutions such as the merchant guilds and the law merchant, which arose prior to the birth of nation-states (and centralized law-making processes) in Europe, enabled the community of medieval merchants to regulate the fluid and evolving context of the commercial revolution.1033

Critically, the commercial environment in Medieval Europe possessed the structural characteristics of close-knittedness (power reciprocity among traders, long-term relationships, adequate knowledge and information) *despite* the transnational dimensions of the trade community. The crucial problem confronted by merchants was the costliness of information about the histories of trading partners, which hampered the commitment necessary to support economic activity. In response to this problem, the institutional

---

setting supplied by the *lex mercatoria* provided a regime of information gathering and dissemination with the capacity of “making of reputation a transferable good, or ‘bond’, within the community of traders”.

The decisions of merchant judges on previous trade disputes were recorded, thereby reducing the costs of monitoring the history of potential trading partners. In this manner, the *lex mercatoria* supported a reputational mechanism that enabled merchants to secure credible commitments.

The reciprocity power among trading partners was such that the merchant who failed to comply with customary standards of behaviour (e.g., by not delivering goods or not fulfilling contractual obligations) was denied access to the trading community. Hence, by excluding defectors from future access to the community’s goods and profit, the reputational mechanism altered the incentives to either cheat or comply with contractual obligations, thereby “placing future contracting in the shadow of the law”.

### 8.1.3.2. Institutional Mechanisms

This section briefly examines two mechanisms that induce cooperation in the absence of a centralized enforcing authority: (a) role-reversibility and (b) reciprocity constraints. These mechanisms prove useful when the incentives of parties to an exchange are *not* aligned. When one of the two parties faces the incentive to defect unilaterally, and the opposing part anticipates this possibility, no exchange will occur and opportunities to exploit the surplus gained from mutual cooperation will not be realized. In these contests, cooperation-inducing mechanisms prove useful in realigning the incentives of parties and facilitating SEN.

### 8.1.3.2.1. Role Reversibility

Role reversibility occurs when the parties to a repeated dyadic exchange face over time the same probability of being on either side of the exchange. This is also called

---


1035 *Ibid*.

1036 Obviously, the principal mechanism used to redress incentive misalignment is the existence of enforcement-contractual agreement. Here, I am discussing the lawlessness of environments, which are characterized by the lack of centralized enforceable institutions. The central question begs whether there are structural conditions facilitating the conclusion of efficient dyadic exchanges in the absence of centralized enforcing institutions.
“stochastic symmetry”, underlining the fact that parties cannot predict on which side of the relationship they will be in future exchanges (e.g., they can be either seller or buyer with unpredictable probability). Under these conditions, parties are subject to symmetric incentives. Alternative rules entail the same expected benefit and costs; therefore, the parties face the incentive to converge upon the same set of rules that maximize their individual benefits (and maximize the aggregate welfare). In short, environments characterized by stochastic symmetry are more likely to be regulated by spontaneously emerged norms of behaviour.

The emergence of lex mercatoria is a classic example of a spontaneous law that emerged under conditions of stochastic symmetry. Merchants acted as both buyers and sellers; therefore, they had no incentive to promote systematically biased rules that favoured one party. As Parisi emphasizes, “[T]he stochastic symmetry of the relationship between medieval merchants eliminated the parties’ incentives to articulate one-sided rules”. That is, “[the] condition of role reversibility changed a structurally asymmetrical situation (buyers versus sellers) into one that was stochastically symmetrical (merchant versus merchant)”.

8.1.3.2.2. Reciprocity Constraints

Structurally asymmetric situations can be redressed through institutional constraints that induce reciprocity by binding the strategic choices of opposing parties. This mechanism is also called “induced symmetry”, emphasizing the fact that the symmetric advantages of cooperation are generated by the presence of institutional constraints that eliminate the incentives to undertake unilaterally defective strategies. Parisi explains reciprocity constraints as “Mechanisms that ensure automatic reciprocity eliminate the access to off-diagonal outcomes. If one players’ strategy is rigidly bound to that of his opponent, the reward for unilateral defection is rendered unobtainable”. This is because the mechanism of reciprocity constraints formalizes a “tit-for-tat” strategy, in

---

1037 Ibid. See also Parisi, “The Formation of Customary Law”, supra note 1029 at 10
1038 See Parisi, “Theory of Spontaneous Law” supra note 1010 at 217-218
1039 Ibid.
1040 Ibid (emphasis is mine).
1041 Ibid at 218.
1042 Ibid at 218-219
which opposing parties with symmetric incentives engage in a defection strategy that punishes unilateral defection.

In general, the principle of reciprocity governs the formation of customary rules in systems where custom is the main source of law.\textsuperscript{1043} The classic example is provided by international law, where the lack of a superior centralized political authority leaves wide room for unilateral defection strategies by single states. To illustrate this point, Parisi emphasizes the case of article 21(1)b of the Vienna Convention of 1969, which regulates the legal effects of the formulation reservation in the process of stipulating treaties among states. This provision establishes that the formulation by state A of a reservation against state B generates automatically the effect of producing a reservation of state B against state A.\textsuperscript{1044} Another interesting example of a reciprocity-inducing mechanism is the creation of patent pools in automobile and aircraft manufacturing industries. These privately made institutional arrangements are designed to facilitate the conditions of close-knittedness by “creating the occasion for repeat-play, reciprocal bargaining, versus more costly one-shot exchanges”.\textsuperscript{1045} A patent pool is an agreement between firms in an industry sector to license to one another one or more of their patents covering the technology of use in the industry. This mechanism of reciprocal licensing concessions significantly lowers the transaction costs through the institutionalization of the repeated-play nature of the exchange between them. Patent pools place future contracting in the shadow of the repeated-play bargaining game between firms.

Finally, it is worth noting an important limitation of reciprocity. In PD situations, parties do not cooperate because (in the absence of the assurance of others actors’ cooperative behaviour) they face an incentive to engage in unilateral defection strategies. In this contest, reciprocity constraints foster cooperation by altering parties’ payoff structures.

\textsuperscript{1043} See Parisi, “The Formation of Customary Law”, supra note 1029 at 13 (“in those areas of the law where custom is recognized as a primary source of law, mechanisms of automatic reciprocity are generally acknowledged among the meta-rules of the system”).

\textsuperscript{1044} “A reservation established with regard to another party […] (a) modifies for the reserving State in its relations with that other party the provisions of the treaty to which the reservation relates to the extent of the reservation; and (b) modifies those provisions to the same extent for that other party in its relations with the reserving State”.

and rendering unattainable the payoff from unilateral defection. Thus, through the imposition of a *symmetric* constraint, a non-cooperative equilibrium is changed into a cooperative equilibrium. This entails the structural limitation of reciprocity-inducing mechanisms: in situations in which the inefficient equilibrium is characterized by symmetric strategies, the imposition of a symmetric constraint does not redress the incentive misalignment between the parties. For example, in a Battle-of-Sex game situation, the inefficient equilibrium is already characterized by symmetric strategies; therefore, the imposition of a reciprocity constraint does not affect the equilibrium of the game.1046

8.1.3.3. The Process of Spontaneous Emergence of Norms

Thus far, I have identified some conditions that facilitate the SEN, but I have not identified yet the *mechanisms* through which norms emerge or whereby new norms supplant old, inefficient ones. Assume that a new social practice becomes available that enables a community to solve a social dilemma and reach a socially efficient outcome. This scenario raises the question of how a social *norm* arises that prescribes the adoption of socially efficient behaviour for each member of the group, that is, how does the society *publicize* the efficient strategy without recourse to a centralized law-making mechanism? The problem is puzzling because, as previously emphasized, it is not the collectivity as such that produces the public good in question (i.e., the norm creation and enforcement), but self-interested individuals. Thus, it must be explained why the equilibrium supported by the existence of a norm is preferable for all members of the community to the non-cooperative equilibrium, even in the absence of a centralized enforcement authority. To answer this question, it is useful to distinguish analytically between “coordination” and “incentive” problems. The *coordination* problem arises from the necessity of coordinating individual decisions to punish inefficient behaviours. That is, assuming that people have incentives to bear the costs of enforcing the norm, the coordination problem involves determining how multiple, simultaneous individual decisions to punish can be coordinated in order to generate a coherent and predictable enforcement process. In

comparison, the incentive problem arises when self-interested individuals are not willing (because they do not have the incentive) to bear the costs of punishing the norm violators.

In this subsection, I briefly discuss three explanatory hypotheses of how norms emerge and evolve spontaneously to overcome both coordination problems and incentive problems. The discussion will enable us to increase our appreciation of the conditions under which SEN can work in the real world as an efficient source of law in the absence of a centralized law-making authority.

8.1.3.3.1. Change Agents and Informational Cascades

The first question to confront concerns how self-interested individuals belonging to a community can acquire the technical and social knowledge necessary to appreciate the benefits of a superior social norm. Ellickson developed an explanatory model that proves useful in this respect. He describes the mechanism of norm production as a market for norms. The demand for norms arises from the need of group members to overcome some negative externality and induce people to internalize the costs of their behaviours. The group members can assume three distinct roles with respect to the process of norm emergence: actors, enforcers, and members of the audience. Actors are people engaging in some primary behaviour having consequences for the welfare of the group. Enforcers observe actors’ behaviours and administer punishments or rewards. Members of the audience observe both actors and enforcers and administer social rewards and punishments to enforcers in the form of either social esteem or disesteem. The supply of new norms is provided by either the actors (through their pattern of behaviour) or the enforcers (through their enforcement reaction to actors’ behaviour).

The key point is that group members are heterogeneous in (i) personal endowments, (ii) discount rates, (iii) technical intelligence (i.e., the ability to assess costs and benefits of alternative pattern of behaviours), (iv) social intelligence (i.e., the ability to forecast people’s reactions to norm change), and (v) leadership skills. Because of these differences, group members play distinct roles in the process of legal change. When some

exogenous shock in the social-economic environment triggers the need for a new superior social norm, some actors or enforcers have incentives to supply a new advantageous social norm by anticipating other group members. Ellickson uses the term “change agents”\textsuperscript{1048} to describe the actors or enforcers who first react to the exogenous shock by supplying the new norm.

Because of their superior technical, social, intelligence, and leadership skills, change agents face lower opportunity costs or higher benefits in the process of legal change. “Self-motivated” leaders anticipate that they will receive greater benefits from the new norms. “Norm entrepreneurs” promote a change in a specific norm; they possess a high level of technical knowledge that enables them, respectively, to appreciate early the cost-benefit advantage of a new social practice and to anticipate that experts among the group’s members will confer esteem on the promoter of the change. Finally, “opinion leaders” do not necessarily receive tangible benefits from the new norm (as self-interested leaders) and do not possess necessarily superior technical intelligence (as norm entrepreneurs). However, they have a high level of social intelligence, which enables them to anticipate which social change would likely receive greater social support. They observe other change agents and decide which legal change to support. As Ellickson emphasizes, “[o]pinion leaders therefore play a pivotal role in determining whether a change agent succeeds in triggering a cascade toward a new norm”.\textsuperscript{1049}

When an exogenous shock determines the social-economic conditions favouring a change of the existing social practices, the self-interested leader and norm entrepreneurs supply new norms that have the potential to improve the social welfare by governing the social dilemma more efficiently. Self-interested leaders advocate social change by undertaking exemplary acts or conferring rewards and punishments to signal the efficiency of the new norm (or the inefficiency of the existing norm). Opinion leaders, if they foresee social conditions that are favourable to promote the transition toward the new norm, start conferring esteem on change agents that have acted as promoters of the change. At this point, some members of the audience will finally be in a condition to appreciate the

\textsuperscript{1048} Ibid. at 41.
\textsuperscript{1049} Ibid. at 45.
benefits of the superior social norm. They will be encouraged by the positive reactions of the opinion leaders to the social change and by the appreciation by technical experts of the new norm. When some members of the audience begin conforming to the new norm, informational and reputational cascades will converge in determining a mass migration toward the new norm. Informational cascades occur because the mass of ordinary members of the audience will interpret the early supply of the norm as a signal of its efficiency and the social approbation of the new pattern of behaviour. Reputational cascades will reinforce the tendency to conform because ordinary audience members will avoid being socially marginalized.

8.1.3.3.2. Norm Internalization

Another explanation of the emergence of norms refers to the process of internalization of norms by target actors. This point is the subject of a vast literature investigating the moral, psychological, and philosophical reasons why people internalize norms.\textsuperscript{1050} Here I draw some conclusions about how internalization affects the incentive structure of individuals and how it helps explain the emergence of norms.

First, from an economic standpoint, internalization provides individuals with additional incentives to behave in the way the norm prescribes, thereby altering the calculus of the psychological costs and benefits of alternative strategies of behaviour. As Cooter notes, “internalization attaches a guilty penalty to violating a norm, which can change the sign of the net psychological benefits”.\textsuperscript{1051} Consider the classic agency game situation in which two parties interact strategically. The principal has to decide whether to invest or not invest, and the agent has to decide whether to appropriate or cooperate. In the absence of an enduring relationship or enforceable agreement, the agent has the incentive to appropriate; the principal anticipates the non-cooperative behaviour of the agent and consequently decides not to invest. The available cooperative surplus is lost because, in the absence of coordinating devices, non-cooperation is the dominant strategy of both


\textsuperscript{1051} Cooter, “Structural Adjudication”, supra note 338 at 1662.
actors. In contrast, the internalization of the norm that prescribes cooperation on the part of the agent reduces the payoff of appropriation, thereby rendering cooperation the dominant strategy. In the agency game, people who internalize the norm will cooperate, not appropriate, even if the objective payoff for cooperating is slightly lower than for appropriating. 1053

Second, internalization provides individuals with additional incentives to participate in the enforcement process. A person who internalizes a norm is more willing to invest resources to take part in the enforcement process for the benefit of the others. Gossip networks and ostracism are relatively cheap ways of sanctioning non-co-operative behaviour and increasing the expected costs of norm violation. Thus, the internalization of norms is a potentially effective mechanism to overcome the “second order” free-rider problem, which plague the enforcement of norms in the absence of a centralized authority.

8.1.3.3.3. Law as a Coordinating Device

Both Ellickson’s model and the theory of internalization of norms are useful because they identify two important mechanisms that lead to norm creation and enforcement. However, they still do not fully demonstrate how—maintaining the assumption of rational choice theory—rational self-interested group members conform to social norms in response to incentives created by themselves. Ellickson’s model explains how group members come to appreciate the advantages of new social practices. However, the decision to conform to norms is explained by informational and reputational cascades: people decide to conform to the new norms because they rely on others’ decisions and because they want to avoid being socially marginalized. In comparison, the theory of norm internalization certainly captures a mechanism that is powerful in generating conformity to social norms, but it largely relies on arguments outside the explanatory domain of rational choice theory. For the purposes of this study, an economic explanation

1052 Ibid at 1663. See also Michael E. Pricea, Leda Cosmides and John Tooby, “Punitive Sentiment as an Anti-Free Rider Psychological Device” (2002) 23 Evolution and Human Behavior, 203 (explaining that one function of punitive sentiments could be to recruit labor for collective actions).

1053 Ibid 1667.
is needed of how self-interested individuals surmount the “coordination” and the “incentive” problems associated with norm creation and enforcement.

Hadfield and Weingast\textsuperscript{1054} recently proposed an insightful explanation of the institutional mechanism that enables the members of a community to reach and support a cooperative equilibrium through collective punishment in the absence of a centralized coercive enforcement body. Their model \textit{links the solution of both the coordination and incentive problems to the structural attributes of law}.\textsuperscript{1055} Here I borrow some insights offered by this model to explain why the cooperative equilibrium supported by the existence of a norm is in some cases preferable for group members to the non-cooperative equilibrium based on individual defection, despite the fact that a centralized enforcement authority is absent.

It should first be recognized that in the absence of a centralized authority ensuring that norm violations are met with an effective punishment, people’s preference for norm abidance is \textit{conditional} in nature. That is, the individual prefers conformity to violation if he or she believes that (i) a sufficient number of other individuals conform to the norm, and (ii) a sufficient number of other individuals expect him or her to conform to the norm, and they are willing to punish him or her in case of violation.\textsuperscript{1056} Let me refer to these two distinct expectations as the “empirical expectation” (the others conform) and the “normative expectation” (the others expect conformity and are willing to punish).\textsuperscript{1057} In my understanding, Hadfield and Weingast’s model demonstrates that individual participation in the collective enforcement of norms reinforces the empirical and normative expectations of the group members, thereby allowing the norms to exist. Let me clarify this point by explaining how norms allow individuals to overcome the problems of coordination and incentive.


\textsuperscript{1055} \textit{Ibid} at 476 (“An important implication of our model is that it provides a link between the attributes of legal order that many intuitively associate with law and the resolution of the coordination and incentive problems that underpin effective collective punishment”).


\textsuperscript{1057} Here I use a simplified version of the more elaborate terminology introduced by Bicchieri (see \textit{supra} note 1056).
First, norms work as a coordinating device because they provide community members with a *common logic* for the identification and classification of wrongful conduct.\textsuperscript{1058} Hence, the supply of a common normative logic enables individuals to coordinate with each other in collectively administering costly penalties to norm violators.\textsuperscript{1059} In fact, the existence of a norm allows for “public and impersonal” reasoning, based on the basis of which individuals can make simultaneous decisions about whether to punish those who engage in non-cooperative behaviours. In brief, norms solve the coordination problem by coordinating individual expectations about how behaviours can be classified for the purpose of the enforcement process.

Second, the enforcement of norms reinforces the expectations of group members *about the likelihood that the inefficient (non-cooperative) behaviour will be punished*. That is, participating in collective punishments allows group members to signal to potential norm violators that the inefficient behaviour is considered a wrongful act, and as such, it is punished according to the common normative logic established by the norm. The enforcement of sanctions reinforces the empirical and normative expectations of potential wrongdoers (and therefore their compliance), thereby raising the expected value of participating in collective enforcement. This is the fundamental reason that self-interested individuals are willing to incur the costs of inflicting punishment, even when the costs of a single, inefficient behaviour does not affect them personally: people *have an incentive to participate in the collective punishment, enabling them to affect other people’s beliefs about (i) the likelihood of punishment in case of violation, and (ii) the likelihood of enforcement by other group members*. The act of enforcement is an efficient signal to other group members that the equilibrium based on the observance of the pattern of behaviour prescribed by the common logic corresponds to each individual’s private interest.

Two clarifications are needed. First, norms can work effectively as a coordinating and incentivizing device because of their institutional characteristics as *public* and *shared* rules of behaviour. The decentralized collective punishment could not function

\textsuperscript{1058} See Hadfied and Weingast, “What is Law?”, *supra* note 1054 at 473 (“A normative classification scheme that designates some actions as ‘wrongful’ (punishable, undesirable)”).
\textsuperscript{1059} *Ibid* at 479 (“A logic is an institution, not a disembodied classification scheme”).
effectively if norms did not possess the specific institutional features enabling them to publicize a normative classifying logic. Hadfield and Weingast identify the attributes of norms as generality, stability, clarity and impersonality. Second, because the common logic supports the incentive to participate in collective punishment (even when the consequences of the single violation are suffered by others), the individual has a diminished incentive to free ride on others’ enforcing efforts. Crucially, the individual incentive to punish norm violators is linked to the effect of the enforcement of sanctions imposed on other people’s beliefs. The failure to administer punishment would send the signal that the equilibrium with punishment is no longer sustainable or individually advantageous, thereby diminishing the deterrence of violations and the incentives to punish. Thus, norm beneficiaries have the incentive to enforce sanctions because they directly internalize (part of) the costs of not punishing norm violators. In brief, the collective action problem exists, but it is mitigated by the incentive to signal the efficiency of the cooperative equilibrium.

8.1.4. The Limits of Spontaneous Norms

Having examined the advantages of spontaneous law and the conditions of its emergence, I now briefly examine the limitations of SEL. First, in the absence of centralized coordinating devices, limited information and bounded rationality might cause individual actors to follow inefficient norms. Information and rationality problems can result in the emergence and consolidation of norms based on the suboptimal use of available information and on the failure to aggregate relevant information. In these cases, social norms can stabilize Pareto-inferior states. Second, the lack of a centralized lawmaker in some cases might exacerbate the tendency of some groups to consolidate standards of behaviours that maximize the welfare of the group at the expense of group outsiders. That is, some norms impoverish outsiders more than they advantage insiders. Examples are
norms of loyalty among members of criminal groups (which strengthen the groups at the expense of outsiders) or norms governing anticompetitive practices undertaken by members of a cartel (which impair competition and extract consumers’ surplus). Third, in some norms, efficiency depends greatly on the number of people who follow the norm, and there is no guarantee that a decentralized process will induce a sufficient number of people to adopt the rule. Because of costly information, parties may not find a way to cooperate and set up an efficient rule. Furthermore, as emphasized earlier, the simultaneous adoption of a new norm by a sufficient number of people might confront the failure of collective action. In short, the spontaneous emergence of law can be affected by serious problems in the public good, both in the formative process and at the stage of enforcing rules, which can result in the suboptimal production and enforcement of norms.

Fourth, and relatedly, SEN is an evolutionary process characterized by randomness and path dependency, which might lead to the stabilization of inefficient norms. Fifth, SEL is subject to the rent-seeking pressure of special interest groups. These two latter points are discussed, respectively, in sections 3 and 4. Finally, the path-dependence effect is exacerbated by the presence of informational cascades. In fact, because of limited information, individual actors might rely excessively on the signal generated by the choices previously undertaken by others although an inefficient first choice might generate a cascade of mistaken subsequent decisions. Furthermore, when people conform to the social norm in order to avoid reputational damage, despite the knowledge that the behaviour is inefficient, reputational cascades can generate negative-sum games with the attendant waste of resources.\footnote{Eric A. Posner, Law and Social Norms (Cambridge Massachusetts: Harvard University Press, 2000) at 40-42.}

### 8.2. Social Choice Efficiency

Like other collective decision-making institutions, bargaining among private actors is susceptible to cycling. To explain this phenomenon, game theory developed the concept
of the *empty core* to demonstrate that under certain circumstances, if private bargaining lacks a core, the negotiating parties might be unable to reach any stable contractual agreement. Consequently, this pattern may cycle indefinitely. In this subsection, I briefly (i) identify the conditions under which the problem of the empty core among private actors arises and (ii) discuss whether SEN might have a comparative advantage (or disadvantage), relative to legislatures, in solving the problem of cycling.

The problem of the empty core occurs in the scenario in which *n* private actors can do something either individually, by forming a coalition or by forming smaller trade coalitions.\(^{1064}\) The possibility of forming *coalitions* and *re-contracting* the bargaining outcome is crucial to explain the problem of the empty core. Each person chooses to form a trade coalition if and only if the coalition outcome makes him or her at least as well off as he or she could have achieved alone or by entering into another trade coalition. This implies that the best outcome available to a coalition sets a limit on the set of outcomes for which the individual is willing to leave the *actual* coalition and join an alternative *superseding* coalition. In the language of game theory, each outcome that is preferred by some other outcome is said to be “dominated” by the preferred outcome. Thus, each individual will be willing to dissolve the actual coalition if and only if the trade outcome achievable with the superseding coalition dominates the outcome achievable with the actual coalition. In essence, each individual chooses the coalition that enables him to gain an “undominated” outcome, that is, an outcome that is preferred to any other outcome that would be available by choosing any other possible coalition. The set of undominated outcomes is referred to as the “undominated” set.

It follows from the preceding explanation that the bargaining process finds a *stable equilibrium* only when the negotiation leads each individual to reach an outcome included in his or her undominated set. The intersection of the *n* undominated sets is referred to in game theory as the “game’s core”. The critical point is that the bargaining process finds a *stable equilibrium if and only if the game’s core is not empty*. That is, negotiation among *n* individuals finds equilibrium when it reaches an outcome included

\(^{1064}\) Lester G. Telser, “The Usefulness of Core Theory in Economics” (1994) 8 2 Journal of Economic Perspectives 152 (“the best outcomes available to a coalition set lower bounds on what its members would be willing to accept as participants in the whole market.”)
in the core. In contrast, in the case of the *empty* core, the negotiation among private decision-makers will *endlessly* cycle. Indeed, if the bargaining lacks a core, whenever a trade coalition selects a dominated outcome, a superceding coalition forms to replace the outcome by a dominant alternative. Each equilibrium reached through private bargaining is *unstable* because the individual with a sub-optimal payoff will be willing to defect to a different coalition.

A simple example serves to clarify the point. Assume three individuals A, B and C, who bargain around three possible trade outcomes corresponding to three alternative states of affairs x, y, and z. The three states of affairs correspond to three different allocations of a fixed amount of money to A, B and C. Contracts are not enforceable, and they can be re-contracted without limitation. In the initial state of affairs x, A and B enter into a contract. Suppose that the preferences of the three individuals are structured as follows:

A: \( x > y > z \)

B: \( y > z > x \)

C: \( z > x > y \)

Because both B and C prefer \( z \) to \( x \), they enter into a new contract to their mutual advantage, moving from state \( x \) to state \( z \), in which A loses his business with B. In the new state of affairs \( z \), because both A and B prefer \( y \) to \( z \), they enter into a new contract to their mutual advantage, and move from \( z \) to \( y \). In the new state of affairs \( y \), both A and C prefer \( x \) to \( y \); thus, they enter into a new contract to their mutual advantage, and move from \( y \) to \( x \), which constitutes the initial state of affairs in which B has an incentive to enter into a new agreement with C, and the process continues indefinitely. The logic of this example can be extended to coalitions with \( n \) individuals.

As the example illustrates, in *redistributive* contests the problem of the empty core among private actors arises. These are contests in which private actors redistribute wealth among themselves but do not produce it. In the language of game theory, games of
distribution are games of pure conflict.\textsuperscript{1065} One situation in which distributive games arise occurs when private actors negotiate to distribute among themselves the benefits or the costs of an “unanticipated windfall or loss”.\textsuperscript{1066} It is crucial to emphasize that in these types of situations, the problem of the empty core arises because parties negotiate \textit{ex post}, that is, \textit{after} the event that generates the redistributive game has occurred. In this situation, the only way to increase the individual payoff is to appropriate parts of others’ surplus. For this reason, in any existing trade coalition, the excluded losing party has the incentive to lure some coalition members into a new agreement that is mutually advantageous for both the excluded party and the defectors.\textsuperscript{1067}

When private bargaining lacks a core, the stability of the outcome can be restored through the creation of \textit{ex ante} legal rules that break the cycle by placing limits on the bargaining process. That is, the solution rests on the creation of cycle-breaking mechanisms \textit{before} the facts arise that generate the redistributive game. The \textit{ex ante} perspective allows parties to devise cycle-breaking rules because they do \textit{not} know how they will be personally affected by the rule; for this reason, they face an incentive to generate utility-maximizing rules. An example of a cycle-breaking mechanism is the liability rule for breach of contract, which confers the right to expectancy damages. Under a liability rule regime, the parties to a contract do not have the incentive to breach the agreement because any gain from defection will be consumed by contract damages.

It is important to recognize that \textit{ex ante} cycle-breaking legal rules can be produced through centralized law-making processes, but they can also emerge \textit{spontaneously}. One might wonder how private bargaining (which is prone to cycles if a core is lacking) may have the capacity to generate the spontaneous emergence of cycle-breaking rules. Is there any contradiction? The answer is simply that \textit{the problem of the empty core is a function of the ex-post law-making dimension, not of the spontaneity of law emergence through repeated practice}. As explained in Chapter 3, \textit{ex ante} mechanisms enjoy a comparative advantage with respect to the problem of the empty core because they enable the

\textsuperscript{1065} Cooter, \textit{The Strategic Constitution}, supra note 219 at 58.
\textsuperscript{1066} This point is discussed in Stearns and Zywicki, \textit{Public choice}, supra note 500 at 118.
\textsuperscript{1067} Stearn observes that a liability rule might not solve the problem of cycling; in fact, under a regime of liability rule, the instability might simply shift to the pre-contractual stage, see Stearns, “The Misguided Renaissance”, \textit{supra} note 330 at 1244.
formation of (anti-cycling) rules before the facts occur that generate the problem of the empty core. That is, the ex ante perspective creates incentives for the participants to use the law production process to select efficient rules that prevent future bargaining to cycle indefinitely. For this reason, a community of people confronted with a frequent empty core problem can devise and support the emergence of a rule that ensures stability for future trade coalitions by imposing limits on the re-contracting process. For example, a customary rule of fairness, prescribing the equal division of value in the case of unanticipated windfall or losses, could work efficiently as an anti-cycling legal mechanism.  

The most relevant question for the purpose of this study is whether the spontaneous emergence of cycling-breaking rules has comparative advantages relative to their production through legislatures and bureaucracies. In this respect, Stearn argues that custom can solve the problem of the empty core “when little is at stake”, but “in more serious cases” ex ante default rules are required (he implicitly refers to legislative default rules). However, Stearn does not provide any reason for why custom would work effectively only with lower interests at stake and why legislators would enjoy comparative advantages with higher stakes. There is no reason to believe that politics would work better than spontaneous emergence of law in devising and establishing cycling-breaking rules. If the law-making process operates from an ex ante perspective, and the lawmaker anticipates the problem of the empty core for future bargaining, then the lawmaker has the incentive to devise rules enabling future contracting parties to reach a stable equilibrium. As Farber and Frickey observe, “the existence of massive cycling provides the basis for a new form of equilibrium, adopted precisely in order to avoid the cycle”. With regard to the ex ante production of default cycle-breaking rules, the relative advantages and disadvantages of politics versus SEL are the same as in other regulated environments. These advantages concern productive, agency and adaptive efficiency, but there is no social choice advantage of politics over SEL. The critical point

is that the *ex ante* dimension determines a relative advantage of *ex ante* processes over *ex post* processes, regardless of whether they are “political” or “spontaneous” in nature.

### 8.3. Agency Efficiency

The centralization of the law-making process, as explained in Chapter 4, dramatically increases the level of agency costs associated with the creation of law. This effect is strictly related to the inefficiencies associated with the principal-agent relationship between the centralized lawmaker and the people subject to law. From this perspective, SEN has a comparative advantage relative to centralized mechanisms. Because, as earlier emphasized, norm producers coincide with norm beneficiaries, *there is no principal-agent relationship between the lawmaker and the people subject to law*. This represents one of the major comparative advantages of SEN over centralized law-making processes.

To clarify, SEL is not immune from rent-seeking pressures or from the related risk of norm manipulation. When a customary rule has emerged and is enforced by the members of a group, it produces binding effects on the community members, similar to the case of legislative rules. Even if the enforcement mechanism is decentralized, the fact remains that the norm violation is punished and a sanction is administered. A necessary condition for a norm to emerge and persist is that a *minimum* number of community members are willing to enforce the norm in cases of its violation. The proportion of the enforcing actors (relative to the community as a whole) that are required to support the emergence of a norm may vary depending upon various circumstances; in any case, norm emergence does not require unanimous support by community members. It follows that, although customary law and social norms rest on widespread consensus, a portion of the population remains, who is subject to the rule but whose preferences or interests are divergent from those underlying the existing norm. This means that despite the absence of a principal-agent relationship between the lawmaker and the people subject to law, there is room for cost externalization by one group at the expense of others. This opens the door to rent-seeking pressures by special interest groups that have the capacity to influence the emergence of a norm with an in-built bias in their favour.
The recognition that organized interest groups may have an incentive to exert pressure on the emergence of norms does not entirely explain why norms and customs are vulnerable to particularistic pressures. SEN requires widespread consensus and people’s willingness to act as norm enforcers, which greatly limits the scope of rent-seeking. The explanation is that the vulnerability of spontaneous norms to particularistic pressures rests on the bounded rationality and limited information of individuals.\textsuperscript{1071} As explained above, the process of norm emergence is triggered by a minority group of change agents that influence the flow of information that is accessible to the general audience. Change agents create the conditions for informational cascades and bandwagon effects that are at the heart of norm emergence. It is easy to see that when interest groups, which enjoy organizational superiority over the unorganized members of the general audience, are in a position to control or influence the change actors, and they can also influence the process of norm emergence.

8.4. Adaptive Efficiency

Norms emerge and persist over time if norm beneficiaries take part in collective enforcement: a critical number of enforcers enable a group to adopt a new norm. If the number of enforcers of a new norm does not reach the critical threshold, then the community remains with the adoption of the existing norm. The willingness of group members to participate in collective punishment depends upon enforcement costs compared to the expected benefits of enforced norms.\textsuperscript{1072} This section discusses the limitations of SEL in terms of adaptive efficiency, focusing first on the cost structure of norm emergence and then on the benefit structure of the spontaneous adaption to new norms.

\textsuperscript{1071} Parisi, “The Formation of Customary Law”, \textit{supra} note 1029 at 22-23.
\textsuperscript{1072} To clarify, I implicitly assume here that the marginal costs of participating in decentralized enforcement does not exceed the marginal costs of participating in other types of collective law-making processes.
8.4.1. Enforcement Costs

The number of actors taking part in the collective punishment of norm violators is a negative function of the enforcement costs.\textsuperscript{1073} Let me call the individuals that participate in the collective enforcement process for any given existing level of enforcement cost “enforcing actors” or “enforcers”. Now assume that the marginal enforcement cost (i.e., the enforcement cost born by the marginal enforcer actor) is a negative function of the number of infra-marginal enforcers.\textsuperscript{1074} That is, for every level of enforcement actors, there is a level of enforcement costs that individuals must be willing to bear in order to participate in the enforcement process.

Now since (i) the number of infra-marginal actual enforcers depends upon the marginal actual enforcement cost and (ii) the marginal actual enforcement cost depends upon the number of infra-marginal actual enforcers, a circular, self-reinforcing, causal mechanism affects the emergence of norms. In fact, a marginal increase in the number of enforcement actors may reduce the marginal enforcement cost and attract new enforcers, thereby triggering a self-reinforcing dynamic that facilitates the spontaneous evolution of the norm.\textsuperscript{1075} However, the same mechanism also can lead to and reinforce an evolutionary stall because a marginal decrease in the number of enforcing actors can trigger a circular self-reinforcing dynamic that reinforces a fall in the number of enforcement actors.\textsuperscript{1076} Whether this self-reinforcing mechanism facilitates or reduces the spontaneous emergence of a norm depends upon a host of factors affecting the shape of the curves depicting the two functions mentioned above (i.e., marginal costs and number of enforcement actors).

\textsuperscript{1073} Cooter, “Structural Adjudication”, supra note 338 at 1670
\textsuperscript{1074} This is consistent with the fact that people’s preference for cooperation is conditional on other people’s cooperative behaviour. Cooter makes the same assumption in Cooter, “Structural Adjudication”, supra note 338 at 1669-1675. In reality, the individual decision to participate in the enforcement process depends upon the belief that (i) a sufficient number of other individuals enforce the norm, and (ii) a sufficient number of other individuals expect him or her to enforce the norm.
\textsuperscript{1075} If the number of enforcement actors increases, the marginal enforcement cost decreases; the lower the marginal enforcement costs are, the higher the number of enforcement actors, which facilitates norm emergence and persistence over time.
\textsuperscript{1076} If the number of enforcement actors decreases, the marginal enforcement cost increases; the higher the marginal enforcement costs are, the lower the number of enforcement actors, which impedes the evolution of the norms.
In order for a norm to be self-reinforcing, the number of infra-marginal enforcers (determined by the level of enforcing costs) should not be less than the critical number of enforcers necessary to support the new norm. All else being equal, if this condition is respected, the self-reinforcing mechanism enables the norm to emerge and persist over time. On the contrary, if the number of actual infra-marginal enforcers is below the critical numbers of enforcers, than the norm is not supported by a self-reinforcing mechanism, and it is prevented from coming into force. Initially, when changing actors introduce a new norm, the number of actual infra-marginal enforcers is below the critical number of enforcers. The process of norm emergence that I have illustrated above enables the community to reach gradually the critical level of self-enforcement that is required to support a new norm. Here I am interested in the environmental conditions that facilitate this process. The literature on this point is thin, and much work remains to be done. It seems to me, however, that the previous analysis of the formation process of SEL provides some useful hints.

The attainment of the critical number of enforcers depends upon (i) the elasticity of the number of the enforcement actors with respect to enforcement costs, and (ii) the elasticity of the enforcement costs with respect to the number of enforcement actors. Some important factors determining (i) and (ii) are likely to be (a) the cost of information in the regulated environments and (b) the receptiveness of general audience members to new technical information. First, in an environment where the costs of disseminating information among group members are relatively low, it is easier for change actors to create the conditions for a cascade toward a new norm. Close-knittedness reduces the information costs, thereby increasing the elasticity of the number of enforcing actors with respect to the enforcement costs. Second, the receptiveness of group members to new technical information may depend on many factors. It is likely that the internalization of norms increases the costs for change actors to challenge the status quo and retards the propensity of group members to appreciate new technical information.1077 Thus, in this latter respect, close-knit environments characterized by strong norm internalization are likely to be less adaptively efficient than those with weaker norm internalization are.

1077 Ellickson, “Evolution of Social Norms” supra note 1047 at 56.
Another factor that undermines people’s receptiveness to new technical information is that people are generally loss averse; they value prospective losses and disadvantages more than they value prospective gains and advantages.

In conclusion, the adaptive efficiency of SEN depends upon the costs of information in the regulatory environment and people’s receptiveness to new technical information. These two variable characteristics are influenced by many factors, such as the degree of norm internalization or loss aversion. The complex interaction of these variables determines, respectively, the elasticity of the number of enforcement actors with respect to enforcement costs and the elasticity of the enforcement costs with respect to the number of enforcement actors. These degrees of elasticity relate to the adaptivity of norms to the demand for regulation.

8.4.2. Moving from a Local to a Global Optimum

One of the greatest obstacles to the spontaneous evolution of Pareto superior norms is associated with the costs of moving from a local to a global optimum. This type of evolutionary stall occurs when the set of individual preferences is non-convex. When a preference set is convex, a single peak represents the global optimum point; therefore, each local improvement reduces the distance to the point of the global maximum. That is, each local improvement leads to a global maximum. On the contrary, when a preference set is non-convex, there are several peaks, each of which represents a local optimum with respect to the local set of preferences, but only one is the global maximum (i.e., the efficient norm). This latter situation is highly problematic because when the group moves away from the local optimum, it experiences a decline in social welfare. If the group keeps moving away from the local optimum, it will reach a point of increasing returns; beyond this point, the move toward the efficient norm will be facilitated by the increasing net benefits of the legal change. However, before reaching the point of increasing returns, it is likely that many group members who face immediate utility losses will abandon the group and decide to adhere to the status quo (i.e., the local optimum).

An example of evolutionary stall caused by the difficulties of moving away from a local optimum is the transition of national states toward the adherence to international measurement standards. The adoption of international measurement standards generates efficiency gains in the long run because it reduces the transaction costs associated with international commercial exchanges or transactions. However, abandoning the local measurement standards generates immediate adaptation costs that individuals might be reluctant to sustain on a voluntary basis. To illustrate this point, Cooter provides the example of the possible transition from the British system of weights and measures to the metric system or from the British rule “drive on the left” to the rule, which is more common in other countries, “drive on the right”. Even if these transitions would probably generate efficiency gains for the country, they face large resistance because upfront utility losses discourage voluntary transitions. There are many other examples of local rules that have become inefficient with the increasing degree of integration of the European Union.\footnote{On this point see Jozef Niznik ed., The Normative Environment of European Integration Social, Political and Cultural Obstacles to Compliance with European Norms (Warsaw: Ifis Publishers, 2008).

In conclusion, in the absence of centralized coordinating devices, the possibility of overcoming evolutionary traps, such as those described above, largely depends upon the group members’ abilities to communicate cheaply with one another and coordinate a collective move toward the efficient legal regime. In these situations, for the reasons discussed in section 4.4, centralized institutions enjoy comparative advantages in terms of adaptive efficiency, relative to spontaneous mechanisms.

### 8.5. Illustrative Examples

The discussion will have clarified that the use of judicial, political and bureaucratic processes involves significant costs in terms of output ineffectiveness, biases in favour of rent-seeking minorities, legitimacy concerns and low adaptiveness. An important research question concerns whether, and to what extent, the shift of law-making activity toward spontaneous law-making processes can mitigate the inefficiencies of alternative
sources of law. In this respect, the analysis conducted in this chapter, which demonstrates that under certain conditions SEN can meet efficiently the demand for law, suggests that the allocation of portions of law-making activity to spontaneous production processes could be an attractive institutional arrangement. Nonetheless, a comprehensive discussion of this fascinating issue would carry us too far afield. However, before proceeding, it is worth mentioning a few examples to illustrate that a better allocation of law-making activity between spontaneous law-making and other sources of law might allow for significant efficiency gains.

*International Commercial Law*

In the second half of the twentieth century, the expansion of transnational activity was under increasing pressure by the state-supplied legal institutions governing international trade. In fact, national regimes proved inadequate to meet the increasing demand for the efficient regulation of transnational commercial activity. During this period, “transnational economic actors increasingly [took] for granted the notion that national regimes [made] it more, not less difficult for them to achieve efficiency and predictability in their relations with one another”.\(^{1080}\) For this reason, during the final decades of the twentieth century, transnational commercial actors generated a new transnational legal system that was based on international customs, usages and practices as well as on international commercial arbitration. This private legal system has been called the new *lex mercatoria* because of its structural resemblance to the medieval Law Merchant: it is a self-regulatory regime of internal commercial trade that is global in reach and has emerged from merchant practices.

For our purposes, two elements are worth noting. First, actors who trade across borders have created this new legal order to avoid the costs and inefficiencies of a system based on national laws and jurisdictions. Second, the new *lex mercatoria* is the result of a *private* law-making process. Trade actors increasingly refer to the principles of the *lex mercatoria* through explicit choice-of-law contractual provisions and select international commercial arbitration procedures (as a method of adjudicating their future disputes) through choice-of-forum clauses. Third, the aim of the new *lex mercatoria* is to develop a

---

unified set of contractual rules regulating transnational commercial activities. Crucially, the standardization of contractual practices is pursued through a bottom-up process, based on the choices of norm beneficiaries and on the work of dispute resolution procedures administered by merchants chosen for their expertise. Thus, the *lex mercatoria* offers an example of legal unification attained without law-making centralization. Fourth, the spontaneous nature of law-making allows for the production of substantive legal rules based on trade usages and practices. This body of law has proved more effective than national contract laws in meeting the needs for certainty and efficiency in international trade. In particular, *lex mercatoria* has proved much more adaptive than national civil and commercial codes to the evolving structure of market relationships. Finally, a system of private, competing, transnational arbitrators has evolved, which has enabled traders to choose among a range of alternative adjudication mechanisms to litigate transnational contract disputes.

*Industry Self-regulation. The Case of International Product Standards*

Thus far, I have assumed that spontaneous law-making rests on decentralized processes of norm creation. However, in some cases, spontaneous law-making can also take the form of an *ex ante* centralized production of law. The classic example is industry self-regulation, in which firms belonging to the same industrial sector entrust the production of norms regulating their economic activity to centralized law-making bodies that act as “private legislatures”. *Ex ante* centralized spontaneous law-making processes allow for the economic advantages of *ex ante* centralization (e.g., economies of scale and scope, legal uniformity, and so on), while, at the same time, reducing the costs typically associated with political law-making (rent-seeking costs, maladaptation costs and low adaptiveness). First, self-regulating actors have superior knowledge of the regulatory issues at stake, relative to political legislators: producers operating in the industry sector are likely to possess greater competence than politicians do in industry production processes.¹⁰⁸¹ Second, “[t]he delegation of law-making powers to the industry improves

---

the quality of law by enhancing the responsiveness of regulators to the *uncertainty* that is inherent in the implementation of institutions". Therefore, the greater the uncertainty in the regulatory environment, the greater will be the expected benefit of self-regulation in terms of the enhanced flexibility of regulation.

These informational advantages are derived at the cost of opportunistic behaviour by industries, especially in economic sectors where there is a sharp divergence of interests between consumers and producers. Indeed, self-regulation is not immune from the agency problems associated with the centralization of law-making. In particular, industry-made regulation entails systematic regulatory biases toward industries, which might be detrimental to consumers to the extent that the latter suffer from organizational disadvantages. However, compared to political processes, the agency problems associated with self-regulation are mitigated by the greater *proximity* of the lawmaker to the people subject to law. For example, the phenomenon of rent extraction associated with political law-making (which I have discussed in sub-section 5.3.2.3.) is drastically reduced in spontaneous law-making processes. Furthermore, private legislatures, unlike public ones, internalize the costs and benefits of the legal rules they produce. This provides an incentive to contain opportunistic behaviour.

An example of industry self-regulation is the creation of product standards by industries to solve compatibility problems in internationally integrated goods. The International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and other private international organizations produce global standards aimed at overcoming technical barriers generated by technological incompatibilities. Examples include differences in voltage standards for electric devices, differences in television broadcast formats, software written for one operating system that is incompatible with

---


1083 Ibid at 522.

1084 In particular, self-regulator entities face the incentive to limit the bias in favour of producers below the level of costs that might trigger a consumers’ reaction.

other operating systems, and so on. These incompatibilities can be the source of economic inefficiencies and welfare losses.\textsuperscript{1086} For example, the presence of technical incompatibilities reduces the variety of integrated market goods; in addition, it increases production costs by hindering the attainment of economies of scale and prevents the exploitation of network externalities. Many problems of technological incompatibilities are solved through the work of private international standardizing bodies (e.g., ISO and IEC). These are examples of how spontaneous private law-making can successfully meet the demand for regulation. Moreover, the case of international standardization illustrates the limitations of public legislative authorities in dealing efficiently with the demand for regulation and the complementary role played by spontaneous law-making processes.\textsuperscript{1087}

Finally, the example of international standardizing bodies also illustrates the limitations of private law-making. First, the illustrated problem of non-convexities often arises in the setting of industrial standards, especially when in cases of technical barriers between old and new technologies. Manufacturers of the old technology tend to oppose the introduction of new compatibility standards that would displace (or increase the retail prices of) products based on the old technology. When convergence on a superior standard is possible only through the cooperative behaviour of the market incumbents, the opposition of the latter might prevent (or retard) the attainment of appropriate compatibility standards. Second, private international standardization can generate significant externality problems. Standards organizations are privately funded, which exposes standardizing bodies to the risk of capture by a restricted number of powerful industries. In turn, this exposes standardization activity to the risk of being a means of restructuring competition. Third, standards organizations are exposed to a pervasive collective action problem. Because the benefits from standardization are indivisible, firms will prefer to let others sustain the costs of the collective effort toward product standardization.

\textsuperscript{1086} Ibid.
\textsuperscript{1087} On this point, see Sykes, “Product Standards”, supra note 1085 at 87-109 (discusses the legislative approach in Europe, the role of the judiciary in Europe and the US federal system with respect to the regulation of product standards).
Medical Malpractice
Compared to agencies and bureaucracies, spontaneous law-making can benefit from greater technical expertise concerning the regulatory issue at stake. This is because the actors operating in the regulated environment participate directly in the production of law. Consider, for example, the area of medical malpractice, which is characterized by high levels of technical complexity and the regulation of which requires scientific knowledge. In many jurisdictions, negligence standards for medical malpractice are based on a doctor’s compliance with medical custom as established by professional associations. Doctors’ informational advantage over potential alternative regulatory entities, coupled with their incentives to establish efficient precautionary standards in response to litigation pressures, probably explains the reliance on medical custom in establishing negligence standards. In the area of medical malpractice, the delegation of the regulatory authority to agencies would dramatically raise agency costs, without significantly increasing the technical expertise or information available to the lawmaker.

Intellectual Property Rights
The protection of intellectual property rights is another area of law in which SEN and PLOs have been capable of operating efficiently, sometimes managing to remedy the shortcomings of centralized law-making processes. A significant comparative advantage of SEN and PLOs in this area depends upon the high heterogeneity of creative practices, which generate a highly heterogeneous demand for legal protection, thereby raising the maladaptation costs of the one-size-fits-all regulations that are typical of centralized law-making processes. Furthermore, spontaneous law-making proves advantageous in reducing agency and adaptive efficiency.

A successful example of SEN in the area of intellectual property rights, which has been created by a community of comedians in the last half of the twentieth century, is the system of social norms that protect comedians against theft. Oliar and Springman analyzed this phenomenon, showing that it has remedied some of the shortcomings of

---

formal law and increased the level of investments in comedic generative activity. Whereas formal copyright did not offer effective protection for stand-up comedians, the informal system of intellectual property norms, created by the community of comedians, has efficiently regulated the ownership, transfer and appropriation of jokes.

Another example is that in many industrial sectors, repeated transactions among firms have generated the spontaneous creation of legal norms to protect intellectual property rights. For instance, Collective Rights Organizations (hereinafter “CROs”) set the rules of exchange for firm industries, thereby allowing for a significant reduction of transaction costs. Compared to centralized, top-down, public law-making processes, these spontaneous rule-producing mechanisms entail significant advantages. Because the outcome of these spontaneous processes is private in nature (i.e., is not collectivized), private actors can produce rules specifically tailored to their specific needs. Those who produce the rules (i.e., the firms operating in the music industry) are those who benefit from and who are regulated by the rules. This allows superior information about local preferences, lower rent-seeking costs, and more flexibility and adaptivity, compared to centralized legislative processes.

Merges provides a useful analytical framework for the study of CROs, which emphasizes the comparative advantages of these privately organized solutions compared to legislators and courts. The author discusses the example of the creation of a digital “encyclopedia of quilting” that requires the producers to buy from others and put together a large amount of material. In particular, the assembly of various materials requires obtaining legal clearance from the different owners of intellectual property rights in the various pieces of material. In this context, if individuals bargained with each owner, they would incur prohibitive transaction costs. In addition, even the organization of an electronic centralized market would not avoid the holdout problems posed by strategic sellers. An alternative solution would be the creation of a compulsory statutory license.

---

1090 I have defined the notion of collective outcome in section 4.3. Here, it is sufficient to recall that the law-making outcome is collectivized when its validity is binding for all members of the community.
1091 Ibid.
However, the statutory provision of a mandatory license would entail a host of problems. First, legislators would have to identify the content of the license (e.g., identify the material object of the license) and establish the relative royalties to be paid to owners. This would entail the difficulty of placing a value on property rights through centralized law-making procedures: politicians would seldom know if the prices they set would be the right ones. The centralized law-making process would establish a one-size-fits-all schedule of royalties that entailed maladaptation costs to the extent that the underlying structure of property rights was heterogeneous. Second, the political process would trigger rent-seeking pressures by various interest groups to influence the contents of the compulsory license. In turn, this would entail rent dissipation and rent extraction, which would deplete the economic surplus. Third, once approved by the Congress (see professor’s handwriting), the schedule of royalties is likely to become locked in. In fact, once the statute was enacted and politicians had assigned the rents to successful interest groups, the mechanism of the transitional gains trap (which I have discussed in subsection 5.4.2.3.) would prevent efficient legal change. In brief, the political law-making process would entail information, agency, and adaptive efficiency costs that would result in the depletion of part of the available surplus.

Unlike a public centralized process, privately organized institutions, such as CROs, “could probably work out something much more in tune with their needs than a [legislative] scheme of one-size-fits-all transactions”. Under the pressure of the repeated need for transactions, the actors involved would have the incentive to establish transaction-cost reducing mechanisms for setting royalties and resolving disputes: “Whatever institutional structures the content owners and users devised, they would reflect the expertise of these industry insiders. Even more important, they could be changed over time by industry participants. For these reasons, private, voluntary organizations of this kind would be superior to state-mandated compulsory licenses.”

1092 The creation of intellectual content to be protected by property rights is highly heterogeneous, and the ex ante centralized structure of the political law-making is ill suited to tailor efficiently the contents of the compulsory license.

1093 I have defined the concept of transitional-gain-traps supra in sub-section 5.4.2.3.
Finally, another interesting example of SEN in the area of intellectual property rights is the creation of patent pools in the automobile and aircraft manufacturing industries. Patent pools are a reciprocity-inducing mechanism that allows for significant comparative advantages over centralized statutorification. Patent pools enable firms to supplant the process of individual bargaining in each single transaction (based on the statutory property rule) with a privately organized mechanism that regulates repeated transactions over time. Under patent pool agreements, the royalties charged for the patented technology are established directly by patents owners, based on the relevance of the technology in the production process. This avoids the informational problems associated with centralized estimates of relative prices. As Merges emphasizes—quoting an excerpt from a Congressional patent-pooling hearing—patent pools are based on the principle that “within the industry, the individual monopoly created by patents is abolished in the form it is provided by statute and a different system is substituted more in harmony with the needs of that industry”. In this respect, the Department of Justice has recognized that patent pools can have significant pro-competitive effects and may improve firms’ competitiveness in an era of rapid technological innovation.1094

Product Liability

Spontaneous law-making might prove advantageous in setting the efficient negligence standard of the liability regime for defective products. The injury caused by a defective product is not occasional but is connected to a transaction between the injurer and the victim.1095 Because the tort occurs in the context of a seller-customer relationship, “firms’ behaviour is influenced not only by potential legal liability, but also by customer's perceptions of product risks, for the latter will affect customers’ willingness to make purchases”.1096 Here, the repeated-play element occurs between firms and the overall class of consumers. This incentive structure allows for a greater role of spontaneous law-making in the area of product liability.

1095 In his classic book about the economics of accident law, Shavell distinguishes between two different situations of firm liability: “the case where victims are strangers to firms” and “the case where victims are customers of firms”, Steven Shavell, Economic Analysis of Accident Law, (Cambridge, Massachusetts: Harvard University Press, 1987) Chapter 3 at 48 and 51 (Emphasis is mine).
1096 Ibid. at 51.
First, the reputational element is intensified in the consumer protection context. Firms know that the reputational consequences of injuring one or more consumers might be significant. Litigation cases involving product liability receive significant attention in the media, especially when they concern widely sold products. In addition, informational and reputational cascades may quickly determine changes in the public’s perception of risk. Second, given the large number of consumers involved, the wrongful conduct is more likely to be detected because monitoring costs are lower. Third, the presence of common regulatory interests among consumers facilitates their coalescence and coordination. These environmental conditions facilitate the functioning of private regulatory systems aimed at mitigating the opportunistic behaviour of manufacturing industries. Even in the absence of a state-mandated liability rule, firms are motivated to invest in product safety to the extent that harms to consumers may cause the reaction of boycotting.

The Japanese product liability system, which was in force until 1995, represents a significant historical example of a nongovernmental product liability system. Until 1995, product liability in Japan was formally subject to a general negligence regime. Despite this form of legal regime, many Japanese firms voluntarily subjected themselves to a strict liability regime. The spontaneous emergence of a strict-liability regime was based on three constitutive elements: safety standards, testing, and insurance. A public entity, the Product Safety Council, set the safety standards for a variety of products. Manufacturers submitted the products to the Council for safety assessments. If the submitted product met the safety standards, the firm could attach a “Safety Goods” (SG) label. In contrast, if the product failed the test, the firm could simply sell it without the label. Manufacturers who wanted to bundle product liability coverage with their products could submit their products to the Council for certification under the SG system, upon payment of a fee. The Council would test the products, and if they met its safety standards, SG insurance would be offered. The Council insured products with the SG label by charging an insurance premium. Under SG insurance, the Council would specify amounts to users injured by defective SG goods. Firms that did not want to bundle

---

1097 In 1973, the Diet enacted the Consumer Products Safety Act and through it established the Product Safety Council. The Act provided for a very limited mandatory regime. The Council was mandated to set safety standards for a few hazardous categories of products and was authorized to ban those products that did not meet the standards.
products-liability insurance with their products could sell their product without insurance coverage. Under this institutional framework, many Japanese firms had incentives to offer insurance coverage as a signal to buyers of the quality and safety of their products. In order to attract consumers, despite formal law establishing negligence standards for liability, many Japanese firms agreed to be subject to a regime of strict liability. The enforcement mechanism was based on a centralized public authority, but the creation of the liability regime was activated by the spontaneous choices of manufacturing industries.

Conclusions
The discussion in this chapter suggests that, under certain conditions, the spontaneous production of law has efficiency advantages over other sources of law:

(i) The decentralized nature of many (but not all) spontaneous processes enables better information about the preferences of the people subject to law; this mitigates the maladaptation costs entailed by centralized processes and generates productive efficiency advantages.

(ii) Spontaneous law-making is based upon direct participation in the production process of norm beneficiaries and target actors. This implies that the lawmaker has better factual knowledge and greater technical expertise with respect to the regulatory issue at stake.

(iii) The absence of delegation entails internalization of the costs and benefits of regulation. This reduces rent-seeking pressures and mitigates agency problems.

(iv) The proximity of the lawmaker to the people subject to law is a source of the comparative advantages of spontaneous law-making (in terms of productive and agency efficiency), even in the form of ex ante centralized private law-making. The example of industry self-regulation illustrates this point.

(v) Although spontaneous law is subject to the cycling problem, politics and other centralized law-making processes do not enjoy any comparative advantages in devising cycling-breaking rules.

(vi) In some regulatory contexts, politicians and bureaucrats do not value legal innovations as highly as private actors do; therefore, they do not have the incentive,
respectively, to pursue legal innovation or to experiment with new regulatory solutions with the speed and flexibility of private actors.

The environmental conditions under which spontaneous law-making constitutes an efficient source of law are summarized in the notion of a close-knit environment, which includes (i) power reciprocity among community members, (ii) continuing relationships and a sufficiently high discount-factor, (iii) adequate knowledge and information about the reputations of community members, and (iv) actors’ ability to promptly punish defection. Structural reciprocity, which facilitates the conditions for cooperation, is associated with environments characterized by role-reversibility. Reciprocity can also be institutionally induced. The process of the spontaneous emergence of law can be explained through the following three hypotheses: (1) change agents trigger informational and reputational cascades converging toward the new norm; (2) norm-internalization enables norm enforcement by overcoming collective action problems; and (3) legal rules supply a common logic for the normative classification of wrongful conduct, thereby operating as a coordinating device. The norms created through spontaneous law-making suffer from limitations rooted in the limited information and bounded rationality of single individuals. In particular, spontaneous norms (i) can produce negative externalities at the expense of group outsiders, (ii) can fail to evolve toward Pareto superior regimes because of evolutionary traps, (iii) can be manipulated by special interest groups, and (iv) can evolve toward inefficiency because of the bandwagon effect, as well as informational and reputational cascades.


9. THE EFFICIENT PRODUCTION OF LEGAL RULES.
INTEGRATING THE ANALYSIS

This concluding chapter formulates the core thesis of this study, identifies its main supporting arguments, and highlights the originality of its contribution to the current literature. In so doing, it emphasizes the unity of analysis throughout the chapters and synthesizes the proposed analytical framework, which systematically incorporates the outcome-process relationship into the economic analysis of law. In the following pages, I refrain from summarizing the many analytical points developed throughout the discussion; for this purpose, I refer to the summarizing conclusions at the end of (and in the main sections within) each chapter.

9.1. Core Thesis

9.1.1. Law Cannot Be Separated By Its Sources

This thesis asserts that the transaction-cost structure associated with alternative law-making processes, as well as its dynamic interaction with the characteristics of the regulated environment, are the key determinants of the efficiency of the law. The proposed analytical framework specifies variables that are predicted to be important in determining the efficiency characteristics of legal rules as they are produced through the available law-making institutions. That is, rather than inquiring into the content of substantive legal doctrines and assessing them against a hypothetical, frictionless world, this study views legal rules as the result of process choices concerning the organization of the sources of law. The theoretical underpinning of my proposed approach is the idea that the law cannot be separated by its sources: In other words, legal rules cannot be fully explained, nor consistently assessed, without careful inquiry into the way in which law is produced.

Part I of this study makes three distinct contributions to the methodological debate on legal efficiency:
(1) it provides methodological guidance for defining efficiency within the legal framework;
(2) it develops the notion of process efficiency to explain and assess the functioning of 
alternative law-making processes; and
(3) it develops a uniform taxonomy of law-making costs.

The discussion in Parts II and III provides two contributions to the comparative 
institutional analysis of law-making institutions:
(4) it compares alternative law-making institutions based on process efficiency criteria; 
and
(5) it identifies institutional and structural variables that are predicted to be important in 
determining the efficiency of the law.

Before these contributions are elucidated, three important implications of this line of 
inquiry should be emphasized.

9.1.2. Law Is Not Allocatively Efficient: Scholars Should Focus More 
on “Inefficient” Law-making

In Part III of this study, I have demonstrated that law-making institutions are generally 
il-suited to generating allocative efficiency. The environmental conditions that allow for 
the production of substantively efficient legal rules are very restrictive and unlikely to 
occur in practice. This leads me to conclude that the law is not as efficient (i.e., 
“allocatively” efficient) as some scholars have contended.1098

The degree of substantive inefficiency of legal rules varies depending upon the 
characteristics of the sources of law relative to the conditions of the regulated 
environment; however, the efficiency comparison remains one among several imperfect 
alternatives.1099 There is no such thing in the real world of law as a tendency toward 
efficiency embedded in law-making institutions. Rather, there are different ways of

bibliography cited supra note 114
1099 See Komesar, Imperfect Alternatives, supra note 7.
producing legal rules, each of which is affected \textit{differently} by pervasive information, public choice and collective action problems, which conduce \textit{differentially} inefficient outcomes. The task of analysts is to identify, in any given context, the organization of the production of law that minimizes the degree of substantive inefficiency—or, stated differently, that reduces the distance between the ideally allocative efficient outcome and the real-world outcome generated by highly imperfect law-making processes. In this sense, the intellectual efforts of legal economics scholars should focus much more on the many sources of legal \textit{inefficiency} that characterize the way in which law is actually delivered.

(Allocative) efficiency is an attribute of the idealized, zero-transaction-costs world; inefficiency, on the other hand, is the daily life of the law. This means that the inefficiency of law, rather than being conceptualized as a “departure” from the Pareto optimality, should be considered the true objective of legal-economic analysis. From this perspective, the purpose of this study is to develop a framework of analysis that enables legal scholars to distinguish both positively and normatively among various manifestations of legal allocative inefficiency.

9.1.3. Process-Efficiency Analysis Complements Output-Oriented Analysis

The methodology set out here does \textit{not} undermine the orthodox method of legal-economic inquiry; rather, it complements it. Output-oriented analysis identifies useful first-best efficiency benchmarks. However, if we agree that the law cannot be separated by its sources and that output efficiency depends upon the environmental and institutional conditions of the production of law, we should follow this logic to its conclusion, \textit{revisiting the very foundations of the efficiency claim}. The claim of the efficiency of law thus becomes a \textit{relative} one, and the point becomes one of weighing the relative advantages of available alternative law-making processes, while taking full account of the pervasive information and public choice problems underlying such alternatives and their
differential impacts on the content of legal rules.\textsuperscript{1100} This process-oriented approach does not eschew the analysis of the output; rather, it conceptualizes output efficiency as the manifestation of choice processes that are at work in the legal-economic nexus.\textsuperscript{1101}

9.1.4. The Need for a General, Systematic Theory of Law-Making

Existing studies on the economics of law-making deal with specific problems concerning the production of law. Although they often provide useful insights, they are not linked to a central analytical core. With the partial exception of Komesar’s work, there is no general, systematic economic theory that identifies the causal mechanisms underlying the interaction between the demand and supply sides of law-making. A “general, systematic” theory identifies regularities and core tendencies across many different institutional contexts. It does not solve specific problems; rather, it provides the principles in light of which these problems can be better solved.

The discussion in this essay has dealt mainly with general principles; it has not offered particular solutions for particular problems. However, solutions to many practical and theoretical problems can be found much more easily using the general principles identified in this study. If it can be demonstrated that the structural approach to law and economics adequately accounts for the relationship between the choice of law-making mechanism and the efficiency quality of the output, then the value of law and economics as a discipline will be significantly increased, since the purpose of social theory is to identify widespread regularities lying beneath superficially diverse behaviors.

To accomplish the task of developing a general theory of law-making, I have relied extensively on the existing literature on law-making dispersed across many areas of social science scholarship. The purpose of the inquiry has not been to advance the scholarship on each single issue; rather, it has sought to identify the various features of the regulated environments, as well as the alternative institutional processes that are predicted to be most relevant in affecting the efficiency of the production of law (as I


\textsuperscript{1101} See Samuels, “Nexus”, \textit{supra} note 30
have defined it). This line of inquiry has built a general, unitary, analytical framework for explaining and assessing the organization of the sources of law and its impact on the efficiency of legal rules.

9.2. Methodological Contribution

9.2.1. Part I: Path to a Definition of Legal Efficiency

(1) This study provides methodological indications of how to construct a theoretically consistent notion of legal efficiency. Since the beginning of the law and economics movement, authors have emphasized the risk of the logical circularity associated with the positional nature of the legal-economic discourse (which inevitably presupposes an normative antecedent specification). In this respect, I have taken the discussion one step further by (i) identifying the problematic implications of output-oriented definitions of legal efficiency, which conventionally overlook the right-based nature of efficiency, and (ii) considering the elements that any definition of legal efficiency should clarify in advance to avoid the risk of logical circularity and the Nirvana Fallacy. In particular, regardless of the definition of legal efficiency one adopts, three issues should be precisely identified ex ante: (i) the output that the law should maximize, (ii) the set of available alternatives for choice, and (iii) the set of relevant interests.

(2) On the basis of the above methodological indications, this study provides a notion of efficiency that enables both a positive analysis and a normative assessment of alternative law-making processes (i.e., process efficiency).

To be clear, the social choice literature has developed various normative criteria for assessing law-making processes. My prospective is different. Although I recognize the conceptual link between legitimacy concerns and the problems discussed in this study, I am less concerned than social choice theorists with the legitimacy foundations of the law-

\[1102\] See, for example, Leff, “Economic Analysis”, supra note 27.

\[1103\] Consider, for example, the normative requirements for collective decision-making identified by Kenneth Arrow, Social Choice, supra note 217.
making process. My chief concern is to identify the obstacles hindering the production of social-welfare-maximizing legal rules. To do so, I identify four fundamental problems that characterize the production of law across institutional settings and historical contexts: (i) the gathering of information and its use to produce legal rules that maximize social welfare; (ii) the misalignment of incentives between lawmakers and the people subject to law; (iii) the aggregation of individual preferences into a stable and rational collective outcome; (iii) the adaption of the production of law to exogenous changes in regulated environments. The structure of law-making processes is conceptualized as an institutional response to these four fundamental problems, which are invariably associated with the production of law. An efficient law-making process is one that (in comparison to its feasible alternatives) better facilitates the production of welfare-maximizing legal rules by minimizing the impact of these problems on the production of law.

More precisely, a law-making process facilitates the production of social-welfare maximizing legal rules when it satisfies the following four conditions:

(i) it minimizes the information costs of the law-making process and maximizes the number of individuals who decide to acquire legal information and comply with the law (i.e., productive efficiency);

(ii) it aggregates individual preferences in order to ensure the rationality of the collective outcome and to be responsive to individuals’ preference intensities (i.e., social-choice efficiency);

(iii) it is responsive to the interests of the people subject to law (i.e., agency efficiency); and

(iv) it optimizes the trade-off between legal certainty and legal change (i.e., adaptive efficiency).

(3) Based on my proposed notion of process efficiency, this study has provided a unified taxonomy of law-making costs.\textsuperscript{1104}

\textsuperscript{1104} The unified taxonomy of the law-making costs is developed in Section 2.3.
This constitutes an important contribution to the existing literature on the economics of law-making, which is dispersed across different areas with no unified taxonomy or commonly agreed-upon explanatory model. It is worth emphasizing that the identification of law-making costs is not a purely taxonomical exercise. It requires a careful and complex investigation of the incentives and constraints of all the relevant actors on both the demand and the supply sides of the law-making process.

My proposed taxonomy constitutes the first step toward a systematic analysis of the cost functions of lawmakers across different institutional contexts. Figure 34 summarizes the law-making costs typology.

**Figure 34 Unified Taxonomy of the Law-making Costs**

<table>
<thead>
<tr>
<th>Information Costs</th>
<th>Supply Side</th>
<th>Demand Side</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Information Gathering Costs</td>
<td>(1) Internal Pressure Costs (e.g., costs associated with appreciating people’s preferences and estimating the distribution of prospective cases with respect to the variables relevant for regulation)</td>
<td>(3) Compliance Information Costs (e.g., costs associated with appreciating the content of legal rules)</td>
<td>(6) Maladaptation Costs (e.g., costs associated with the inability of the lawmaker to meet the heterogeneity of local preferences)</td>
</tr>
<tr>
<td>(2) Measurement Costs (e.g., cost-benefit analysis costs)</td>
<td>(2) Costs of Reducing the Agency Costs (e.g., costs associated with the separation of powers, checks and balances, bicameralism, and federalism in the political process; <em>ex ante</em> and <em>ex post</em> control costs in the bureaucratic process; and so on)</td>
<td>(4) Compliance Conduct Costs (e.g., opportunity costs of obeying the law)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5) Participation-Information Costs (e.g., costs associated with litigation, costs associated with participation in the electoral process, and so on)</td>
<td></td>
</tr>
<tr>
<td>Agency Costs</td>
<td>(3) External Pressure Costs (e.g., rent-seeking costs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4) Outcome Agency Costs (e.g., monopolistic rent costs in the political process; bureaucratic and legislative drift in the bureaucratic process; and so on)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 34 Unified Taxonomy of the Law-making Costs
### Social-Choice Costs

<table>
<thead>
<tr>
<th></th>
<th>(1) Decision-Making Costs (e.g., bargaining costs associated with vote trading in the legislative arena)</th>
<th>(2) Decision-Making Costs (in the case of private legal ordering)</th>
<th>(3) Costs Associated with Involuntary Redistributions from Winners to Losers (e.g., costs imposed by high per capita minorities on low per capita majorities in politics).</th>
</tr>
</thead>
</table>

### Adaptive Costs

<table>
<thead>
<tr>
<th></th>
<th>(1) Dynamic Implications of Production, Agency and Social-Choice Costs</th>
<th>(2) Adaptation Transaction Costs</th>
<th>(3) Adjustment Costs</th>
<th>(4) Resistance Costs</th>
<th>(5) Maladaptation Costs</th>
</tr>
</thead>
</table>

The availability of a unified taxonomy of law-making costs provides the conceptual framework for measuring the various components of process efficiency. For example, the agency efficiency of a given law-making mechanism is measured by the sum of the agency costs (indicated in the second row of the table in Figure 34) relative to the agency costs of the available alternative law-making processes. These costs are a function of (i) the supply and (ii) the demand sides of the production of law, as well as of (iii) the characteristics of the law-making outcome. From this perspective, process efficiency corresponds to the minimization of the relative costs (of alternative law-making processes) of producing the least relatively inefficient output. The traditional methodological perspective focusing on the social-surrplus-maximizing outcome is not entirely dismissed; rather, it is radically transformed into a more complex efficiency calculation. First, the efficiency of the output is still relevant, but it is interpreted from a second-best perspective: that is, as the minimization of the difference in surpluses between the ideal allocative efficient outcome and the real outcome produced by one of the available law-making processes. Second, the outcome is a function of the other law-making costs; this implies that substantive efficiency is a function of the other determinants of process efficiency (on the supply and demand sides). Simply stated, the
difference between the conventional and the structural approaches is that the output of the law-making process is measured and assessed in terms of inefficiency minimization and is a function of law-making costs.

9.3. Comparative Institutional Analysis

Part II and Part III have focused entirely on identifying the variables predicted to be important in determining the efficiency of the law. This line of inquiry is organized along the lines of a supply-demand model. The demand side summarizes the incentives, constraints and preferences of the people subject to law, as well as the features of the regulated environment (i.e., the structural characteristics of the situations from which the need for law arises). The supply side summarizes the institutional features of the law-making process that are predicted to significantly affect the structure of the law-making costs.

The analysis of the supply and demand side of alternative law-making processes has corroborated the hypothesis (formulated in Part I) that the efficiency of the law depends upon the structure of law-making costs, which, in turn, depends upon the interplay between the demand for and the supply of law. Stated differently, only a careful inquiry into the dynamic interaction between (i) the environmental regulatory conditions; (ii) the preferences, incentives and constraints of the people subject to law, and (iii) the institutional features of the law-making process enables us to explain and assess legal efficiency.

The following two subsections synthesize the essential institutional and structural features that provide the foundation of a general unitary framework of the efficient production of law.

9.3.1. Part II: Fundamental Law-Making Dimensions

Part II develops a four-dimensional framework of law-making: Two dimensions characterize the supply side (i.e., the ex ante vs. ex post dimensions, as well as the the degree of centralization of the law-making process), and two dimensions characterize the
demand side (i.e., the degrees of frequency and homogeneity of the cases to be regulated), which are assumed to affect the structure of the law-making costs. I have referred to these four variables as the “fundamental” institutional and structural dimensions of law-making. They are fundamental in the sense that they determine the core tendencies of law-making more than any other dimension.

The framework developed in Part II provides three distinct contributions to comparative institutional analysis:

(i) it explains the mechanisms through which the ex ante versus ex post dimensions affect the productive efficiency of law-making by focusing, in particular, on the impact of legal homogeneity and frequency on the production information costs of law-making;
(ii) it provides arguments supporting the hypothesis that the ex ante versus ex post dimensions of law-making affect the other components of process efficiency (i.e., agency, social choice and adaptive efficiency); and
(iii) it distinguishes the economic effects causally connected with the ex ante / ex post dimension from those related to the degree of centralization of the law-making process; this leads to a demonstration that the “combination” of the ex ante / ex post dimension and the degree of centralization of law-making crucially affects each of the four components of process efficiency.

Let me briefly synthesize these three important points.

(i) The ex ante / ex post dimension of law-making significantly affects the production information costs of law-making.

First, although the existing literature recognizes that the ex ante / ex post dimension affects the information costs of law-making, the causal mechanisms explaining this relationship remain poorly understood. This study has provided an explanation of these mechanisms. First, the discussion is taken one step further by investigating the demand for ex ante legal precision expressed by the people subject to law. Although this point is crucial to determining the efficient allocation of law-making between ex ante and ex

---

1105 See supra note 257
1106 See 3.1.
post processes, it is not sufficiently investigated in the economic literature. I have contributed to filling this gap by developing a simple model that provides one explanation of how the level of ex ante precision of legal rules affects individuals’ compliance decisions. I have argued that the individual demand for ex ante legal specificity is a function of the expected net benefit to be gained from legal compliance, which is, in turn, a function of the individual probability of a law’s application (i.e., the probability of an individual undertaking conduct that falls within the scope of the law) and of the frequency per actor (i.e., the number of times the individual undertakes the course of action covered by the law). These findings are relevant if one considers that an excess of ex ante legal specificity incentivizes individuals to remain rationally ignorant about (and to not comply with) the content of legal rules, whether or not these rules are substantively efficient. This point has potential political economy implications. It suggests, for example, that the size of “underground economies” in many countries could be progressively reduced—and that a more efficient formal legality could be restored—through a reformation of the institutional design of the sources of law, independent of the substantive legal doctrines adopted.

Second, the discussion has provided a better understanding of the supply side of the production of legal rules by focusing on the impacts of legal homogeneity and frequency on the information costs of law-making. While it is widely recognized in the literature that heterogeneity and frequency affect the relative advantages of ex ante versus ex post law-making, many aspects of the distinction remain poorly understood. In this respect, the discussion has shown that, while frequency affects average information costs, homogeneity affects the shape of marginal information costs and amplifies the magnitude of the economies of scale effect. Alternative combinations of the degrees of legal homogeneity and frequency generate the relative advantages and disadvantages of ex ante

\[\text{\textsuperscript{1107}}\text{Therefore, the substantive efficiency of legal rules is only one part of the complex structure of incentives behind individual compliance decisions. Lawmakers can influence the magnitude of compliance-information costs (and, therefore, individuals’ compliance decisions) through their choice of the level of ex ante legal specificity of legal rules. This choice is procedural in nature. It does not affect the allocation of legal entitlements, but it has an impact on the costs of appreciating the contents of the legal rules. See 3.1.2. and 3.1.3.}\]

\[\text{\textsuperscript{1108}}\text{See 3.1.4.}\]

\[\text{\textsuperscript{1109}}\text{See 3.2.1.}\]
versus \textit{ex post} processes.\footnote{1110} In the case of \textit{frequent} and \textit{homogeneous} cases covered by the law, \textit{ex ante} law-making benefits from economies of scales, while \textit{ex post} law-making suffers from diseconomies of scale, due to what I refer to as the “repetition effect” (i.e., an increase in fixed production costs).\footnote{1111} By contrast, when facts are \textit{heterogeneous}, \textit{ex post} law-making benefits from a relative advantage over \textit{ex ante} law-making because the total fixed costs of law-making tend to decrease (i.e., the magnitude of the “repetition effect” associated with \textit{ex post} law-making is smaller than the magnitude of the increasing effect on \textit{ex-ante} fixed costs due to heterogeneity).\footnote{1112}

Third, the discussion has provided \textit{a better understanding of the causal mechanisms whereby homogeneity and frequency affect the costs of law-making}. In particular, I have linked an understanding of the economic effects of legal homogeneity to the structure of legal norms and, in particular, to the scope of the factual predicate of these norms. From this perspective, I have related legal homogeneity to the marginal productivity of \textit{ex ante} law-making.\footnote{1113} The more the demand of law is \textit{homogenous}, the greater the number of cases falling within the scope of the application of law. Therefore, \textit{holding the level of ex ante specificity of the law constant, higher homogeneity increases the marginal productivity of ex ante law-making.}\footnote{1114} The practical consequence is that, at high levels of homogeneity, it is economically efficient to \textit{anticipate} the law-making activity at the promulgation phase; on the contrary, at low levels of homogeneity (and given the low productivity of \textit{ex ante} law-making), it is economically efficient to leave the specification of the legal content to the adjudication phase.\footnote{1115} In addition, I have explained how the degree of homogeneity of the demand for law affects the level of maladaptation costs. In particular, the optimal level of \textit{ex ante} specificity depends upon the optimal trade-off between the maladaptation costs associated with the \textit{ex ante} regulation of the most

\footnotesize
\begin{footnotes}
\item[1110] See 3.2.2.
\item[1111] See 3.2.1.
\item[1112] \textit{Ibid.}
\item[1113] See 3.4.2.
\item[1114] \textit{Ibid.}
\item[1115] See 3.2.1.
\end{footnotes}
infrquent cases and the increase in production-information costs associated with the *ex post* regulation of similar cases.\footnote{See 3.4.3.}

Finally, with regard to legal frequency, the discussion of *ex ante* versus *ex post* law-making elucidates two points that under-explored in the economic literature: i) the relationship between aggregate frequency and frequency *per actor* and ii) the effect of the degree of *ex ante* precision of law on legal frequency. First, while aggregate frequency is related to the magnitude of the economies of scale effect, the frequency *per actor* (which is a determinant of the individual utility function) is related to the behavioural effect generated by norms.\footnote{See 3.5.1.} Second, aggregate frequency and frequency *per actor* are related to the legal definition of the *scope of application* of the law. Namely, by focusing on the formal definition of the scope of the legal norm, it is, in many cases, possible to predict the magnitude of the flow of legal cases that will be covered by the norm (i.e., aggregate frequency), as well as whether the regulated behaviour is concentrated in a few actors or widely dispersed among the set of the people falling within the scope of application of the law (i.e., frequency *per actor*). Put shortly, the way in which the law *defines* its scope of application affects both the aggregate frequency and the frequency *per actor* of the prospective cases to be regulated.\footnote{See 3.5.2.}

Finally, by combining the analysis of the demand for and supply of *ex ante* legal specificity, I have discussed the lawmaker’s choice of the optimal levels of *ex ante* legal specificity (i.e., the level of specificity that maximizes the number of people who choose to acquire legal information). The classical theory of monopoly provides insights on this point and helps to clarify the peculiarities of the production of legal rules, in comparison to the production process traditionally assumed by microeconomic theory.\footnote{See 3.3.}

The table in Figure 35 summarizes the main results of the discussion about legal homogeneity and frequency.
### Figure 35 The Degree of Homogeneity and Frequency

<table>
<thead>
<tr>
<th></th>
<th>Demand and Supply</th>
</tr>
</thead>
</table>
| **Homogeneity**      | (1) It reduces the magnitude of the “repetition effect” associated with the *ex post* dimension, thereby reducing the fixed costs of law-making.  
                        (2) It increases the marginal productivity of *ex ante* law-making due to lower marginal production-information costs.  
                        (3) It decreases the maladaptation costs associated with *ex ante* law-making.  |
| **Heterogeneity**    | (1) It reduces the marginal productivity of *ex ante* law-making and increases the maladaptation costs associated with prospective law-making.  
                        (2) *Ex post* law-making entails lower fixed production-information costs and lower maladaptation costs.  |
| **Aggregate Frequency** | (1) It affects the average production information costs, thereby affecting the magnitude of economies of scale.                               |
| **Frequency per Actor** | (1) It affects the expected benefit from compliance, thereby increasing the behavioural effectiveness of the law.                           |

(ii) *The ex ante / ex post dimension of law-making affects the other components of legal efficiency (i.e., social choice, agency and adaptive efficiency).*

On this point, the discussion has proceeded by linking many of the insights provided by public choice and social choice literature to the *ex ante* production of law. First, the *ex ante* dimension poses a problem of inter-temporal inconsistency, which significantly raises the agency costs associated with law-making. The subsequent discussion of politics and bureaucracy has provided supporting arguments for this point.  

Second, the *ex ante* dimension avoids the cycling problem associated with zero-sum, purely conflictual games. The subsequent discussion of spontaneous law-making has confirmed and further elucidated this point.  

Finally, the *ex ante* perspective affects the adaptive efficiency of law-making. The comparative analysis of adjudication versus politics shows that prospective legal rules tend to increase both resistance and maladaptation costs, which are sources of disadvantages in comparison to *ex post* retrospective, fact-specific rules.

---

1120 See 3.2.1.  
1121 See 3.4.3.  
1122 See 3.7.  
1123 See 3.6.  
1124 See 3.8.
(3) The second fundamental institutional dimension of law-making is the degree of centralization of the law-making process. The economic effects causally connected to the ex ante / ex post dimension should be distinguished from those related to the degree of centralization of the law-making process. It follows that the “combination” of the ex ante / ex post dimension and the degree of centralization of law-making crucially affects each of the four components of process efficiency.

As I have repeatedly emphasized, ex ante law-making can be either centralized (i.e., politics and bureaucracies) or decentralized (i.e., repeated private bargaining). The law and economics literature typically fails to analytically distinguish the cost-saving effects generated by the ex ante / ex post dimension from those associated with the degree of centralization of law-making. In this respect, the discussion in Chapters 3 and 4 has demonstrated that the economizing effects associated with ex ante law-making differs from and should not be conflated with those depending upon the degree of centralization.\(^{1125}\) The table in Figure 36 summarizes the distinct effects of these two fundamental law-making dimensions in terms of their process efficiencies, thus demonstrating that each of them differentially affects the various components of the efficient production of legal rules.

**Figure 36 Fundamental Institutional Dimensions of Law-Making**

<table>
<thead>
<tr>
<th>Productive Efficiency</th>
<th>Ex Ante versus Ex Post</th>
<th>Degree of Centralization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The ex ante / ex post dimension affects: (1) the magnitude of compliance-information costs and (2) the structure of production-information costs.</td>
<td>The centralization of law-making entails: (1) economies of scale, (2) legal uniformity, and (3) the regulation of externalities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The decentralization of law-making entails: (1) greater incentives for the lawmaker to meet people’s preferences, (2) greater incentives for the lawmaker to experiment and explore innovative legal solutions, and (3) the reduction of</td>
</tr>
</tbody>
</table>

\(^{1125}\) See, in particular, Section 4.1.
The ex ante dimension avoids the cycling problems associated with private bargaining via an empty core.

The centralization of law-making entails problems regarding the aggregation of the preferences associated with the collective nature of the decision-making process and the collectivized nature of the outcome.

The ex ante dimension poses an inter-temporal inconsistency problem, which is independent from the degree of centralization.

The degree of centralization affects the intensities of both internal and external pressures.

Prospective legal rules (which are general in nature) tend to increase resistance costs, since a greater number of people is immediately affected by the rules. It also increases maladaptation costs, since rules are necessarily applied to “classes” of cases. Both of these cost effects are a function of the degrees of frequency and homogeneity of legal demand. Resistance costs are also affected by the frequency per actor.

On the one hand, the degree of centralization may increase resistance and maladaptation costs, since the heterogeneity of the legal demand increases. On the other hand, the degree of centralization increases the ability of the lawmaker to provide a focal point around which individuals can coordinate their collective migration toward superior legal rules.

9.3.1.1. Hypothesis To Be Tested in Part II

The discussion synthesized above allows the identification of the following hypotheses, which are to be tested against the results of comparative institutional analysis.

(i) Under conditions of high homogeneity and frequency of the demand for law, the ex-ante centralization of law-making allows for significant economies of scale, which significantly reduce the average (production and compliance) information costs.

(ii) The ex ante centralization of law-making entails a significant increase in the agency costs of law-making.
The various combinations of ex ante versus ex post and centralized versus decentralized law-making affect the structure of the social-choice problem confronted by the lawmaker. However, there are no unambiguous relative advantages of ex ante centralization over ex post decentralization, or vice versa, in terms of social choice efficiency.

The ex ante centralization of law-making reduces the magnitude of adaptive and information costs associated with legal change; however, it increases the magnitude of resistance and maladaptation costs.

9.3.2. Part III: Law-Making Institutions

Parts II and III are strictly related. While Part II focused on a four-dimensional analytical framework, Part III has adopted a lower level of abstraction, thereby including a greater number of institutional and structural variables. The comparative analysis has proceeded through an examination of four stylized law-making models: (i) adjudication, (ii) politics, (iii) bureaucracy and (iv) spontaneous law-making.

The logical connection between Parts II and III can be easily understood by observing that the four models discussed in Part III (and some variants examined in the discussion) can be placed in a Cartesian space, in which the ex ante/ex post dimension is represented on the abscissa and the degree of centralization is measured by the ordinate. Figure 37 provides a graphic representation of this point.

---

1126 They differ in the level of abstraction, but the object of inquiry does not change.
Figure 37 Law-Making Institutions

- (Adjudicating) Agencies
- Adjudication in “Made” Orders
- Legislation
- (Rulemaking) Agencies
- Private “Legislatures”

- Adjudication in “Grown” Orders
- Collective Bargaining
- Private Bargaining

- Social Norms and Customs
The Introductions of each chapter in Part III have identified the institutional features that (beyond the ex ante / ex post dimension and the degree of centralization) most affect process efficiency. The table in Figure 38 summarizes these additional institutional features, which characterize the four law-making models examined in this study.

**Figure 38 The Additional Institutional Features**

<table>
<thead>
<tr>
<th>Fundamental Dimensions (Part II)</th>
<th>Additional Institutional Variables (Part III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Politics</td>
<td>(1) Ex ante</td>
</tr>
<tr>
<td></td>
<td>(2) Centralization</td>
</tr>
<tr>
<td></td>
<td>(3) Political Representation (Electoral Process + Majoritarian Decision-Making + Outcome Collectivization).</td>
</tr>
<tr>
<td>2) Bureaucracy</td>
<td>(1) Ex ante</td>
</tr>
<tr>
<td></td>
<td>(2) Centralization</td>
</tr>
<tr>
<td></td>
<td>(3) Technical Expertise</td>
</tr>
<tr>
<td></td>
<td>(4) Political Insulation</td>
</tr>
<tr>
<td>3a) Adjudication (Grown Orders)</td>
<td>(1) Ex post</td>
</tr>
<tr>
<td></td>
<td>(2) Decentralization</td>
</tr>
<tr>
<td></td>
<td>(3) Private Litigation</td>
</tr>
<tr>
<td></td>
<td>(4) Adjudicating Courts</td>
</tr>
<tr>
<td></td>
<td>(5) Doctrine of Precedent (i.e., Persuasive Value of Precedents)</td>
</tr>
<tr>
<td>3b) Adjudication (Made Orders)</td>
<td>(1) Increasing Relevance of the Ex Ante Dimension (Due to the Doctrine of Stare Decisis)</td>
</tr>
<tr>
<td></td>
<td>(2) Increasing Degree of Centralization</td>
</tr>
<tr>
<td></td>
<td>(3) Private Litigation</td>
</tr>
<tr>
<td></td>
<td>(4) Adjudicating Courts</td>
</tr>
<tr>
<td></td>
<td>(5) Doctrine of Precedent (i.e., Stare Decisis)</td>
</tr>
<tr>
<td>4) Spontaneous Law-Making (PLOs)</td>
<td>(1) Absence of Centralized Lawmaker</td>
</tr>
<tr>
<td></td>
<td>(2) Repeated Dyadic Contacts</td>
</tr>
</tbody>
</table>

As demonstrated in Part II, the efficiency of law largely depends upon the colocation of the process whereby legal rules are produced in one of the quadrants delimited by the Cartesian axis, which is depicted in Figure 37. This important—and, generally, poorly understood—core tendency of the law-making process interacts with the additional institutional features identified in Figure 38. It has been shown that further institutional mechanisms, along with a number of structural conditions, either exacerbate\(^{1127}\) or counterbalance\(^{1128}\) the core tendencies identified through the four dimensional law-making models. In this sense, there is an overlap between the two distinct levels of

\(^{1127}\) E.g., the ex ante centralization of law-making increases agency costs or counterbalances them.

\(^{1128}\) E.g., the ex ante centralization increases productive efficiency under conditions of frequency and homogeneity.)
analysis. Many mechanisms and cost effects identified through the analysis of adjudication, politics, bureaucracy and spontaneous law-making are simply manifestations of the mechanisms already identified in the four-dimensional law-making model developed in Chapters 3 and 4. By comparison, in other cases, the discussion of adjudication, politics, bureaucracy and spontaneous law-making reveals additional dynamics that mitigate or counterbalance the core tendencies previously identified. The present and following sections briefly examine this point.

(1) Ex ante centralization allows for significant economies of scale. However, in politics and bureaucracy, these should be weighed against the disadvantages associated with, respectively, (i) the high irrationality and limited information capabilities of the political decision-making process and (ii) the tendency toward (a) overregulation, (b) x-inefficiency and (c) the output ineffectiveness of the bureaucratic processes.

Part II led to the conclusion that ex ante law-making processes enjoy significant advantages in terms of productive efficiency when the legal demand is homogenous and frequent enough to allow for economies of scale. In this respect, Part III has shown some additional institutional features characterizing the ex ante centralized processes (e.g., political representation or bureaucratic insulation) may entail production costs that outweigh the economizing effects emphasized in Part II.

Chapter 5 has shown that high levels of irrationality in the decision-making process affect political law-making. With regard to the demand side, voters (i) deal with a Knightian-type uncertainty, (ii) do not have incentives to become informed and (iii) instead have incentives to behave irrationally. With regard to the supply side, politicians deal with severe informational problems. I have maintained that the institutional mechanisms designed to mitigate the information problem (e.g., information from interest groups, committee systems, delegation to bureaucracies) are unlikely to lead political law-making toward the efficient production of legal rules. Second, the legislative process is likely to be inefficient in regulatory environments characterized by (i) highly technical or (ii) highly dispersed information.
Chapter 6 has shown that the informational advantages of technocratic law-making should be weighed against the tendencies displayed by bureaucratic processes toward: (i) overregulation, (ii) cost ineffectiveness and (iii) output ineffectiveness. The environmental conditions under which the expected benefits from technocratic law-making increase - in such a manner as to outweigh the above tendencies toward productive inefficiency - are related to the level of outcome uncertainty, which is often a function of the complexity of the regulatory issue at stake. Under outcome uncertainty, the legislators’ risk aversion increases the benefit expected to derive from the delegation to technocratic law-making conducted by experts.

(2) Pure decentralization entails relative advantages in terms of productive efficiency under the restrictive conditions of close-knittedness.

Part II led to the conclusion that, at high level of legal heterogeneity, pure decentralization enjoys comparative advantages over centralized processes, since it allows for greater responsiveness of the production of law to local preferences. Comparative institutional analysis has better specified this conclusion by demonstrating that pure decentralization can generate the spontaneous production of efficient legal rules only under the very restrictive conditions of close-knittedness.1129 In particular, in order for norms to emerge and persist over time, it is crucial that the magnitude of enforcement costs is not so high as to prevent the attainment of the critical number of enforcers necessary to support the spontaneous emergence of a norm. The emergence of a self-reinforcing mechanism is facilitated by the low costs of information about group members’ histories and reputations, as well as by the high degree of receptiveness of general audience members to the technical information required to appreciate the efficiency of superior norms.1130

---

1129 See 5.1.3.1.
1130 See 5.4.1.
(3) Politics and bureaucracy exacerbate the increase in agency costs associated with the ex ante centralization of law-making.

The comparative analysis of alternative law-making institutions has confirmed the conclusion that ex ante centralization generates increasing agency costs.

First, adjudication in centralized “made orders” entails higher agency costs than adjudication in decentralized “grown orders”. Furthermore, the doctrine of stare decisis accentuates the ex ante dimension of judicial law-making, which further increases the rent-seeking pressures aimed at influencing the evolution of precedents.

Second, political law-making generates increasing agency costs because the high level of centralization entails a separation between politicians and voters, which generates both external pressures from interest groups and opportunistic behaviours from self-interested politicians. The increased pressure to capture rent and extract surpluses from the law-making process is reflected in the structure of the agency costs of political law-making. External pressures generate monopolistic-rent costs, rent-seeking costs, and rent-dissipation costs. Internal pressures generate increasing augmentation costs (i.e., the strategic increase of transaction costs by politicians) and rent–extortion costs. The ex ante perspective of politics exacerbates agency costs by creating a scope for moral hazard by politicians (i.e., agency slack). Finally, in addition to the ex ante centralized dimension, three further institutional features exacerbate the agency costs associated with political law-making: (i) the rational ignorance of the electorate as a whole, (ii) the majoritarian character of political decision-making, and (iii) the collectivized nature of the legislative outcome.\textsuperscript{1131}

Third, delegation to agencies and bureaucracies increases agency costs. Bureaucracy entails a second level of “technocratic” ex ante centralization, which adds to “political” ex ante centralization. This entails bureaucratic drift costs, as well as costs related to the strategic use of delegation by politicians, thereby exacerbating the agency problem associated with ex ante centralization.

\textsuperscript{1131} See 7.3.1.2.
Spontaneous law-making significantly decreases agency costs (although it is not immune from the agency problem). Even in the context of private legal orderings, as the degree of centralization increases and private law-making adopts an *ex ante* perspective (e.g., industry “self-regulation”), agency costs increase significantly.

(4) The social-choice problem differentially affects law-making institutions. Law-making is always path-dependant and structurally induced. The production of legal rules is subject to the risk of incoherent legal outcomes. In particular, every law-making institution is prone to the problems associated with (i) intransitive cycling, (ii) agenda influences, and (iii) path dependence. However, the different features characterizing alternative law-making institutional designs can be a source of comparative advantages for some institutions over their alternatives.

First, the “reason-based” nature of judicial decision-making is a source of comparative advantages for adjudication over politics (with respect to social choice efficiency). In many cases, the rules of legal reasoning ensure transitivity in adjudication, independent of the full transitivity of social preference orderings. In addition, other stabilizing features of the judicial process (e.g., the doctrine of precedent) mitigate the social choice problem. The adjudication process is, by its very nature, “path dependent” and subject to “agenda influence”; however, adjudication path dependence tends to be non-cyclical and evolutionary in nature, which entails advantageous economizing effects.

Second, politics is driven by a preference-based rationality, which poses a problem of cyclical, sequential path dependence. Furthermore, the collective nature of politics (i.e., collective process and collectivized outcome) associated with majoritarian decision-making exacerbates the social-choice problem. The conditions under which political

---

1132 See 5.3.
1133 See 5.5.
1134 See 6.3.1.
1135 See 6.3.2.
majoritarian decision-making can reach a stable, non-path-dependent equilibrium are very restrictive.1136

Finally, private bargaining is subject to the problem of intransitive cycling in the case of empty core bargaining. Nonetheless, there is no reason to believe that spontaneous law-making suffers disadvantages in relation to politics and other centralized processes.1137 This is because the relative advantages associated with dealing with the problem of the empty core are connected to the ex ante dimension of law-making, and private bargaining is an ex ante process.

(5) The comparative analysis of alternative law-making institutions has confirmed that the ex ante centralization of law-making reduces the adaptive and information costs associated with legal change; at the same time, however, it increases resistance and maladaptation costs.

It was argued in Part II that the ex ante centralization of law-making increases both resistance and maladaptation costs associated with legal change (to the extent that legal demand is heterogeneous and that both aggregate demand and frequency per actor increase). According to comparative institutional analysis, the political representative mechanism exacerbates this tendency for three reasons: (i) an increasing number of veto-players might block efficient transitions, (ii) an increasing number of rent-seeking groups have incentives to promote inefficient transitions, (iii) outsider minorities who are unable to block inefficiencies might be forced to bear the major costs of innovations and (iv) political short-sightedness and transitional gain traps prevent efficient legal change.1138 At the same time, however, it must be recognized that the offsetting capabilities of legislators can reduce resistance costs.1139

The adjudication process has relative advantages over politics with regard to reducing maladaptation and resistance costs, depending upon the degree of centralization of the

1136 See 7.2.3. and 7.2.4.
1137 See 5.2.
1138 See 7.4.2.
1139 See 7.4.1.
adjudicatory system. In particular, decentralized adjudicatory systems facilitate legal change for three reasons: (i) the levelling-of-the-playing-field effect associated with judicial process; (ii) the incremental nature of judicial legal change; and (iii) the fact that change-seeking actors, who are a minority in the political process, might be advantaged in adjudication.\footnote{See 6.4.1.} As the \textit{ex ante} centralized dimension of adjudication increases (due to a centralized adjudicatory system and the binding quality of legal precedents), adjudication tends to share the same disadvantages as politics in terms of increasing resistance and maladaptation costs.\footnote{See 6.1.2.3.}

The analysis of the spontaneous emergence of law-making and private legal orderings has confirmed that decentralization enjoys comparative advantages only in those environments where group members are (i) receptive to new technical information and (ii) able to cheaply communicate with one another and coordinate a collective move toward efficient legal regimes. When these conditions are absent, pure, decentralized law-making is subject to the risk of being impeded by evolutionary traps and manipulated by interest groups; hence, it evolves toward or may stabilize the emergence of inefficient legal rules.

The next four tables (Figures 39 through 42) synthesize the discussion of the impact in terms of the process efficiency of the demand- and supply-side features of alternative law-making processes. In so doing, they show that the additional institutional features identified in Part III either exacerbate or counterbalance some of the dynamics already identified in Part II.
Figure 39 Process Efficiency Characteristics of Politics

<table>
<thead>
<tr>
<th>Politics</th>
<th>Demand Side</th>
<th>Supply Side</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Productive Efficiency</strong></td>
<td>Obstacles to efficiency:</td>
<td>Three institutional features assist musicians in dealing with the Economic Calculation Problem:</td>
</tr>
<tr>
<td></td>
<td>(1) voters’ Knightian uncertainty,(^{1142})</td>
<td>(1) information received from interest groups;</td>
</tr>
<tr>
<td></td>
<td>(2) voters’ rational ignorance(^{1143})</td>
<td>(2) the committee system</td>
</tr>
<tr>
<td></td>
<td>(3) voters’ rational irrationality(^{1144})</td>
<td>(3) delegation to bureaucracies</td>
</tr>
<tr>
<td></td>
<td>(4) short-run vs. long-run legal certainty(^{1145})</td>
<td>These institutional arrangements do not conduce the production of efficient law.</td>
</tr>
<tr>
<td><strong>Social-Choice Efficiency</strong></td>
<td>Three problems of majority rule:(^{1147})</td>
<td>1146</td>
</tr>
<tr>
<td></td>
<td>(1) preference intensity problem</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) political externalities (i.e., involuntary redistribution associated with non-unanimous decision criteria)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) cycling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two stabilizing institutional features of the legislative process:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4) structurally induced equilibrium (^{1148})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5) vote trading</td>
<td></td>
</tr>
<tr>
<td></td>
<td>However, a structurally induced equilibrium is unlikely to correspond to the efficient outcome and is likely to be manipulated by the agenda setter. Vote trading is subject to strong structural limitations and restrictive utility requirements.</td>
<td></td>
</tr>
<tr>
<td><strong>Agency Efficiency</strong></td>
<td>(1) Group dominance entails:</td>
<td>(1) Agency slack (^{1152})</td>
</tr>
<tr>
<td></td>
<td>(1a) monopolistic-rent costs (^{1150})</td>
<td>(2) Transaction cost augmentation (^{1153})</td>
</tr>
<tr>
<td></td>
<td>(1b) rent-seeking costs (^{1151})</td>
<td>(3) Rent extraction (^{1154})</td>
</tr>
<tr>
<td><strong>Adaptive Efficiency</strong></td>
<td>Disadvantages: (^{1155})</td>
<td>Advantages: (^{1156})</td>
</tr>
<tr>
<td></td>
<td>(1) an increasing number of veto players block efficient transitions</td>
<td>(1) focal point provision</td>
</tr>
<tr>
<td></td>
<td>(2) an increasing number of rent-seeking groups promote inefficient transitions</td>
<td>(2) legislators’ capabilities of compensating losers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disadvantages: (^{1157})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) political short-sightedness</td>
</tr>
</tbody>
</table>

\(^{1142}\) See 7.1.1.1.1.  
\(^{1143}\) See 7.1.1.1.2.  
\(^{1144}\) See 7.1.1.1.3.  
\(^{1145}\) See 7.1.1.2.  
\(^{1146}\) See 7.1.2.  
\(^{1147}\) See 7.2.2.  
\(^{1148}\) See 7.2.5.2.  
\(^{1149}\) See 7.2.5.1.  
\(^{1150}\) See 7.3.1.2.4.  
\(^{1151}\) See 7.3.1.2.5.  
\(^{1152}\) See 7.3.2.1.  
\(^{1153}\) See 7.3.2.2.  
\(^{1154}\) See 7.3.2.3.  
\(^{1155}\) See 7.4.2.  
\(^{1156}\) See 7.4.1.  
\(^{1157}\) See 7.4.2.
outsiders are often unable to block inefficiency

Figure 40 Productive and Agency Efficiency Characteristics of Bureaucracy

<table>
<thead>
<tr>
<th>Bureaucracy</th>
<th></th>
</tr>
</thead>
</table>
| **Productive Efficiency** | Advantages:  
(1) Technical expertise and greater cognitive resources accentuate the informational advantages associated with the \textit{ex ante} centralized dimension of law-making. Technocratic centralization (unlike political centralization) is more than a cost-saving technique; it is an accuracy-enhancing mechanism.  
Insulation from the political process entails disadvantages:  
(1) Overregulation\textsuperscript{1159}  
(2) Cost ineffectiveness\textsuperscript{1160}  
(3) Output ineffectiveness\textsuperscript{1161}  
(4) Lower receptiveness to local preferences.\textsuperscript{1162} |
| **Agency Efficiency** | Two sources of increased agency costs associated with technocratic law-making:  
(1) Bureaucratic drift, plus associated costs of \textit{ex ante} and \textit{ex post} control\textsuperscript{1163}  
(2) Strategic use of delegation by politicians (except in the case of credibility-enhancing delegation)\textsuperscript{1164} |

\textsuperscript{1158} As I anticipated in the Introduction to the thesis, I have not discussed the social choice efficiency or adaptive efficiency of bureaucratic processes.  
\textsuperscript{1159} See 8.1.1.1.  
\textsuperscript{1160} See 8.1.1.2. and 8.1.1.3.  
\textsuperscript{1161} See 8.2.  
\textsuperscript{1162} Ibid.  
\textsuperscript{1163} See 8.2.1.  
\textsuperscript{1164} See 8.2.2.
**Figure 41 Process Efficiency Characteristics of Adjudication**

<table>
<thead>
<tr>
<th>Adjudication</th>
<th>Demand Side</th>
<th>Supply Side</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Productive Efficiency</strong></td>
<td>Demand-side explanations of adjudication identify structural variables that are most likely to influence the efficiency of law-making through adjudication (see <em>infra</em> table in Figure 45).(^{1165}) However, these do not support the “efficiency of common law” hypothesis. There is no unambiguous evolutionary tendency toward the production of efficient legal rules that can be inferred from an <em>a priori</em> analysis of the demand side of the adjudication process.</td>
<td>Obstacles to efficiency: (1) Due to the economic calculation problem, it is impossible for courts to achieve efficiency (2) A “selection” bias in the flow of cases brought before judges prevents courts from being fully informed(^{1166}) (3) The balancing assessment of reasons is more costly than the subsumption in statutes; this implies less replicability of judicial precedents compared to statutory rules(^{1167}) (4) The decisional perspective of judges is that of the efficiency of the “average” case(^{1168}) Advantages of the doctrine of precedents: (1) Decision-making costs saving effect (2) Judicial specialization (3) Outcome correctness (jury theorem effect). The degree of polycentrism of the adjudicatory system accentuates the informational advantages of adjudication.</td>
</tr>
<tr>
<td><strong>Social-Choice Efficiency</strong></td>
<td>Adjudication enjoys comparative advantages over politics due to four stabilizing features of the judicial process.(^{1169}) (1) Unlike legislatures, courts do not set their own agendas. (2) A judge is never confronted with the option of reconsidering a rule that has been excluded by a prior decision; therefore, there is no cycle because no rejected alternatives are reconsidered. (3) Although adjudication is not completely immune from the problem of cycling, the preference-based rationality that characterizes adjudication is a source of comparative advantage over politics. (4) Parties can negotiate around common law rules (which is not the case for statutory rules).</td>
<td></td>
</tr>
<tr>
<td><strong>Agency</strong></td>
<td>(1) Adjudication entails a lower access</td>
<td>(1) More limited rent-offering power</td>
</tr>
</tbody>
</table>

\(^{1165}\) See 6.1.1.  
\(^{1166}\) See 6.1.2.1.2.  
\(^{1167}\) *Ibid.*.  
\(^{1168}\) *Ibid.*.  
\(^{1169}\) See 6.3.2.
<table>
<thead>
<tr>
<th>Efficiency</th>
<th>Cost threshold than politics, which, in some cases, enables access to the law-making process of interests that would be excluded by the political arena.\textsuperscript{1170} for judges than politicians\textsuperscript{1171} (2) Courts become insensitive to external influence at lower levels of influence expenditures. \textsuperscript{1172} This entails reduced rent-seeking costs and outcome-agency costs. (3) Despite the fact-specific nature of the adjudication process, the reason-based decision-making used by courts enables a higher degree of generality in the production of law than political law-making does.\textsuperscript{1173}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive Efficiency</td>
<td>Decentralized adjudicatory systems facilitate legal change for three reasons: (1) Change-seeking actors that are minorities in the political process might be advantaged in adjudication because of the levelling-of-the-playing-field effect\textsuperscript{1174} (2) The incremental nature of judicial legal change\textsuperscript{1175} (3) No lawmakers’ inertia\textsuperscript{1176} The following disadvantages must be considered: (4) The doctrine of precedent may generate a lock-in effect\textsuperscript{1177} (5) The dichotomous nature of the outcome might increase agency costs\textsuperscript{1178}</td>
</tr>
</tbody>
</table>

**Figure 42 Process Efficiency Characteristics of Spontaneous Law-making**

<table>
<thead>
<tr>
<th>Productive Efficiency</th>
<th>Spontaneous law-making is based upon direct participation of norm beneficiaries and target actors in the production process. This enables: (1) Better information about the preferences of the people subject to law,\textsuperscript{1179} (2) Better factual knowledge and greater technical expertise of the regulatory issue at stake \textsuperscript{1180} However, spontaneous law-making is subject to the following limitations:\textsuperscript{1181} (1) The limited information and bounded rationality of individuals participating in the norm formation process (2) Externality problems (3) Public good problems affecting the formative process of norms (4) Randomness and path dependency, which are exacerbated by informational and reputational cascades</th>
</tr>
</thead>
</table>

\textsuperscript{1170} See 6.2.1. \textsuperscript{1171} See 6.2.2.1. \textsuperscript{1172} Ibid. \textsuperscript{1173} See 6.2.2.2.1. \textsuperscript{1174} See 6.4.1.1. \textsuperscript{1175} See 6.4.1.2. \textsuperscript{1176} See 6.4.1.3. \textsuperscript{1177} See 6.4.2.1. \textsuperscript{1178} See 6.4.2.3. \textsuperscript{1179} See 6.5. \textsuperscript{1180} Ibid. \textsuperscript{1181} See 5.1.4.
Social Choice Efficiency

Spontaneous law-making is subject to the cycling problem in the case of the empty bargaining core. However, politics and other centralized law-making processes do not enjoy any comparative advantages in devising cycle-breaking rules.\textsuperscript{1182}

Agency Efficiency

The absence of delegation mitigates the agency problem by reducing rent-seeking pressure. However, spontaneous law-making is subject to the risk of manipulation by interest groups.\textsuperscript{1183}

Adaptive Efficiency

Greater incentives for norm beneficiaries (who participate directly in the law-making process) to pursue legal innovation and to experiment with new regulatory solutions. However, two problems undermine the adaptive efficiency of spontaneous law-making:
(1) High enforcement costs and related collective action problems\textsuperscript{1185}
(2) Evolutionary traps—in particular, the difficulties of moving toward a global optimum in the case of the non-convexity of individual preferences\textsuperscript{1186}

9.3.3. Parts II and III: The Structural Variables

This section emphasizes the contributions provided by this study to a better understanding of the structural variables affecting the production of law.

Many contributions dispersed across various areas of the literature have recognized the importance of structural variables in determining the efficiency of law. However, this recognition remains limited to the discussion of specific legal problems. The mechanisms linking institutional features and environmental conditions remain significantly under-explored in the literature. In this respect, the comparative institutional analysis conducted in Parts II and III has made a first step in the direction of a systematic investigation of the environmental conditions under which each source of law enjoys relative advantages in terms of process efficiency.

Part II inquired into how legal homogeneity and frequency affect the interaction between the demand and supply sides of the law-making process, thereby determining the comparative advantages of various combinations of ex ante / ex post and centralized / decentralized law-making.

\textsuperscript{1182} See 5.2.
\textsuperscript{1183} See 5.3.
\textsuperscript{1184} See 5.5.
\textsuperscript{1185} See 5.4.1.
\textsuperscript{1186} See 5.4.2.
The discussion in Part III adopted a lower level of abstraction, thereby incorporating additional supply side-features into the model. The same approach was adopted with respect to the structural variables. The discussion in Part III identified numerous characteristics of the regulated environments that interact differentially with these additional institutional features. The purpose of the analysis was to identify the environmental conditions under which alternative sources of law enjoy relative advantages in terms of process efficiency. The tables in Figures 43-46 summarize the main results of the discussion.

**Figure 43 The Structural Variables Affecting Politics**

<table>
<thead>
<tr>
<th>Politics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Productive Efficiency</strong></td>
<td>Politics enjoy advantages in cases of high frequency and homogeneity of legal demand. However, in these cases, politics suffer from informational disadvantages when the regulatory environment is characterized by: (1) highly technical or (2) highly dispersed information. In case (1), technocratic law-making has comparative advantages. In case (2), decentralized processes have advantages.</td>
</tr>
<tr>
<td><strong>Social-Choice Efficiency</strong></td>
<td>Majoritarian decision-making generates a stable equilibrium outcome in the two following scenarios: (1) single-dimensional issue space + single-peak preferences (2) multidimensional issue space + Plott-condition (pairwise symmetry of preferences)</td>
</tr>
<tr>
<td><strong>Agency Efficiency</strong></td>
<td>The magnitude of rent-seeking costs is affected by: (1) the distribution of per capita stakes (2) the structure of the political conflict (3) the characteristics of the market for regulation (4) issue visibility The magnitude of agency slack is affected by: (5) incumbent self-publicity (6) political competition (7) public-policy intelligentsia (8) news media Agency slack is reduced by the following institutional features: (1) the division of power</td>
</tr>
</tbody>
</table>

1187 See supra Figure 38. 
1188 See 7.1.2. 
1189 See 7.2.3.1.1. 
1190 See 7.2.4.2. 
1191 See 7.3.1.2.6. 
1192 See 7.3.2.1.1.
| (2) a bicameral legislature  
| (3) the fragmentation of political parties  
| (4) rules governing legislative proceedings, which impose constraints on legislators’ strategic behaviours  
| (5) the public financing of campaigns  

Augmentation costs are exacerbated by:  
(8) the complexity of the issue  
(9) the enhancement of political job security  
(10) pay-offs to lawmakers, which are financed by third parties receiving benefits from transaction cost augmentation  
(11) the existence of a justification for the measure appealing for public opinion  
(12) the perceived importance to constituents of the regulatory issue  
(13) the publicity of the issue  

The augmentation costs tend to be reduced by:  
(14) media attention  
(15) institutional mechanisms aimed at reducing the degree of autonomy of politics (e.g., separation of powers, judicial review, bicameralism and so on)  

Rent creation and rent extraction depend upon:  
(16) the elasticity of industry supply with respect to changes in production costs  
(17) the elasticity of industry demand with respect to changes in price of goods  

### Adaptive Efficiency

Politics enjoy comparative advantages in promoting legal change when:  
(1) the balance of power among affected interest groups is in favour of socially efficient legal change  
(2) the visibility of the issue acts as a catalyst for forces pushing for the legal change.  

---

**Figure 44 The Structural Variables Affecting Productive and Agency Efficiency of Bureaucratic Processes**

**Bureaucracy**

| Productive Efficiency | The expected benefits from technocratic law-making both increase and tend to outweigh the tendency toward productive inefficiency under the following environmental conditions:  
| (1) outcome uncertainty  
| (2) technical complexity of the regulatory issue at stake  

**Agency Efficiency**

| The following variables affect the structure of the agency costs entailed by technocratic law-making:  

---

1193 See 7.2.3.1.2.  
1194 See 7.3.2.2.2.  
1195 Ibid.  
1196 See 7.3.2.3.2.  
1197 7.4.2.2.  
1198 As I anticipated in the Introduction to the thesis, I have not discussed the social choice efficiency or adaptive efficiency of the bureaucratic processes.  
1199 See 8.2.1.3.1.
(1) **outcome uncertainty**
(2) politicians’ **risk aversion**
(3) **Conflict of preferences** between politicians and bureaucrats
(4) **Information asymmetries** between politicians and bureaucrats

Blame-shifting delegation is likely to occur in cases of:
(1) Concentrated costs and widespread benefits for the people subject to law (if the political benefits outweigh the costs of delegation)
(2) Highly conflictual, purely distributive issues

---

**Figure 45 The Structural Variables affecting Adjudication**

<table>
<thead>
<tr>
<th>Adjudication</th>
<th>The direction of the evolution of judge-made law is determined by:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Productive Efficiency</strong></td>
<td>(1) the degree of substantive inefficiency in existing law</td>
</tr>
<tr>
<td></td>
<td>(2) parties’ incentives to establish favourable legal precedents</td>
</tr>
<tr>
<td></td>
<td>(3) parties’ willingness to invest in litigation</td>
</tr>
<tr>
<td></td>
<td>(4) parties’ available information</td>
</tr>
<tr>
<td></td>
<td>(5) the relative costs of settlements and litigation</td>
</tr>
<tr>
<td></td>
<td>(6) the relative costs of collective action for different interest groups</td>
</tr>
</tbody>
</table>

The relative degrees of replicability for statutory and jurisprudential rules depend upon:
(7) the relative clarity of statutes versus precedents
(8) the relative difficulties of statutes and precedents in solving conflicts with, respectively, previously enacted rules and previous legal decisions
(9) the relative degrees of maturity reached through the application over time of enacted law and case law
(10) the degree of heterogeneity of the cases to be regulated

Finally:
(11) the degree of polycentrism of the source of law system accentuates the relative advantages of adjudication

<table>
<thead>
<tr>
<th>Social Choice Efficiency</th>
<th>The occurrence of intransitive cycling depends upon the structure of preferences of the individual members of the adjudication body.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agency Efficiency</td>
<td>The intensity of the agency problem in adjudication depends upon:</td>
</tr>
</tbody>
</table>
<pre><code>                                                             | (1) the characteristics of the constitutional framework |
</code></pre>

---

1200 Ibid.
1201 Ibid.
1202 See 8.2.1.3.2.
1203 See 8.2.1.3.3.
1204 See 8.2.2.1.
1205 Ibid.
1206 See 6.1.1.
1207 Grant, “Precedents”, supra note 890 at 7.
1208 6.1.2.3.4.
1209 6.2.2.3. To be clear, the constitutional framework within which courts operate is a structural variable (not a choosing variable), to the extent that it is taken as a given. Under this assumption, adjudication has greater relative advantages over politics in polycentric than in centralized constitutional systems. By contrast, if we regard the constitutional framework as an institutional variable, than we should conclude
(2) the differential abilities of interest groups to engage in effective litigation

Adaptive Efficiency

(1) The degree of strictness of the doctrine of precedent should be calibrated to the level of adaption information and transaction costs in the regulatory environment

(2) The rate of change of the regulatory environments

<table>
<thead>
<tr>
<th>Figure 46 The Structural Variables Affecting Spontaneous Law-Making</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spontaneous Law-making</strong></td>
</tr>
<tr>
<td><strong>Productive Efficiency</strong></td>
</tr>
<tr>
<td>The environmental conditions under which spontaneous law-making constitutes an efficient source of law are summarized in the notion of close-knittedness, which includes (a) power reciprocity among community members, (b) continuing relationships and sufficiently high discount factors and (c) adequate knowledge and information about the history of potential trading partners.</td>
</tr>
<tr>
<td><strong>Social Choice Efficiency</strong></td>
</tr>
<tr>
<td>Bargaining among private actors is susceptible to the empty core problem in purely redistributive contexts (i.e., zero-sum games).</td>
</tr>
<tr>
<td><strong>Agency Efficiency</strong></td>
</tr>
<tr>
<td>The agency problem is likely to be accentuated in economic sectors in which there is sharp polarization between consumers and producers.</td>
</tr>
<tr>
<td><strong>Adaptive Efficiency</strong></td>
</tr>
<tr>
<td>The attainment of the number of enforcers critical to support self-reinforcing mechanisms depends upon: (1) the receptiveness of general audience members to the technical information required to appreciate the efficiency of the norm (2) the costs of communicating among community members information regarding the availability of a new norm.</td>
</tr>
</tbody>
</table>

9.4. Normative Contribution

The main purpose of this study has been to identify and explain the core tendencies of law-making processes predicted to be important in determining the efficiency of the law. From this perspective, the analysis in this study has been predominantly positive in nature. However, since efficiency is a concept with unavoidable normative overtones, a positive efficiency analysis can be neither immune to underlying normative assumptions nor insensitive to the normative implications of its conclusions.

that the degree of polycentrism of the organization of the sources of law should be increased in order to improve the agency efficiency of adjudication.

Ibid.

See 6.4.2.1.

See 6.4.2.2.

See Chapter 1.
First, my proposed definition of process efficiency rests on a number of well-specified normative assumptions. For example, as emphasized in Chapter 2, I build on the normative assumption that legal rules should seek to maximize social welfare.\textsuperscript{1214} I also assume that cost-minimization is a normatively desirable goal of the institutional designs of law-making. Further, I assume that law-making should be responsive to the preferences and interests of the people subject to law and that law-making outcomes should be rational and stable. I also assume legal predictability and certainty to be normative goals.

These normative assumptions constitute the logical basis for conducting a positive analysis of law-making. That is, without these normative criteria, it would be impossible to develop a taxonomy of the law-making costs that constitute the logical basis of a comparative institutional analysis of alternative law-making processes. For example, in order to qualify a given expenditure as an agency cost, an existing definition of a desirable agency relationship is needed. Thus, to define interest group pressure as “disproportionate” (and, as such, a source of agency “costs”), an underlying normative benchmark against which one can assess the “proportionality” of interest group pressure is needed. To conclude this point, the positive efficiency analysis of the law-making process is made possible by the prior specification of underlying normative assumptions. As Myrdan famously observed: “There can be no view except from a viewpoint. In the questions raised and viewpoint chosen valuations are implied”.\textsuperscript{1215}

\subsection*{9.4.1. Decentralizing and Depoliticizing the Production of Law}

Although the discussion has not focused on providing detailed policy guidelines, its analysis of the strength and limits of alternative sources of law has suggested some general prescriptive principles for the institutional design of law-making.

\textsuperscript{1214} For a critical assessment of this normative assumption, let me refer to Bertolini, “The Long Quest”, \textit{supra} note 1100, and the bibliography cited there.

First, *ex ante centralized law-making should be viewed as the solution of last resort in the organization of the sources of law.* This study challenges the widely held assumption that the principal legal devices by which societies can resolve pervasive social problems are either legislation or regulation. The environmental conditions under which these two processes are likely to enjoy relative advantages over alternative sources of law are very restrictive (see Figures 43 and 44) Moreover, even in such cases, *ex ante* centralized law-making, along with the additional institutional features of politics and bureaucracy (respectively, representative mechanisms and bureaucratic insulation), entails a dramatic increase in both production information and agency costs, which should be carefully weighed against its informational advantages.

Furthermore, politics and bureaucracy are relatively inefficient and ineffective sources of law. On one hand, the politicization of the law-making process outside of the restrictive circumstances identified in this study significantly undermines the substantive efficiency of the law. On the other hand, the delegation of law-making power to agencies constitutes a rational response to the informational inefficiencies of political processes in regulatory environments, which are characterized by outcome uncertainty and technical complexity of regulatory issues. However, technocratic law-making is ill-suited to remedying the misalignment of incentives between the *ex ante* centralized lawmaker and the people subject to law. As repeatedly emphasized, more bureaucracy entails more expertise, but not necessarily less “*ex ante* centralization”. Technocratic law-making does not allow for greater responsiveness to the preferences and interests of the people; rather, it simply shifts the agency problem from legislative to bureaucratic drift.

Second, *the efficiency of the law is advanced by the degree of polycentrism (“competition”) among the sources of law.* All sources of law are pervasively plagued by information and public choice problems, which are predominantly associated with the *delegation* of the law-making authority by the people subject to law. One of the most effective ways of reducing the inefficiency of the production of law is increasing the degree of *elasticity* of the demand for norms with respect to the adequacy of the law-making outcome. The most effective way to increase demand’s responsiveness to quality
outcomes is to ensure the presence of close substitutes for the existing lawmaker. Where close substitutes are not available, internal and external pressures on the law-making process will increase dramatically.\textsuperscript{1216} By contrast, where substitutes are close, if the outcome falls significantly short of the substantive efficiency parameter, the expected pay-off to the people subject to law decreases, which, in turn, generates incentives to opt out of the inefficient law.\textsuperscript{1217}

\textbf{9.5. Directions For Future Research}

As the Introduction to this study anticipated, the proposed analysis has been preliminary in nature. A full economic understanding of the production of law should undertake three further steps, which constitute guidelines for future research: (i) an inquiry into the institutional determinants of people’ preferences, (ii) an analysis of the dynamic interaction among sources of law, and (iii) an examination of the biases and heuristics affecting the cognitive processes that underlie the production of legal rules.

\textbf{9.5.1. The Historical Determinants of the Law-making Process}

First, the theory developed here treats people’s preferences (on both the demand and the supply sides) as strictly exogenous. Nothing in the theory explains why people hold the preferences they do. This is a major limitation of the theory. Rational actor theories cannot explain why norms are accepted or internalized. A full range of issues concerning the production of legal rules is rooted in historical processes and cannot be explained through rational actor models. Since the production of law is a fundamental constituent of the culture of a society, a full understanding of the production of law requires the study of the history of the production of law and the analysis of the evolution of alternative law-making processes from a comparative, legal, anthropologic perspective.

\textbf{9.5.2. The Institutional Determinants of People’s Preferences}

The comparative institutional analysis has implicitly assumed that people do not change their expectations or preferences during the decision-making process. However, real-

\footnotesize{\textsuperscript{1216} See, in particular, 4.3 and 6.1.2.3. \textsuperscript{1217} Ibid.}
world law-making processes exert significant effects on people’s preferences: People often change their preferences and expectations about law-making outcomes both during and because of their participation in the process. Hence, one promising line of inquiry is an investigation into whether and how law-making processes act as sources of individuals’ preferences and expectations by filtering and shaping them.

9.5.3. The Dynamic Interaction among Sources of Law.

The comparative institutional analysis developed in this study rests on the unrealistic assumption that each law-making institution (i.e., politics, adjudication, bureaucracy, and spontaneous law-making) acts as a “monopolist” with respect to the production of law. Although this analysis is helpful in capturing the comparative advantages and disadvantages of alternative law-making processes, it may lead to results that do not effectively model real-world economic phenomena. The impact exerted by the dynamic interactions between sources of law on law-making outcomes has not been investigated. In this respect, it is important to emphasize the importance of further inquiry into how the interactions among law-making institutions affect the efficiency properties of the law.

Alternative law-making design strategies may prove efficient, depending on the type of interaction. The interaction between two sources $x$ and $y$ in a given regulatory setting can be characterized in three different ways, depending on whether these sources: (i) are complementary sources, (ii) substitute for each other, or (iii) crowd out each other. The complementarity between sources of law calls for an institutional design that facilitates coordination among them. Substitute sources will tend to “compete” with each other for the allocation of law-making power, which might either increase or undermine the efficiency of the production of law, depending on a set of varying circumstances. Finally, the presence of a crowding-out effect generates a tendency toward monopoly in the production of law, which suggests the adoption of institutional arrangements aimed at reinforcing the plurality of the sources of law.

The analysis provided in this study rests largely on the assumptions of the rational choice theory of action.\textsuperscript{1218} This methodology allows for a nuanced understanding of the trade-offs associated with a choice between alternative law-making processes. However, it is not, in itself, sufficient to capture the complexity of institutional choices.\textsuperscript{1219} One of its major shortcomings is the failure to account for the relationship between the institutional design of law-making and the imperfections of the cognitive processes underlying the production of law. Future research should incorporate into the general theory of law-making an analysis of how the biases and heuristics characterizing law-making actors’ decision-making processes undermine process efficiency (or, alternatively, how they can contribute to the efficient production of law).\textsuperscript{1220}

I suggest two lines of inquiry. The first should investigate the ways in which the fundamental institutional dimensions of law-making affect the cognitive processes of the lawmaker. For example, the \textit{ex post} perspective entails a fact-specific approach, which can be distortive with regard to the way in which legal issues are presented. In particular, the \textit{after-the-fact} standpoint entails a higher probability of the “vividness bias”, or “hindsight bias”, which occurs when vivid information affects the importance attributed by the decision-maker to factual elements, independent of their legal relevance. In addition, \textit{ex post} law-making might fail to consider broader policy issues, focusing instead on factors that should not be considered from a broader law-making perspective. It is also true that the \textit{ex post} perspective involves significant advantages. The case-by-case method necessitated by the \textit{ex post} perspective allows one to address the same legal issue \textit{repeatedly} with different facts. This allows one the chance to review a whole range of previous decisions made in actual cases before a final pronouncement. Further, the possibility of dealing repeatedly with the same legal issue under a variety of factual

\textsuperscript{1218} In Chapter 1, I emphasized the limitations of the neoclassical theory of production for the purposes of a comparative analysis of law-making institutions.


\textsuperscript{1220} On biases and heuristics in lawmaking see, generally, Gerd Gigerenzer & C Engel, eds, \textit{Heuristics and the Law} (Berlin: Massachussets Institute of Technology and Freie Universitat, 2006)
circumstances allows the decision-maker to adopt a perspective that avoids logical pitfalls.

The *ex ante* perspective allows a broader view in many cases, which results in a comparative advantage when the issues that underlie the demand for regulation are “polycentric problems” (i.e., issues that cannot be resolved independently and sequentially). However, the risk of the *ex ante* perspective is “overconfidence” in judgment by the lawmaker. It is easier for the *ex ante* lawmaker to fall into the delusion that he or she has all the information needed to make good decisions. Such a propensity can produce judgment overconfidence, which is exacerbated by a lower propensity to revisit past experiences. Many other biases and heuristics might be discussed in relation to *ex ante* vs. *ex post* law-making. 1221

A second suggestion for future research concerns the behavioural shortcomings of the alternative law-making institutions. This is an important point for understanding real-world phenomena, since some biases are causally related to the characteristics of the “institutional arena” within which individuals interact. For example, legislatures tend to be vulnerable to *availability cascades* and, therefore, in some circumstances, may be poorly suited sources of regulation. 1222 The judicial process is affected by the *hindsight bias*, which, in many cases, undermines the rationality of the legal assessment (e.g., the assessment of the causal link between negligent conduct and the magnitude of the damage imposed on a victim is often influenced by a lack of technical expertise within a jury, which reinforces a strong hindsight bias). 1223 Both legislatures and courts tend to be

---


affected by *context-dependent* decisions, which may vary depending upon the set of options available.\textsuperscript{1224} The preferences of private parties with regard to a bargaining process are generally influenced by the *endowment effect*.\textsuperscript{1225} As a result, completely alienable entitlements might not be traded, even in presence of low transaction costs.\textsuperscript{1226} Such “stickiness” of legal entitlements may present an obstacle to enhancing an efficient form of private ordering. These examples suggest that the behavioral approach provides useful insights for a more accurate understanding and comparative assessment of alternative law-making institutions.

