The Form of the Firm: A Critical and Normative Theory of the Corporation

by

Abraham Andrew Singer

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Political Science
University of Toronto

© Copyright by Abraham Andrew Singer 2016
The Form of the Firm: A Critical and Normative Account of the Corporation

Abraham Andrew Singer
Doctor of Philosophy
Department of Political Science
University of Toronto
2016

Abstract

This dissertation explores the corporation from the perspective of normative political theory. As I argue, the corporation is not easily located within the theoretical matrix of liberal democracy. Because of this, a political theory of the corporation is required in order to evaluate whether our corporate institutions are consonant with the liberal democratic values we affirm. The first half of the dissertation explores the Chicago School’s political theory of the corporation. In this view corporations are taken to be the result of market-like individual voluntary initiatives, and therefore should be left alone for the sake of efficiency and liberty. The Chicago School argues for this, however, at the expense of coherence and sociological rigor. The better way to approach the corporation, I argue, is by seeing it as enabling what I call “norm-governed production.” On my account, firms overcome market inefficiencies by drawing on social and cooperative scripts that allow for cooperative activity that the market under-produces or does not produce at all.

This analysis opens up an avenue for prescription and critique that does not deny the economic aspects of the corporation; if we admit that the efficiencies produced by the corporation entail normative scripts, and not merely voluntary contracts, then we can ask whether or not such scripts are compatible with our notions of fairness and justice. We can also ask whether changing scripts in the direction of fair or just relationships can be done without sacrificing too much in the way of corporate efficiency. In the second half of my dissertation I leverage this account of corporate efficiency into an alternative political theory of the corporation. I argue that our legal and political institutions ought to be restricted so as to protect corporations’ vulnerable parties and to encourage
ownership patterns that empower those who make the most personal investments. I further argue that, from the perspective of business ethics, the corporation ought to be reconceived as an agent in maximizing efficiency while also being an agent for mitigating social injustice.
Acknowledgments

The underlying, animating idea in this dissertation is that activity done cooperatively, and not merely in concert, changes the nature of that activity in morally significant ways. My own personal moral belief—which I neither defend nor draw attention toward—is that such acts of cooperation ought to be strived for, and that a life properly lived is a life dedicated to this type of activity. I like to think that in completing this dissertation I have vindicated both these ideas. The completion of this dissertation was only possible because of the immense help and support I received from others; this cooperative context made what would otherwise be onerous and atomizing, ultimately, a life-affirming task and achievement. Completing this thesis, and doing so with my sense of self more or less intact, is therefore owed to the many people who helped me in various fashions along the way.

To begin with, I must thank my committee. My co-supervisors, Joseph Heath and Joseph Carens—my “not-so-average Joes”—remain models of scholarship to which I will always aspire. Joseph Heath first got me interested in the topic and has continued to help me see the importance and complexity in many of these ideas. The breadth of his knowledge, which spans many disciplines, has kept me on my toes and opened me to ideas I might not have encountered otherwise. Joseph Carens’s uncanny ability to get right to the heart of any argument or idea made this dissertation more lean and coherent than I could have mustered on my own. More incredible is his manner of doing this, which is at once both blunt and caring; for me this is the archetype for how political theory ought to be done. Perhaps more importantly, both Joes gave incredible support personally and professionally, offering me advice and help that extended well beyond their function as supervisors of my dissertation. Peggy Kohn was a huge part of this project as the other member of my committee. She remains one of the sharpest and most intelligent people I have ever met; her insights and criticism made this thesis much stronger than it otherwise would have been, and her approval and encouragement always left me elated. A special thank you to my examiners, Ed Andrew and Jacob Levy, who stepped into this role under heavy time constraints and whose comments will be hugely important going forward. Together my committee was responsive, responsible, and attentive.

I am so grateful for my colleagues in the University of Toronto Department of Political Science. Having such people as colleagues and friends meant that my work was always constructively challenged, that my horizons were always broadened, and that I was ultimately affirmed and encouraged to continue. There are too many people to list here but a few people
require special thanks. I thank Luke Melchiorre for being a patient roommate and a constant source of laughter and insight. I thank members of my cohort, Steve Trochimchuck, Hamish Van der Ven, Matt Lesch, Marion Laurence, Heather Miller, Jessica Soedrigo, and Andrew McDougall for helping me through the rigmarole of a graduate program by providing a nurturing and supportive community. Thank you to Noaman Ali and Igor Shokheidbrod for being models of how to mix intellectual endeavor and political intention without watering down either.

I was so privileged to be able to be a part of the political theory community at the University of Toronto, which includes some of the smartest people in the business, all of whom engage each other in a spirit of camaraderie and ecumenism. They also, thankfully, know how to cut loose and dance. Thank you to Mauricio Suchowlanski, Stephanie Murphy, Emma Planinc, Simon Lambek, Ella Street, Jaby Matthews, and Taylor Putnam for last minute printing help. Although she was not a member of my committee I learned a great deal from Simone Chambers through coursework and through conversation, and I remain deeply influenced by her thinking and her approach to writing. Through my time at U of T I became more interested in the history of political thought than I ever thought I would be. This is in part due to Clifford Orwin and Ryan Balot opening their classroom and reading group doors to a curious Rawlsian philistine like myself; I thank them for being so welcoming and encouraging. It also has much to do with the influence of my dear friends, Lindsay Mahon and Lincoln Rathnam; in admitting this I concede to losing an argument we began years ago. For this and for so much more I thank them both. And I would be remiss if I did not acknowledge the significant contribution of my colleague, neighbor, and friend Kiran Banerjee whose patience, generosity, willingness to read drafts, and ample supply of midnight whiskies helped bring this dissertation to completion.

I was helped by many people outside the University of Toronto. Thank you to my UIC folks who started me out on this path and continue to push and enlighten me: Isaac Balbus, Stephen Engleman, Zach Gebhardt, Vanessa Guridy, Heidi Lawson, Nawojka Lesinski, Norma Moruzzi, Jason Stodolka, and Sultan Tepe. I wrote chapter 2 of this dissertation while a visiting scholar at the Quinlan School of Business at Loyola University Chicago. I thank Alexei Marcoux for making that possible. I wrote drafts of Chapter 4 and Chapter 7 while a visiting student at the Normative Orders cluster at Goethe University in Frankfurt. I thank Rainer Forst for that opportunity, and the incredible group of people there who created such a great atmosphere and who gave significant feedback on drafts: Daniel Callies, Beth Kahn (on both chapters!), Lisa Herzog, Esther Neuhann, Nate Adams, Julian Culp, Merten Reglitz, Max Hugendubel, Laura
Roth, Johannes Schulz, and Eszter Kollar. Also for beer, soccer, and making me feel at home in Germany: Tijana Milunovic, Alexandra Shabaneecki, Dorte Huhnert, and, I guess, Matze. Closer to home, Chris MacDonald gave me feedback on early drafts at the Canadian Business Ethics Research Network PhD Workshops. He is a great source of positivity and encouragement, and a unique example of a successful academic who is also competent at making things happen. For that, for his professional advice, and his friendship, I thank him dearly.

I should note that my dissertation might have been completed earlier had it not been for the strike that I took part in with CUPE 3902. That remains one of the most amazing and most difficult experiences of my life. Though the time and stress that came with having a leadership position may have disrupted the process of finishing my thesis, it was also a reminder of what collective action involves and why we engage in it, and therefore reinvigorated my interest in my dissertation. For that reminder, and the inspiration that I got from everyone’s selflessness, I thank every single person who participated in that strike in all their different ways. I thank fellow members of the strike committee and those who contributed with centralized logistics—Abe Nasirzadeh, Craig Smith, Victor Lorentz, Golbarg Rekabtalaei, Banafshe Beizaei, Kiran Banerjee, Alex Ivovich, Lama Mourad, Lincoln Rathnam, and Jeremy Hurdis, among many others—for keeping an eye on the big picture through trying times, for all the legwork and preparation that went largely unseen and unrecognized, and for being friends throughout it all. I also thank folks who served as picket captains and worked tirelessly on the ground in organizing pickets and mobilizing our members. There are far too many to mention but in particular I want to thank the Delta Squad: Joe Curnow, Caitlin Henry, Anjali Halferty, Rebecca Bartel, Omar Sirri, Kevin Edmonds, Madeline Whetung, Ellie Adekur Carlson, and Rachael Kimmerling. Thank you for being so strong, so dedicated, so caring, and showing me how to fight the good fight while being sensitive to what others are going through.

I admit that I often felt caught between both of these groups of people—those involved in the more bureaucratic and logistical aspects of things and those trying to maintain democratic engagement and mobilization—and sometimes felt like I was part of neither. Obviously, tensions arose, tensions that were never resolved in our strike. But I am convinced that these were the right tensions to have; they reflect the very real commitment of everyone participating, as well as the extraordinary undertaking that is large-scale collective action. How best to deal with the tension between these two imperatives is a question that underlies my entire dissertation. For making me fully aware of what is at stake in this kind of question, I thank everyone who struck.
I was lucky enough to have friends and support from outside the academy who kept me from falling into irredeemable nerdiness throughout this process. Thank you for the people who were there when I first became passionate about scholarship, going all the way back to New College, particularly Bill Cameron, Patrick Crummey, Melissa Fourage. Thank you to the expanding circle of lovely folks that first met in orbit of the Lowther Bloc of CCRI: Liz McLaughlin, Toby Bowers, Kaitlyn Cinovskis, Lindsay Denise, Emma Lander, Suzanne DeWeese, Emily O’Brian, Emmet Ferguson, and Astrid Idlewild. Our time together in CCRI first inspired my interest in economic organizations and I am so proud of the family we’ve built and continue to build for ourselves. Thanks to Claire Lobenfeld and Kelli Korducki for their love, encouragement, and for being better writers than I will ever be; being a fellow traveler with you as we strike out into the professional world of letters has been a gift. Thank you to Florence Noel and Alex Kondo for being able to pick up a conversation from a year prior seamlessly and for giving me a sense of continuity and constancy. Thank you to Matt Bruno, Ben Rudow, and Lucas Frankel for, among so many other things, the annual pilgrimages to the Garden. Merci beaucoup: Diane Eros, for good spirits and good food, and Patricia Mangeol, for her keen intellect and different perspective. A special thank you to Andrew Budziak, for being a model of (ir)responsibility and never letting me become an academic hermit; to Boaz Beeri, for always making me feel like what I was doing was somehow noble and epic; and to Jon Medow, who first got me interested in political theory, and who I secretly am still always hoping to impress.

I am fortunate to have a beautiful family that has been so loving and supportive. Thank you to all the Keenans, Pressers, Frees/Friedlanders, and Singers who encouraged me along the way! A special thanks to Emily Hallock for being both a cousin, a comrade in theory, and a friend. Thanks to my grandmas—Nana Keenan, Grandma Terry, Grandma Jean, and Grandma Sid—for instilling in me curiosity, passion, and care, and being models of strong and compassionate women. Thank you to my brothers Sam and Patrick who are two of the funniest and smartest people I know. There are very few days that go by where I don’t find myself chuckling at some joke you once cracked. I love you both. And thank you to my parents—Mom, Dad, June, and Barry. I am blessed to have four loving parents who have given me so much in so many particular ways. The theory of this project was influenced by a piece of wisdom that my father has passed down, which he has dubbed Singer’s Law: “There’s always a schmuck.” The writing of this project was influenced by what my mother once told me was the first rule of writing, the adverse of what was written above the temple of Delphi: “Know thy audience.”
Thanks to my partner, Luisa Rollins, for her encouragement, care, and patience throughout this whole process. Despite my inability to make a mixtape for her in a timely manner, she has been my most thoughtful reader and critic, my traveling companion, my roommate, my co-chef, the person who keeps me in line by calling me on my nonsense, and the person who encourages and soothes me when things feel scary and hopeless. No habría podido hacerlo sin ti. Te quiero.

And, finally, some people’s influence on me has outlasted their time on this earth. I would like to dedicate this dissertation to a few of them in particular: to Stanley Free for teaching me how to appreciate that which is well-crafted and beautiful; to Frederick Singer for encouraging me to be inquisitive and curious; and to Cindy Bruno for exemplifying relentlessness in the face of adversity.
# Table of Contents

Acknowledgments .......................................................................................................................... iv

Table of Contents ........................................................................................................................... ix

List of Tables .................................................................................................................................. x

List of Figures ................................................................................................................................ xi

Introduction ..................................................................................................................................... 1

Chapter 1. Assumptions and Approaches ..................................................................................... 27

Chapter 2. The Productivity Component of Transaction Cost Economics ................................. 49

Chapter 3. The Political Theory of the Corporation of the Chicago School ................................. 90

Chapter 4. Norm-Governed Productivity .................................................................................... 146

Chapter 5. Relational Corporate Law ......................................................................................... 193

Chapter 6. The Architecture of Corporate Governance .............................................................. 227

Chapter 7. Business Ethics: Market Failures and Justice Failures .............................................. 261

Conclusion .................................................................................................................................. 300

Bibliography ............................................................................................................................... 315
List of Tables

Table 2.1 The Productivity Component of Smith, Mill, and Marx .............................63
Table 2.2 The Productivity Component of the Transaction Cost Approach in Comparison ......87
Table 3.1 Berle and Means’s Normative Component .................................................109
Table 3.2 Managerial and Chicago School Normative Components ..........................139
Table 3.3 Transaction Costs and Law and Economics Productivity Components ..........140
Table 4.1 The Productivity Components of the Chicago School and the Norm-Based Model .191
Table 7.1 The Market Failures Approach to Business Ethics .......................................280
Table 8.1 The Productivity Component of the Chicago School .................................302
Table 8.2 The Normative Component of the Chicago School .................................302
Table 8.3 The “Norm-Governed” Productivity Component .................................303
Table 8.4 My “Normative” Component .................................................................305
List of Figures

Figure 1 The Relationship between Shareholders, Directors, and Management ………………..11
Figure 7.1 The Place of Business Ethics within a Scheme of Social Justice …………………..283
Introduction

The corporation is undoubtedly one of the most important institutions in modern liberal capitalist societies, controlling a very large amount of wealth and, in turn, commanding a great degree of social and political influence. Yet for much of the 20th century, mainstream political theory and political philosophy has had little to say about it. I attempt to address this oversight. In this dissertation I engage with and criticize the theory of the corporation developed by economic theorists, in order to better understand how political theory can speak to, and learn from, questions that arise when considering the modern business corporation. By using methods of political theory to assess the economic and legal theories of the corporation, I attempt to offer answers to two important-yet-largely-unasked normative questions about this institution. From the perspective of liberal political theory, is the power within the modern business corporation legitimate? With qualification, I answer “no.” Could it be made to be legitimate? With qualification, I answer “yes.” The qualifications make up the bulk of this dissertation.

One must start with law and economics scholarship because that is where the bulk of reflection about the corporation has taken place. Whereas normative political philosophy since John Rawls has been concerned with macro-level questions about egalitarianism in society and the contours of its “basic structure,” and normative business ethics has focused on micro-level concerns of particular market actors’ duties, questions about the institution of the corporation and its organization have largely been dealt with in economic departments and law schools. As we will see there are exceptions to this (the main one being theories of workplace democracy, which offer important critical evaluations of the corporation’s hierarchical structure). Still, economic theorists, particularly in the Chicago tradition of law and economics, have developed theoretical accounts of the most important aspects of the modern corporation’s institutional structure: the nature of corporate hierarchy; the separation of ownership and control; the primacy of the
shareholder; managerial discretion; and government regulation. All of this has been generated from a relatively simple theory of the corporation as a nexus-of-contracts, a bundle of efficiency-improving and freely-chosen arrangements that individuals haven imposed upon themselves. The result has been a well-developed theoretical explanation of the corporation, the power of which comes from its parsimony and nuance.

My dissertation has two parts. The first attempts to articulate and assess this economic justification of the corporation in normative terms. That is, while economic theorists sometimes offer their theories of the corporation as mere explanations of why the corporation has emerged in its current form, their analyses also often offer a justification and defense of a particular vision of the corporation. From this perspective, the corporation as it currently exists is, on the whole, morally justifiable and therefore ought to be largely left alone by other social institutions, especially the state. Any political theory of the corporation, then, must start with these economic theories, and take seriously both their explanatory and normative claims. Chapters 2 and 3 of my dissertation attempt to do precisely this. By placing the development of the Chicago School’s theory of the corporation in the context of economic intellectual history, we can see the power, strength, and novelty of its contribution, and better understand the nature of its claims.

In the second part of my dissertation I offer a critique of this economic theory of the corporation and a limited counter-proposal. I concede the primary significance of efficiency to an understanding of what the corporation is and ought to be, and I contend that critical accounts of the corporation must take these economic concerns more seriously than they do. However, I also contend that these dominant economic accounts of the corporation make a crucial error; they overlook how the corporation’s contribution to overall economic efficiency depends upon norms and social relationships within the corporation. Taken simply as a positive explanation of the
corporation, Chicago’s contractual account is wanting. As a result, its normative force is also weakened. In chapter 4 I argue that force, power, and compulsion play an inevitable role within the corporation because of the way norms and social pressure within the corporation work. For that reason, contrary to the contractualist theory of the corporation, individuals’ preferences cannot be assumed to be exogenous to the corporation. Because people’s preferences and behavior will be altered by the corporate institution in which they participate, we cannot assume that such an institution can be justified as the outcome of free consent by all those involved in the corporation. Therefore, corporate power requires a normative assessment in terms of the values and relationships corporate organizations promote.

I therefore contend that the corporation must be assessed not only in terms of its contribution to overall efficiency but also in terms of the kind of relational values it generates. Any particular corporation must be efficient enough to warrant its existence in the marketplace, but ought also to avoid cultivating relationships that undermine the values of liberal democracy. As a result I offer three different normative proposals for the contemporary corporation. I offer an account of corporate law that recommends treating the corporation as an entity distinct both from the individuals who contract with it and the government that recognizes it (chapter 5); I offer an account of corporate governance that shows why a range of background legal, regulatory, and financial institutions are needed to render corporate relationships more legitimate (chapter 6); and I offer an account of business ethics, which highlights managerial duties with respect to both market efficiency and liberal-democratic legitimacy (chapter 7). Overall my goal is to show that the moral legitimacy of the corporation as an institution in a society committed to liberal democratic values depends upon a range of regulatory, legal, and ethical background
conditions that can constrain the organization and activity of corporations; only some of these are satisfied in contemporary liberal societies.

Although I attempt to offer my own outline of what a political theory of the corporation looks like, this project might best be understood primarily as an internal critique of the Chicago School and its theory of the corporation. This theory holds such sway and influence in intellectual, legal, and political discourses, that such a critique is not only warranted but necessary if we wish to achieve a more just regime of corporate organization. Because of this, I take the Chicago School’s account seriously and try to show where it is most convincing and illuminating, before showing why I think it ultimately fails on its own grounds. My own proposals—though perhaps more limited than some other radical interventions in the study of political economy and the workplace—do call for radical departures from the status quo, but do so while taking the dominant economic discourse seriously.

Throughout all of this, one can detect a subplot involving the scholar Ronald Coase. As we shall see, Coase is both a foundational figure in the development of the transaction cost account of the firm and the Chicago School of economics; the former coming from his 1937 article “The Nature of the Firm” and the latter coming from his 1960 article “The Problem of Social Cost.” Part of my contention about the Chicago School is that they combine these two different insights in ways that privilege the latter. In chapters 2, 3, and 4 I try to disentangle the various contributions of Coase in order to show that the Coasian insights with regards to the firm are distinct from—and actually at odds with—his contributions to the economic theory associated with the Chicago School. This rescuing of Coase runs parallel to my engagement and criticism of the Chicago School: what I take to be most laudable about the Chicago School approach is generally what is associated with the Coasian theory of transaction costs; what I take
to be most problematic about the Chicago School is associated with their misinterpretation and/or misuse of his theory. As a result, in chapter 4, my critique of the Chicago School also includes a novel reading of Coase’s work, highlighting his distinctive views on institutions and moral psychology.

Before getting into the meat of my analysis, I must bend to scholarly convention and make explicit the intellectual context of this study and its methodology. In chapter 1 I offer an account of what a political theory of the corporation would be and the key questions such a theory should answer. I therefore delay a discussion of my methodology until then. For the remainder of this introduction I first offer a definition of the corporation. I then situate this study in its scholarly context. I should note that much of the literature to which this study relates gets discussed in detail in the chapters to come. I therefore leave much out of my literature review: classical and Marxist political economy (discussed in chapter 2), 20th century managerial and economic theory (chapter 3), and theories of workplace democracy (chapters 4 and 5). In fact, the majority of the literature I review in the introduction is not really literature on the corporation at all; for the most part I review literature in political theory that does not deal with the corporation, and discuss how my approach either complements or conflicts with those studies. This will both make clear some of the research communities that I wish to address, as well as give a better understanding of how my intellectual commitments influence my approach to the problem of the corporation.

The Corporation: Some Key Features

To begin with, I should mention that I am concerned specifically with the modern business corporation. In their influential account of corporate law Hansmann and Kraakman have
put forward five basic features of the standard public corporation, which are essentially constant across most legal jurisdictions. These are worth explaining as they paint a basic empirical picture of the corporation as it exists in most liberal market economies. Understanding these five features also help us to understand why the corporation is generally understood to promote efficiency and productivity. I here explain these 5 features, only offering some qualification to the 5th.

1. Legal Personality

Following the U.S. Supreme Court’s 2010 *Citizens United* decision this is probably the most infamous feature of the corporation. It is also likely the most misunderstood. Legal personality does not mean that the corporation is in fact a natural person, or that it has the same rights and privileges as natural persons. As we will see in Chapter 6, there are various theories interpreting what the corporation’s legal personality implies theoretically and politically. However, at a minimum all legal personality means is that a corporation is an entity that the law recognizes as capable of bearing certain rights and responsibilities. In particular, corporations have the ability to own property and enter into binding contracts; they also can be sued, taxed, and be made to fulfill obligations of debt and contract.

The significance of this is that when we say the corporation can do, or be required to do, any of these things, we are drawing a distinction between the corporation as an entity and those parties who work on its behalf or own shares in it. If the corporation owns property then this property is not to be seen as jointly owned by its various patron groups -- it literally belongs to

---

the corporation. A holder of some 45% of the corporation’s shares does not own any of the corporation’s property, nor does the CEO who directs the use of corporate property, nor the workers who might use that property. Similarly, when the corporation enters into a contract, agreeing for instance to pay a certain amount to some person, then it is the corporation that it bound by this contract, and which then owes this person the money. This idea—commonly called “asset partitioning”—is the key aspect of legal personality; corporate assets are held to be distinct from the assets of the corporation’s patron groups.

2. Limited Liability

Though not a sufficient condition, the asset-partitioning aspect of legal personality is a very important enabling condition for shareholders to have limited liability. If the corporation is a legal person capable of owning its own assets then this allows for the possibility of arranging things such that claims against the corporation can be made on only those assets owned by the corporation. Often it is said that corporations have limited liability, but this is actually to misspeak; it is more accurate to say that the shareholders who invest in the corporation have limited liability (as do, it should also be noted, most corporate employees). All that shareholders can lose on claims made against the corporation is the money initially invested in the

---

2 I use the term “patron group,” following Hansmann, to refer to any of the groups that directly deal with the corporation (customers, shareholders, creditors, workers, raw goods suppliers, etc.). The term is not meant to evoke the patron-client relationship, and is similar to how many scholars use the term “stakeholders.” I use the term “patron group” for two reasons. First, it differentiates people who deal directly with the corporation in some formal way with those who are simply affected by a corporation’s actions (the way “stakeholder” sometimes means). Secondly, it is a helpful term because it forces us to explain why one particular group of contractual relationships ought to be singled out as having a special role within the corporation; we start with the idea that, say, workers and shareholders are both corporate patron groups, and then must explain why one or the other (or both or neither) should be given special privileges and responsibilities.
corporation; because they have limited liability, their personal assets are safe from seizure. ³ Shareholders are not liable for claims against the corporation, despite their ownership of the corporation’s shares. Only those assets invested in the corporation are subject to claims by creditors.

Again, it is worth stressing that the legal personality of the corporation does not require that limited liability be granted to its investors. Early corporate statutes granted the corporation rights-bearing status but still left shareholders proportionately liable for the corporation’s debts. Limited liability began catching on throughout the 19th century as a means to stimulate the United States’ developing manufacturing base as a response to the diminishing of British imports. ⁴ Its effectiveness in this aim—witnessed both by the spread of limited liability throughout all the United States and the meteoric rise of the United States as an economic power—is due to its incentivizing investment by shifting risk. More will be said about this in chapter 6, but it is worth quickly explaining the economic logic here.

Absent the stipulation of limited liability, would-be shareholder will be hesitant to invest, due to the possibility of being personally on the hook for significant corporate losses. The result is that a significant amount of potential investment will be foregone due to normal risk-aversion; limited liability reduces the risk for shareholders by enabling them to invest without having to worry about complete personal ruin. It also allows shareholders to further reduce their risk through diversification, by investing in many different firms and sectors. (Without limited liability, this strategy would increase, rather than decrease, risk.) Though risk is reduced for

³ As always, caveats abound. There are instances where courts “pierce the corporate veil” and assign liability for certain debts to patron groups. This, however, is the exceptional case, and there is no clear consensus on whether a principled or doctrinal justification for veil-piercing exists.

shareholders it is not necessarily reduced generally. As David Moss notes, limited liability actually implies a tradeoff between the magnitude of shareholder loss and the probability of creditor loss.\(^5\) That is, though shareholders are protected from complete devastation of their private assets, parties who loan money to the corporation, or have some other financial claim on it, become more likely to suffer losses; limited liability restricts the extent of their claims and therefore reduces the available assets that can be used to pay back an obligation. Why does this result in a net-positive effect on investment? It would seem that we would be freeing up equity capital at the expense of debt capital. However, because people tend be disproportionately averse to large losses compared to small losses, the slight decrease in debt capital (or its higher cost) will be outweighed by the increase in shareholder investment. As Moss puts it, “shareholders may well gain more satisfaction from the sharp reduction in their maximum possible loss than creditors lose from the corresponding increase in their probability of a more limited loss.”\(^6\) The historic economic success of limited liability to stimulate investment seems to offer a prima facie reason to think that there is good economic sense to such a legal innovation. Whether such a regime can be justified from the standpoint of justice, however, will be explored later in the dissertation.

3. Transferable Shares

Again, asset partitioning is an important enabling condition for there to exist the transferability of shares. Because personal and corporate assets are distinct, with the latter being owned by the corporate person, it is possible for the corporation and its operations to remain

\(^5\) Ibid 74.
\(^6\) Moss, 75.
relatively constant despite shareholders changing. Shareholders can be understood simply as owning two distinct rights-claims upon the corporation: the right to claim dividends when they are distributed; and voting rights on particular corporation decisions—generally for the board of directors, as well as major and extraordinary strategic decisions like mergers and buyouts. That these shares are transferable means that corporate management, operations, sales, and the like continue uninterrupted despite these rights being transferred to other people. The actual identities of shareholders normally have no effect on the ordinary business or day-to-day operations of the corporation, and actual shareholders need not weigh in on such affairs. According to the standard economic story, this allows for greater liquidity of shareholder capital, enabling greater diversification for shareholders, and greater possibility to attract investment for corporate ventures.

4. Centralized Management with a Board of Directors

The last three characteristics have dealt with how the corporate entity and its assets are distinct from other patrons of the corporation. However, there is also a general tendency for how these various parties relate to one another and the corporation. Essentially, the corporation is represented by the multi-member board of directors. The main task of the board of directors is to appoint and monitor the corporation’s management, who oversee both its day-to-day operations and strategic direction. They are hired by and accountable to the board of directors. The board of directors, on the other hand, is elected by the corporation’s shareholders, generally at stipulated election periods. The result is a structure defined both by formal linkages of accountability

---

7 Though, as we will see in chapter 6, stock classes can be differentiated so that some stockholders have different claims on dividends or attenuated voting rights.
between shareholders, directors, and management, and a formal distinction of identity and function between all three: shareholders elect the board of directors but they generally do not determine corporate direction by plebiscite; the board of directors hires corporate management, but does not generally engage in the ordinary business of the corporation; and neither management, directorship, nor the shareholders are understood to be the corporation, but rather relate to it either as representatives (the board of directors), agent-employees (management), or rights-claimants (shareholders). I depict this model below in Figure 1.

**Figure 1. Relationship between Shareholders, Directors, and Management**

![Diagram of corporation's relationship between shareholders, directors, and management]

---

5. Ownership by Capital Contributors

This last aspect of the corporation is perhaps the most characteristic and also the one that needs some qualification. This trait in fact encompasses two distinct claims: 1) the shareholders
of the corporation are its owners; and 2) people come to be shareholders by virtue of contributing capital (or, because of #3 above, by buying shares from those who have contributed capital). The first claim has been challenged by many, from various ideological stripes, who argue that the corporation in fact has no owners. The corporation owns its assets and various parties have rights and duties relating to that entity, but nobody can be said to own it. These claims are most likely correct, but for now they are not terribly important. Whether shareholders are in fact owners of the corporation might have moral consequences, depending on whether we assign normative importance to ownership; from a functional perspective however, it is not important whether or not shareholders are literally the owners of the corporation. All that matters is that we recognize them as having rights commonly associated with ownership, namely control rights and rights to residual profits.

It must be emphasized that the second claim—that these rights are to be allocated to those who provide capital—is an empirical trend and not a conceptual necessity. As Hansmann notes, the ownership rights of the corporation can be allocated to customers, workers, or raw goods suppliers just as easily—we refer to such corporations as cooperatives. Indeed, to break the aura of distinction between corporations and cooperatives Hansmann refers to corporations as “capital coops”—cooperatives that have simply allocated ownership rights to its capital providers. As we will see in chapter 6, I think this characterization is slightly misleading. For now, however, all we need note is that ownership rights in the corporation are generally assigned to those who invest capital, but this need not be the case. As a result the corporation is much more open to

---


organizational variety than is generally thought—since most think of cooperatives and corporations as utterly distinct. Furthermore it may be possible that such organizational variety could exist to a greater degree than is reflected in the current make-up of corporate structures. This will be revisited in chapter 5.

Although exceptions exist, these five traits make for a useful definition of the business corporation: all business corporations will have these five traits, and all entities that have these five traits will be business corporations. As I have already suggested, the commonly accepted justification for the corporation is that it increases investment and economic productivity. I will argue later that such a justification is neither obvious nor sufficient. It is not obvious because the organization of the corporation runs counter to the way the economic advantages of the market are conceived. We therefore need a more nuanced story about the economic benefits of the corporation than their mere ability to stimulate investment. This justification is also insufficient because even the most efficient and productive institutions might be immoral or unjust in some other respect. The corporation therefore needs to be defended in terms of some value other than efficiency such as legitimacy, morality, or justice. As I explain in chapter 1, my dissertation is organized around providing these two justificatory accounts: an account of corporate productivity and efficiency, and an account of the corporation’s normative legitimacy (or the lack thereof).

Politics vs. Economics
Recently Jacob Levy has offered a political theoretic account of intermediary or “meso-level” institutions like corporations.\(^\text{10}\) Levy argues that within liberal thought there are generally two ways of understanding the intermediary corporation. “Rationalist liberalism” views these institutions suspiciously as potential sources of domination, and therefore contends that states must intervene into the life of these institutions in order to protect individuals. “Pluralist liberalism,” on the other hand, views such institutions as being important because they provide individuals with the ability to resist state overreach. Where rationalists see the protection of liberty as residing within the state’s ability to limit and affect the internal life of meso-level institutions, pluralists see the protection of liberty as residing in the meso-level institutions’ ability to limit and affect the reach of the state. Levy claims that these basic tendencies are irreconcilable; the tension between the pluralist’s celebration of intermediary institution and the rationalist’s skepticism is a tension inherent in liberalism.

Levy’s account has been influential for my own, and I will discuss it in more detail below, and in chapter 6. I recount the argument here, however, because the distinction between viewing the corporation as an extension of individual liberty or as something harmful to individual liberty maps on well to the distinction between economic approaches to the corporation and political approaches; the former sees the corporation as simply an extension of individual endeavor, and the latter sees the corporation as an extension of state power. The tension between these two approaches to the corporation—and, in some sense, to social phenomena more generally—is a tension that I try to navigate throughout the dissertation. My general contention, similar to Levy’s, is that both the economic emphasis on individual incentives and the political emphasis on power and social structuring are necessary for

understanding most things; we should be hesitant of approaches that attempt to subsume one fully into the other.

I will note here that by virtue of focusing on the business corporation, I focus on precisely the type of corporation that Levy excludes from his study, and I do not engage with those that he does: the church, the university, the guild, or the city. It should also be noted that from a historical perspective such a distinction is somewhat arbitrary. Prior to the mid-19th century the distinction between a business corporation and, say, a municipal corporation was quite murky. While early American corporations included familiar institutions like banks, insurance companies, and canal-digging companies, they also included Harvard College, Philadelphia, the Connecticut Medical Society, and the Massachusetts Charitable Mechanic Association. There was legally little to no distinction between these corporations, and a similar lack of discrimination on the part of their critics. While many were outraged by the granting of the corporate charter to the Bank of New York, this was not what founding American republican Sam Adams denounced as the “introduction of aristocracy” and antithetical to republicanism. Adams was referring to the incorporation of the city of Boston.

The development and rise of the business corporation as distinct from other corporations points to a theme that has become common fare for critics of capitalism: that the separation of the political from the economic is not a given, but is in itself a political project. Indeed, the distinction between politics and economics is fairly recent. Blaug notes that such rigid categorization was not possible in classical political economy due to the central place that

population management held in the postulates and purposes of economic inquiry; as we will see in chapter 2, it was not until the so-called marginalist revolution of the late 19th century that population as such could be seen as a variable, as opposed to the object, of economics.\textsuperscript{14} Stark further contends that this distinction was solidified by disciplinary lines in what he calls “Parson’s pact,” an agreement between Talcott Parsons and Lionel Robbins that served to distinguish the field of economics from sociology.\textsuperscript{15}

Many political scientists have sought to overcome this distinction by attempting to subsume political science within economics, developing formal models of political phenomena and viewing things like elections and legislative activity as cases of economic agents acting within markets for specific political goods. I do not wish to disparage this vast field of scholarship; in the pages that follow I draw heavily on economic models and studies that start with the rational actor assumption. Still, my dissertation fits better in the tradition that tries to overcome the distinction in the opposite fashion, by showing how economic institutions, behavior, and actions are in fact political by nature and can best be explained by the political scientist’s analytic concern with power.

One method of troubling the political-economic distinction is historical in nature. This approach focuses on various developments in the 19th and 20th centuries that shine light on how the modern capitalist economy was constructed and produced as a political and ideological project. Polanyi’s famous work on the transformation to industrialism demonstrated that the market economy was a conscious large-scale social project, created for the benefit of particular

class interests, and decidedly not a natural phenomenon.\textsuperscript{16} Similarly, Scott and Graeber have historicized economic concepts of property and exchange, showing how these practices are thoroughly modern and developed out of the large-scale ambitions of powerful actors.\textsuperscript{17} Complementing these inquiries into the historical development of market capitalism are the studies of the historical development of economic thought and concepts employed by the economics discipline. The most influential of these is surely Foucault’s Lectures at the College de France, which have stimulated interest in his genealogical account of neoliberal economics and its relationship to particular practices of government.\textsuperscript{18} I will discuss Foucault’s account of Chicago economics in more detail in chapter 3. For now all we need to note is that a number of scholars have contributed similar types of ideas, including historical understandings of how the concept of “the economy” was created,\textsuperscript{19} and the linkages between the rise of rational choice theory and the cold war.\textsuperscript{20} These historical accounts all serve to “breach the self-evidence”\textsuperscript{21} of the distinction between economy and society, economics and politics, or economic science and political science. As a result we are able to understand the distinction itself as something political.

While each of these writers places emphasis on a different part of the story—the political, the academic, the ideological, and so on—they all share a common narrative of markets and economic logic being extended into domains previously untouched by such analysis; similarly, they all share an understanding of free-market economics, and the ideologies used to justify them, as constructed political projects. In a similar manner, I am arguing that the common contemporary view of the corporation as merely an economic institution—fully understandable and justifiable on economic grounds—is the result of a particular political project. Now, the fact that we can historicize something and show it to be the result of political factors does not mean we have shown it to be bad or unjust; contingency does not imply illegitimacy. However, it does raise the question of whether it is a good political project, a question that is harder to ask if we take something for granted or naturalize it. My project is distinct from the historical studies above in putting less emphasis on genealogy and more emphasis on the normative question of whether the development of the business corporation is a good project, and whether its economic account and justification is a compelling one. Though I take care to historicize the theoretical developments with which I grapple (especially in chapters 2 and 3) so that we can understand their political purposes and orientations, my emphasis is on presenting these economic practices and theories in the most charitable manner, and then appraising and criticizing them from a normative perspective.

The Limits of Rawls

---

By emphasizing principled normative criticism and prescription over historical or discursive critique, this study can be thought of as being in dialogue with the tradition of political philosophy developed since Rawls. Insofar as almost all Anglophone normative political theory aimed at institutions is influenced by Rawls, placing my study within that tradition may seem as trivial or platitudinous as saying a rock band is influenced by the Beatles, or a jazz trumpeter plays in the tradition of Louis Armstrong. More substantively, this study can generally be placed within a particular reading of the Rawlsian project. Following people like Ron, Laden, Chambers, and Forst, I view Rawls as a critical theorist whose key achievement was articulating the contours of a more substantively-egalitarian political order using norms and ideas already immanent in liberal democratic societies. Both in terms of the aim and the method of critique, this study is very much influenced by the Rawlsian project. As I explain in chapter 1, I build my “political theory of the corporation” upon normative commitments that I take to be implicit in our own liberal social orders, in order to show how contemporary corporate governance fails on those terms. I therefore take my approach to the corporation as roughly Rawlsian in that it criticizes a key institution in liberal democracy on the grounds of egalitarian values that are inherent in liberal society itself.

Because of this influence one might expect to me take Rawlsian normative concepts, like the veil of ignorance or the difference principle, and apply them to the corporation in order to derive my critique. After all, in addition to being concerned with institutions and how they might be structured fairly, Rawls also recognized and made powerful arguments on behalf of efficiency.

---

in political economic institutional arrangements. Rawls, then, seems the perfect candidate for offering normative direction on the corporation. What makes this task difficult is that Rawls was relatively silent regarding questions of corporate governance, corporate law, and business ethics, as well as on a host of other questions about the internal dynamics of the corporation.

This silence is presumably due to the restricted application of Rawls’s principles of justice to the “basic structure” of society. Now, there is much debate over precisely what this basic structure encompasses. In contrast to monists like G.A. Cohen, I do not think the basic structure can include all pervasive forms of power. This is because of Rawls’s commitment to articulating his theory in “political” terms, a commitment that entails distinguishing between public institutions that are necessary for governing social cooperation and other kinds of association, as well as between individuals’ private comprehensive doctrines and the principles appropriate for governing a pluralistic society. This requires a more limited understanding of what the basic structure can be. This is all despite Rawls’s impulses to the contrary; in responding to Okin’s criticism that his theory cannot address gendered inequality in the family, Rawls contends that the principles of the basic structure restrict the organization of family life. If this were true this would give us reason to think that the corporation could be included in the basic structure. However, on the next page, Rawls distances himself from such an idea: “the principles of political justice are to apply directly to this [basic] structure, but they are not to apply directly to the internal life to the many associations within it, the family among them.”

He later notes that another one of these associations whose internal life is not to be affected directly by the basic structure is the business firm, an echo of the explicit distinction he draws between the basic structure and the “rules of corporate associations” in his earlier Theory of

---

Justice.25 The result is that the basic structure must be understood narrowly as the legal and statutory institutions of society—that is, the state—so that the principles of justice can be “political” in the sense that Rawls means it.

With this reading of the basic structure, it is not clear how the corporation could be brought within the purview of Rawls’s theory. As we will see in chapter 4, there have been attempts to assimilate the corporation within Rawlsian theory, some of which are fairly nuanced and sophisticated, but which I nevertheless will contend ultimately fail. In the end, Rawls casts his critical net too narrowly, leaving institutions like the corporation outside of his analysis.

Insofar as my argument offers compelling critical and normative accounts of the corporation, we may see this as a friendly critique of the Rawlsian paradigm: in order that we might bring philosophical critique to bear on all of the relevant institutions and practices of our social order we must abandon the methodological focus on the “basic structure,” and look to meso-level institutions as important sources of legitimacy and/or potential domination. The challenge is doing this in such a way as to respect Rawls’s correct concern for reasonable pluralism, the burdens of judgment, and the limits to political justice.

Both Political and Economic

By rejecting the straightforward application of Rawlsian principles to the corporation, I mean to signal my resistance to a more general impulse. This is the impulse of responding to the increasing emphasis on the economic by trying to fold everything back into politics and government to the exclusion of other considerations. As I noted earlier, this is similar to the rationalist impulse that Levy discusses; if we see the corporation as an extension of government

then we can use government to constrain the corporation. This impulse has shown itself most strongly in recent literature on normative business and management studies. One view characteristic of this impulse sees the corporation simply as equivalent to a governmental agency and sees corporate executives as something very much like governmental officials or holders of public office. Another line of research contends that, given the transnational condition in which the state has receded in importance, corporations now operate like states in many instances, requiring them to take on the moral constraints of political legitimacy. As a countervailing force to the mainstream emphasis on profit-maximization, these normative theorists have felt compelled to assert the primacy of social and political values as a means to bring the corporation within the purview of theoretical reflection.

In political science such a view is exemplified by David Ciepley’s ground-breaking article on “the political theory of the corporation.” Ciepley attempts to avoid the politics-centric view; indeed, as he announces in his title, he wishes to situate the corporation “beyond public and private.” Ciepley’s overarching claim is a welcome one: despite their dominant place in liberal orders, corporations are not in fact liberal institutions:

While not simply private, corporations are also not simply public, since, unlike armies and government bureaus, neither their financing, staffing, nor direction come from government. Indeed, corporations transgress all the basic divides that structure liberal treatments of law, economics, and politics: government/market, state/society, privilege/equality, status/contract, as well as liberalism’s master dichotomy of public/private. Corporations are not of liberalism and cannot be satisfactorily assimilated to its categories. Instead, they need to be placed in a legal and policy category of their own—neither public, nor private, but “corporate”—to be governed by distinct norms and rules, so as to render them more intelligible, more accountable, more

---

responsible, and more productive. Developing this category of the corporate will be the central task of the political theory of the corporation going forward.\textsuperscript{29}

This could very well serve as the introductory epigraph for the study that follows. I too seek to show that the corporation as currently structured is not easily captured by liberal values or categories, that it sits uneasily between public and private, government and market. I also want to argue that corporations should be dealt with as corporations, as a distinctive category of social institution.

Despite this clarion call for a new approach to the corporation, Ciepley later argues that corporations are “government colonizers” of the market, and puts central emphasis on their governmental provenance. Although Ciepley wants to contend that corporations are something distinct from both state and market, he does this is by playing down their private or contractual elements, and describing them as, primarily, “governing entities.” While Ciepley provides compelling arguments for seeing the corporation as having a governmental nature, he is far too dismissive of their private or economic bases. On Ciepley’s own terms it is not clear why the corporation could not be understood in purely public terms, as some sort of autonomous governmental entity that pursues ends on its own terms in a bounded manner.

The missing aspect of Ciepley’s argument is an answer to the question of why individuals choose to incorporate in the first place. We can accept Ciepley’s historical account of the corporation as a governmental and mercantilist institution as well as his arguments that the corporation, today, cannot be made sense of only in economic or contractual terms; there is more than a gain of truth in this and in the governmental views of the corporation mentioned above. But it is not as though we still live in the days where corporations are imposed on populations by

\textsuperscript{29} Ibid 140.
monarchical prerogative or legislative grant. The great ascendency of the corporation has come in the age of general incorporation laws, where individuals freely choose to incorporate, as well as choose to develop innovative ways of structuring the corporation. The fact that government creates a particular legal vehicle does not mean that citizens will use it; corporations, then, come into being not only through governmental fiat, but individual initiative as well. And insofar as voluntary individual action is crucial to the actual creation of corporations, there has to be a private or economic story to tell about corporations in order for it to be complete.

To put it differently, the corporate entity cannot be understood solely in rationalist terms as an extension of government, or in pluralist terms as merely an extension of individuals. The former cannot account for individual initiative in the corporation’s development, and the latter cannot account for the kinds of power and compulsion that develop in these kinds of institutions. We must understand corporations as distinct from both states and individuals, and we can do this by understanding how corporations have both political and economic facets.

In this thesis I give an account of the particularly corporate nature of the business corporation by first attending to this economic story. In chapter 2 I explicate the Coasian theory of the firm, which goes a long way to explaining why individuals would choose to incorporate for purely economic reasons—or, to put it even more sharply, why corporations continue to exist in the market despite their non-compulsory nature. The crucial point here—which Ciepley does not touch upon—is that corporations continue to exist and survive in the market because they provide important efficiency gains over markets in coordinating production. The problem with the Chicago school’s theory of the corporation, however, is that they make efficiency gains the entire story, as I show in chapter 3. Chapters 4 through 7 attempt to articulate a political theory of the corporation that takes both sets of concerns seriously—the private, contractual, and
economic on the one hand, and the public, legal, and governmental on the other. I address the question of how to reconcile the public and private aspects of the corporation head-on in the discussion of corporate law in chapter 6.

I believe the best way to criticize the deleterious effects of economic processes is not to attribute them solely to some nefarious project. The critic should first seek to understand those economic phenomena, and try to understand why they exist and persist in precisely the economic terms they wish to criticize. Of course, none of this is to deny the crucial importance of politics to such processes; I would not have written the dissertation that I have if I did not think that economic processes are shaped and structured by the three I’s of politics—interests, institutions, and ideas. Still, if we attribute everything to political influence then we risk reacting to the reductionism of economic imperialism by responding in kind. Instead we should begin with charity, seeking to understand the logic underlying economic theories, and attempt to interpret them on their own terms, as they would want to be understood. With such an understanding we will be in a better position to offer critique, likely finding our initial political impulses tempered in the process by insights we had not thought of before.

This conviction underlies my general approach to this dissertation and its structure. I first go through the economic theories of the corporation. Though I ultimately show them to be flawed, I contend that their emphasis on efficiency is something that normative and critical theorists must take seriously. I then develop my own theory in the last three chapters; in doing so I try to take seriously both the economic and moral concerns that one should have with respect to the corporation.
The hope is that such a study will contribute to our normative understanding of this central institution and perhaps serve to inspire some change towards a better and more just economic order. In the process if I can make the economic theories of the corporation more legible to political theorists—or, more improbably, make the moral concerns of political theorists more legible to economists—then this will seem to me no small feat.
Chapter 1. Assumptions and Approaches

When setting out on a project one often looks for exemplars to follow. In the canon of political theory, by far the most thorough investigation of the relationship between the political economic order and those economic enterprises which exist within it is Marx’s *Capital Vol. 1*. The first two sections of the book (“Commodities and Money” and “The Transformation of Money into Capital”) are dedicated to understanding the general processes of exchange and distribution in bourgeois society’s “sphere of simple circulation”; the rest of the book is dedicated to the process of production that happens within the economic enterprise. Presenting them as being in contradiction with one another, Marx claimed to have uncovered the contradictions underlying capitalism, contradictions that would eventually undermine the capitalist mode of production itself.

Since the days of Marx there have been numerous political and philosophic inquiries into the subject of the first section of *Capital Vol. 1*, the nature of capitalism, its institutions and tendencies, and the normative dilemmas it raises. In contrast, there has been precious little by way of inquiry into the subject of the second half of *Capital*, a philosophical and systematic understanding of the productive enterprise’s *internal life*. This project, by considering the internal dynamics of the corporation from the perspective of political theory, attempts to address this situation, albeit in a very different manner than Marx did. In order to do this I first set out some basic presuppositions that, I hold, are the normative bases for the liberal democratic order—the “exclusive realm of Freedom, Equality, Property, and Bentham” according to Marx—in which we live. I start with them not so much because I think they are correct (though I do think there is much to be said in their favor) but because I take them to represent the moral framework of our liberal democratic order. By holding these norms and institutions constant we will be able to understand more precisely the types of puzzles the corporation raises and, hopefully, new ideas of how corporations ought to be structured in liberal democracy. This is roughly the approach Carens has proposed for answering normative questions that address real and practical
problems without foreclosing the possibility of a more radical critique.\textsuperscript{1} Therefore, I present these circumstances not simply as a historical accident but as having a normative basis that must be taken seriously. By first laying these out as clearly as possible, I will be able to establish what these principles must demand of the corporation and its organization.

For now we can restrict these suppositions to two: 1) liberal democratic institutions and norms—individual liberty, formal equality, and democratic decision-making—are to be preferred to others; and 2) there is a form of capitalism—a market-based political and economic order—that satisfies our intuitions regarding justice. In the two sections that follow I will unpack and explain these presuppositions respectively. After this I will briefly explain how these presuppositions create something of a puzzle for understanding the corporation, and then delineate the main questions needing answers in order to resolve this puzzle. A coherent answer to all of these questions, I argue, constitutes a political theory of the corporation, the subject of the remainder of my dissertation.

**Presupposition 1: Liberal democratic norms and institutions are preferable to other alternatives.**

This first presupposition is meant to articulate the normative basis for the liberal democratic approach to the relationship between individual and society. Now, the canons on liberalism and democracy are long and complex and I don’t pretend to do them justice here. Primarily what I wish to do is provide workable understandings of what I consider to be the three main pillars of liberal democracy—individual liberty, formal equality, and democratic procedures for collective decision-making— and the moral intuitions I take to underlie them. These are not meant to be defences of liberal democratic norms against other norms—I am not here trying to vindicate a Mill or a Kant against a Nietzsche or a Marx—but brief statements about what I see as the basic and minimal moral intuitions,

\textsuperscript{1} In so doing I follow the basic methodology used by Carens to explore the ethics of immigration. See “Presuppositions and Political Theory” in Joseph Carens. 2013. *The Ethics of Immigration*. Oxford: Oxford University Press.
norms, and practices underlying liberal democracy.

**Individuality**

There are two components to this individuality. On the one hand there is the normative approach to politics whereby political society is seen as justifiable insofar as it protects the moral primacy of the individual and her liberty. On the other hand there is the empirical and methodological approach to social sciences whereby “all social phenomena are in principle explicable only in terms of individuals – their properties, goals, and beliefs.” For this reason, liberal democracy has generally grounded its legitimacy in the consent of rational individuals (whether real or hypothetical) for the purpose of preserving the plural values and ends individuals will wish to pursue.

The Rawlsian articulation of liberalism can be seen as doing precisely this. By presenting a stylized account of how individuals would choose a certain scheme of social cooperation based on a particular articulation of their “properties, goals, and beliefs,” Rawls argues that our liberal institutions and norms can be seen as the result of a hypothetical social contract to which we as individuals would consent were we properly positioned to do so. Furthermore, he argues that this agreement is in harmony with the diverse ends and reasonable pluralism by which a free and modern society will be characterized. Where the contractarian original position of *Theory of Justice* gives way to the overlapping consensus of *Political Liberalism*, we find an emphasis less on the hypothetical consent of individuals and more on how the content of such a doctrine can be endorsed by people holding radically different understandings of the good. Presenting people as both rational and reasonable, Rawls claims that we are best able to pursue our various goods in a liberal society governed by a political conception of justice which does not rely on metaphysical doctrines.

---

3 I am particularly fond of Ian Shapiro’s articulation of this idea. See Ian Shapiro. 1986 *The Evolution of Rights in Liberal Theory*. Cambridge: Cambridge University Press.
As I have said in the introduction, Rawls’s theory is not suited to the task of dealing normatively with the corporation. Still, for articulating the normative ideas implicit in current practices and institutions, Rawls is quite helpful, particularly because he claimed his theory to be doing precisely this.\(^4\) There are two features of the Rawlsian project to which I want to draw attention. The first is that our understanding of liberal political institutions like rights and constitutions are grounded in the acknowledgement that society is a scheme of cooperation made up of distinct individuals. Because cooperation benefits everyone involved, everyone must in principle be able to consent to the particular scheme of cooperation at hand; in arguing that behind the veil of ignorance we would agree to the liberty principle, Rawls contends that there are certain dimensions of our selves that we would want protected, regardless of the benefits and harms particular states of the world would bring us. We reject approaching the problems of liberal society with a utilitarian calculus because the interpersonal comparisons of utility such a calculation requires would be empirically fraught and morally unpalatable, due to its blindness to the distinction between persons.\(^5\) As a result, the normative defense of liberal institutions like rights presented here is distinct from someone like Mill’s who argues that we should understand rights as those conventions we adopt because their general acceptance tends to increase overall utility.\(^6\) Whereas Rawls's is based on a concept of the individual and what the individual seeks (albeit a highly abstract concept), Mill's account of rights is not from the perspective of the individual but from the perspective of what increases the welfare of a society taken in the aggregate. It is not clear why a right, once defended on consequentialist grounds, could not be trespassed or made trivial in the face of a utilitarian calculus that showed it to be in the general interest to do so.


The second point from the Rawlsian perspective is that the privileging of particular individual interests above some aggregate notion of social welfare leads to the acceptance of reasonable value-pluralism. That is, because we view certain things like freedom of conscience, expression, and religion as protected individual rights, political society must be such that people who use these freedoms for different ends can coexist and cooperate with one another. Just as there can be no one metric according to which the rights and liberties of citizens may be subject, the liberal political project aims for the ideal that there is no one right way to live one’s life or to pursue the good.⁷

*Formal Equality*

While the permissibility of different forms of material inequality are debated amongst various forms of liberalism, I take it as a given that liberalism endorses the formal equality of persons both in a legal sense and in a moral sense. Legally, this idea is best captured in the phrase that all are “equal before the law.” This phrase implies that all citizens⁸ within a state are subject to the same laws to the same extent, and that the same due process of law will, or should, be applied to all in the same manner. In the eyes of the law, and to the extent that persons are affected by the law, all are equal as citizens.

More than this, however, I take the liberal project to also imply a kind of moral equality amongst persons. Here the contrast with the legal equality of citizens draws our attention to the more fundamental idea of liberalism, namely that there is no basis according to which we ought to treat one person as being of greater moral worth, or being of greater dignity, than any other. While there are many ways of justifying this—by reference to people's abilities (as Hobbes claims), divine provenance (as Locke claims), common rational capacities (as Kant claims), or for want of some religious or

---


⁸ There is good reason to think that liberalism implies the equal recognition of non-citizens as well (see Carens, *Ethics of Immigration*, Ch. 5). The view being presented here is friendly to this idea but does not require it all; at minimum, liberal democratic norms implies the formal equality of citizens at the very least.
metaphysical agreement upon which cooperation can be based (à la Habermas or Sen)—it is impossible to imagine a concept of liberal democracy that does not contain the basic notion that people are deserving of equal treatment from others.

This follows from what was said above regarding the protection of individuality. If the individual has moral primacy, and this moral primacy is articulated in a discourse of rights, then this means that people must approach all other people with the same restraints such rights imply, thus necessitating an idea of legal equality. Furthermore, if such protection of individuality necessitates the preservation of value pluralism, then our cooperative social endeavours cannot be grounded on one particular notion of the good; instead we must presume some notion of equal worth to exist prior to political convention, at least as a pragmatic manner, thus requiring some notion of moral equality. Now it should be made clear that I am not claiming that liberalism cannot entail a commitment to a more robust notion of equality, either of resources, of utility, of capabilities, or some other equalisandum. Many (including myself) think that liberalism, in order to be coherent, must also entail a commitment to a certain level of material equality. However, certainly not all liberals do, and even those who do still seek to defend such a claim with respect to formal equality and individual rights.

**Democratic Procedures**

From the foregoing we get the liberal idea that when we approach the question of social cooperation, we must begin with the baseline assumption of equality amongst individuals. Yet the institutions necessary for coordinating cooperation and collective action require some forms of hierarchy that will inevitably confer status upon some so as to create inequality both in terms of status and in terms of obligation. What is a liberal to do? For this reason, democratic procedures are also a key component of liberal democracy (as the name would imply).

Now, again, there are various strains of democratic thought ranging from the substantive, to the
participatory, to the market-oriented. I take Dahl's understanding of democracy to be the most basic model of 20th century views of democracy, one endorsed as a minimum, whatever other features we might demand of democratic institutions. Dahl and Lindblom describe democracy both in terms of a goal and a principle. The goal of democracy is the “condition of political equality,” the political actualization of liberal formal equality wherein “control over governmental decisions is shared so that the preferences of no one citizen are weighted more heavily than the preferences of any other one citizen.” This leads to the principle of majority rule: “governmental decisions should be controlled by the greater number expressing their preferences in the ‘last say.’” As Dahl later clarified, the moral bases of this democratic goal are the liberal ideas of “equal moral dignity,” “autonomy,” and “strong personal equality”; affirming these ideas behooves us to strive for democracy, even if in the last analysis we fail to meet its high standards.

Importantly for Dahl, the fulfilment of this goal is cast in terms of institutions and procedures, not in terms of discourse-conditions or policy substance. The intuition here is that institutional and procedural standards are better for analyzing real-world democracies than more substantive measures of democracy; these provide a more realistic goal toward which democratizing communities can aspire, while not necessarily contradicting the goals of advocates desiring more substantive or discursive visions of democracy. Perhaps more importantly, the emphasis on procedure and institution takes seriously the prior claim that protection of the individual must also be a protection of value pluralism. Saffon and Urbinati have argued that due to liberalism’s “diverse, at times irreconcilable, opinions, imposing a substantive standard to democratic decisions may threaten freedom.” This necessitates a


10 Robert Dahl. 1989. *Democracy and its Critics*. New Haven: Yale University Press. Famously, Dahl offered the term “polyarchy” for those states which approached, but failed to reach, the ideal institutional fulfillment of democracy. This term does not seem to have caught on in the same way that the rest of his ideas have.
minimal set of procedures that can be appealed to in inevitable instances of conflict.\textsuperscript{11} These minimal procedures ensure that citizens have the ability to formulate preferences, signal their preferences to fellow citizens and the government, and “have their preferences weighed equally in the conduct of the government, that is, weighted with no discrimination because of the content or source of the preference.”\textsuperscript{12} Procedural democracy thus allows for different ideas and policies to compete for public support while also being sensitive to issues of individual freedom and equality.

The liberal affirmation of democracy is therefore not borne from a concern for solidarity, or giving voice to a “general will,” but from the protection of individual freedom and equality. The basic impulse is that when individuals are subject to a situation in which their freedoms might be trespassed or in which inequality obtains, there ought to be procedures that can effectively allow them to communicate their preferred ends and have those preferences count as equal to others in the formulation of collective decisions. Whatever else liberals might demand of democracy—for instance a particular idea of justification or deliberation which, again, is something to which I am sympathetic\textsuperscript{13}—there must be procedures in place that can protect against infringement upon individual freedom and equality.

**Presupposition 2: There is a good form of capitalism that satisfies our intuitions regarding justice.**

The first presupposition is meant to articulate the basic moral intuitions underlying our political institutions that protect individual rights, legal equality, and democratic control through participation and competitive elections. This second presupposition is meant to articulate the moral intuitions


underlying the institutions through which we allocate, produce, and distribute goods. The way that I have stated this presupposition will probably seem unsatisfactory for at least three reasons. First, because there are many forms and institutional designs of capitalism (perhaps as many forms of capitalism as there are capitalist societies), this presupposition might be taken to be essentially vacuous. It thus invites the question as to what I mean precisely by capitalism and whether such a definition is over-encompassing. Secondly, this presupposition is vague with regards to what precisely is taken to be good about capitalism. Third, beyond merely claiming that there is a good that capitalism can achieve, one would want to know why capitalism is particularly well-suited to promoting this good. Here, I answer these three questions in turn: Conceptually, what is my definition of capitalism? Normatively, what is the good which capitalism promotes? Empirically, why should we believe capitalism capable of promoting this good?

Defining Capitalism

Part of the difficulty of defining capitalism is that the term gets used to refer to very distinct things. “Capitalism” is sometimes used to refer to a specific historical period—the one in which we currently live—that followed the transcendence and abolition of feudalism. In this view we might say that we don't define capitalism, but rather the historical period of capitalism develops a logic that comes to define us, our practices, and our very way of understanding things.\(^\text{14}\) Related to this is a view of capitalism as an ideology, as opposed to a particular set of institutions and practices. We find this in people like Gramsci and Foucault who emphasize cultural and discursive elements of our world to explain why power is situated and distributed the way it is. In this tradition capitalism is primarily a mode of thinking that structures our actions and preferences for the benefit of particular empowered

interests. While not unrelated to these two ways of understanding capitalism, I am primarily concerned with capitalism as a complex of institutions. While any given institution might have existed at various times throughout history, their coincidence is particular to our historical period, and reflects particular normative and ideological commitments.

What institutions constitute a capitalist order? There is no off-the-shelf definition, but a look at various scholarly approaches to capitalism yields a workable consensus which, I think, will be sufficient for my purposes.\textsuperscript{15} This definition would basically include six institutional prerequisites for a capitalist order: 1) labor power is treated as a commodity that implies both that there is a market for labor power and that laborers work for wages (in contrast to peasant labor, or slave labor, where labor is not traded on a market, and where laborers are not remunerated with wages); 2) the idea of wages implies that the primary mode of exchange is money; 3) allocation of both productive resources and finished goods is done primarily through the competitive determination of prices; 4) productive goods (or the “means of production”) are owned privately (as opposed to collectively or nationally), and can therefore be exchanged and valuated using the price mechanism; 5) profits, or the surplus of productive enterprise, accrue to the owners of productive capital; and 6) property rights are protected and enforced by the state.

I take these criteria to be fairly intuitive and plausible on their face. It seems hard to imagine a capitalist order not bearing these six features, and there does not seem to be another feature which all forms of capitalism seem to necessarily have in common. Importantly, however, this definition still admits of variation. I have left out, for instance, the role of the welfare state in capitalism, as well as the role of regulation, or national protection. This is because I take seriously the political science insight

that there exists a variety of capitalisms, all of which combine the six above-mentioned institutions in varying ways, and some which contain other core influential institutions. This definition therefore admits of the various sizes of the welfare state, differences in tax regime, degree to which property rights are attenuated, relative influence of labor in fiscal and economic policy, and so forth, while not admitting of every kind of political economic order. Under this definition, for example, 17th century Ireland and present-day Cuba could not be considered capitalist societies. This definition therefore not only seems plausible but also seems to admit of the necessary variation while avoiding conceptual stretching.

What is the good promoted by capitalism?

To this normative question I answer simply that the good to which capitalism strives is the good of efficiency. That is, capitalism emphasizes the moral nature of satisfying individuals' preferences—maximizing their utility—by maximizing the productive use of resources. The implicit moral argument of capitalism is that because the human condition is one of scarcity, and because the ability to lead a good life (however defined) requires a material basis, a political and economic order that does not concern itself with the effective use of material resources will fail to allow people to pursue the good lives in the ways they wish. Put better, to be relatively indifferent to efficiency is to be relatively indifferent to human welfare; this is to be indifferent to the degree to which people can pursue the good life. This is something, I contend, that any critic of capitalism must take seriously, as notable 20th century socialists Oskar Lange and Abba Lerner implored in the 1930s.17


17 Oskar Lange claimed that socialists were indebted to the Austrian economist von Mises for forcing them to recognize the “importance of an adequate system of economic accounting to guide the allocation of resources.” For this reason, Lange added as a joke, “a statue of Professor Mises ought to occupy an honourable place in the great hall of the Ministry of Socialization or the Central Planning Board of the socialist state.” Oskar Lange. 1936. “On the Economic Theory of
Note that I said “satisfying individuals' preferences” and not satisfying more preferences generally. Against an aggregationist approach, I take the normative standard of efficiency to which capitalism aspires to be the Pareto standard, which says that a state of affairs can be ranked as preferable if, all things being equal, one person considers herself better off and nobody else considers themselves worse off than they did before.\(^\text{18}\) Where one person is made better off and another is made worse off, the Pareto principle of efficiency remains silent. What makes the Pareto principle significant is that it is generally satisfied by trade. If, without being coerced, two people exchange some good or service for another, each of them must prefer or at be least indifferent to the state of affairs existing after the exchange in comparison to the one before it. Therefore, there is good reason to think that every freely engaged exchange of goods entails an efficiency gain and results in an allocation superior to that existing prior to the exchange. Market exchanges, then, have a strong implicit argument from efficiency in their favour, granted the assumption that such exchanges are done without coercion.

Capitalism's privileging of efficiency, then, should be understood in terms of a respect for the distinctiveness of persons, and an affirmation of the basic good in improving people’s material welfare. This is in sharp contrast to the Foucaultian critique that capitalism’s emphasis on the efficient satisfaction of preferences constitutes a form of discipline.\(^\text{19}\) While there is an important empirical and social-theoretic insight in this idea, I do not think that this view undermines the value of efficiency. Insofar as we think that political society ought to protect and advance human welfare, we must be

---


\(^{19}\) Foucault argues that the neoliberal age brought about a view of *homo oeconomicus* wherein the subject is understood as an “entrepreneur of himself,” not simply a consumer but a producer of his own satisfaction. Feeding into a view of human capital, Foucault argues that this disciplines people to structure their lives in ways which fit into the logic of growth. See Michel Foucault. 2008. *The Birth of Biopolitics: Lectures at the College de France, 1978-1979.* Trans. Graham Burchell. New York: Palgrave Macmillan, 226. I discuss Foucault’s interventions into the theory of neoliberalism in chapter 3.
committed to efficiency as a social end. Obviously, this does not mean that efficiency is the only social commitment, or even the most important commitment. But it must be an essential one.

I want to draw attention here to another important contrast. I have argued that the normative defence of capitalism lies in its efficiency-advancing properties and not in its liberty-advancing properties. This isn’t to say that we ought not be concerned with liberty; as the first presupposition made clear, we should be and are concerned with individual liberty. However, I stated the preference for individual liberty as a separate presupposition precisely because I do not think that this liberty is the good of capitalism in two senses: first because a non-capitalist regime might place an equally high regard on individual liberty; and second because capitalism might not be terribly good at promoting this liberty, but still might be redeemable on welfarist grounds. This is precisely why the good of liberalism was stated separately, and prior to, that of capitalism; liberal institutions are necessary to protect these rights from many things, one of which just might be capitalist institutions.

This is in contrast with people like Nozick and sometimes Hayek who claim to support capitalism on grounds that such a system maximizes human liberty. This libertarian defense of capitalism²⁰ views the market as the way of promoting social cooperation and cohesion most in line with liberty. For Nozick this is embodied in his defining of justice narrowly in terms of just acquisition and just exchange and not in terms of how distributions match a preconceived notion of desert or obligation—what he refers to as patterned conceptions of justice. Capitalism is therefore preferred not for its welfare-related benefits but because it allows people to pursue the goods they want with minimal constraint. That is, by not restricting “capitalist acts between consenting adults,” the capitalist market best secures Nozick's maxim of justice: “from each as they choose, to each as they are chosen.”²¹ For

²⁰Here I take libertarianism to be a philosophical perspective that uses the maximization of negative liberty as the metric by which social arrangements and institutions are judged, and thus not simply a political project regarding the size of a welfare state.

Nozick liberty is always justice-preserving, and because capitalism enables the most liberty in economic affairs, it is the most just economic institution.

One difference between these two views is that they have direct political and practical consequences in terms of the type of capitalism endorsed. I began by noting that there are different varieties of capitalism. When presented with a question about the welfare state, for instance, the coercive taxation involved will be greeted with distrust if not straightforward hostility by a more libertarian-minded defender of capitalism, while potential efficiency-gains of universal social insurance schemes will be embraced by welfarist defenders of capitalism. However, more importantly for our purposes here, I also believe the welfarist account to be more coherent than the libertarian account. This is because the libertarian is not able to explain why we should have a coercive state enforcing and protecting the property rights necessary for a market-based society. Indeed, the libertarian must explain from where these property rights come in the first place; absent a religious or natural rights foundation (of which a modern libertarian ought to be weary), the question of how and why property rights arise looms large. The answer to this question, ultimately, is that private property rights are the most efficient way to protect productive endeavour and allow people to trade, buy, and sell. Once conceded, it is not obvious why we should cease to be concerned with efficiency after the setup of property rights. Indeed it seems that enabling the trade and sale of property would just be an extension of this. Therefore it is not obvious that a libertarian can get away from welfarist stances in order to justify the foundational element of their liberty-enhancing scheme.

How and Why does Capitalism promote this good?

The empirical question then is why we should believe that capitalism is particularly efficient and welfare-enhancing. Here I take Hayek’s criticism of planned economies and analysis of the role played by “knowledge” in political economy as the best account of the welfare-enhancing qualities of capitalism. For Hayek, the problem to be solved in political economy is the question of “how to secure
the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know....It is a problem of the utilization of knowledge not given to anyone in its totality.”

Capitalism’s virtue is the price mechanism, which conveys information about what consumers want and what producers have in more effective ways than other alternatives. Instead of requiring one or a group of individuals to do the Herculean task of figuring out how best to produce and for what ends, the prices produced by markets have the effect of conveying this information spontaneously:

It is more than a metaphor to describe the price system as a kind of machinery for registering change or a system of telecommunications...The marvel is that in a case like that of a scarcity of one raw material, without an order being issued, without more than perhaps a handful of people knowing the cause, tens of thousands of people whose identity could not be ascertained by months of investigation, are made to use the material or its products more sparingly; they move in the right direction.

Price information allows producers to ascertain the relative demand for a given product, and to adjust their production accordingly; similarly, prices signal to consumers the relative scarcity of a given good, which allows them to alter their consumption habits accordingly.

Because of this, more efficiency gains can be achieved given a certain set of resources and preferences in an economy coordinated by a price mechanism than by a planning committee. Effective transmission of accurate information in the form of prices allows for the greatest number of possible Pareto-improving transactions to occur. Pigou captured this insight in a slightly different but helpful way. For Pigou, prices convey information that allows the marginal personal interest of any particular economic action to equal marginal gain to social welfare. We can say that the cost of a good is the price I must pay to obtain it. The virtue of capitalism’s competitive price mechanism is that, by reflecting the relative scarcity of a particular good, the cost it imposes on the purchaser is equal to the marginal social cost she imposes on society by removing that good from the general stock; similarly the benefit one


23 Ibid, 527.
gets by selling a good is equivalent to the marginal social benefit she gives by adding that good to the general stock. For example, if I buy a steak from a grocery store, I will pay a certain price for it. This price reflects both the cost of producing the steak but also the relative scarcity of steak in relationship to demand. By paying a premium on the steak I am paying the cost of removing one steak from our social reserves of steak. All this is to say that prices communicate relative scarcity, allowing both consumers and producers to decide more effectively upon, and enact, their preferred economic behaviour considering the context.

This is a slightly different argument than what we might call motivation-based arguments for capitalism. By this, I mean the idea that capitalism is most efficient because it motivates people to work harder than they would in some other social system because capitalism alone ties direct material incentives to productive behavior. While the lineage for this line of thinking dates back at least as far as Bentham, it receives its most significant 20th century articulation in the work of Schumpeter. As will become clear throughout the dissertation, I contend that that account of capitalism is unpersuasive. I argue in chapter 4 that people are motivated by various concerns, material incentives being only one among them; furthermore, even if such incentives were the only source of human motivation, the material incentives proffered by capitalist markets sometimes create perverse situations where people do not work as productively as they might otherwise. This isn’t to say that incentives are not important. Indeed, the point made above about prices and the information they convey would become meaningless if we did not recognize that producers and consumers have budget constraints, constraints within which they need to make decisions that contribute to an efficient economy. The point I am making here is that price-related material incentives are important mainly for the manner in which they orient economic behavior, and less so for how they stimulate such behavior.

Now it should be acknowledged here that many socialist thinkers have conceded this point and acknowledge this informational strength of capitalism. It is for this reason that many on the left have attempted to create schemes whereby the equality of collective ownership is created but the efficiency of market pricing is preserved. Whatever the viability of the market-socialist project, the point to be gleaned here is that the virtue of capitalism lies in the information it allows to be conveyed and the types of activity it enables.

A Political Theory of the Corporation and Its Components

I have attempted to articulate the normative understandings of the liberal political ideas and the capitalist modes of allocation, production, and distribution in which we find ourselves. Put together we might state the basic ideas in this way: we must approach and treat each other as equal individuals and therefore must pursue institutions that protect this individuality and equality. Once these basic things are secured we should seek to create institutions that allow us to pursue our aims to the greatest degree possible. Liberal democratic institutions attempt to do the first, and capitalist institutions the second.

The Puzzle

Stated in this way we can see the types of trouble the corporation raises for us. On the one hand the institutional features of capitalism find their defense in their capacity to maximize efficiency, done primarily by virtue of the information-expressing properties of the price mechanism. Yet, the corporation—a key component of capitalist markets—does not use the price mechanism to organize itself internally. Instead we find planning, hierarchy, and bureaucracy. Either the corporation violates the efficiency mandate of capitalist institutions (in which case we should wonder why it exists) or it does not (in which case we should wonder whether the price mechanism is in fact the best way to attain

efficiency). Furthermore, within the corporation we find the rejection of individuality in favour of corporate enterprise, we find hierarchy instead of equality, and we find democratic procedure restricted to a very few members of the corporation, if it can be said to exist at all. The corporation, at first glance, seems to fly in the face of the normative underpinnings of both liberalism and capitalism.  

For Marx, of course, this was precisely the point. The “equality, freedom, property, and Bentham” of the market were, upon entering the confines of the private institution, left at the door. The inconsistency of principle and practice was not something to be resolved but to focus upon; the corporation was the institutional manifestation of capitalism's contradictions, and would prove to be the vehicle through which capitalism was transcended. This is a perspective I wish to avoid. In the next chapter I explain this argument in greater detail and show that, despite the considerable power of his narrative, we have several reasons to doubt Marx's account of capitalism and the corporation. Furthermore, I want to start with the assumption that the corporation can be seen as a coherent and consistent component of a liberal capitalist institutional order. Starting with that assumption, however, I argue that such a commitment demands a fairly radical reordering of our current institutional practice of corporate capitalism.

Instead of seeing the corporation as a contradiction of capitalism, I want to suggest we see it as a puzzle in need of theoretical solution. A political theory of the corporation, therefore, must attempt to resolve the corporation's apparent violation of both liberal and capitalist norms. It must offer a rational reconstruction of why a corporation exists in the market in the first place. That is, even if we can give a specific historical narrative of why and how a particular corporation came to exist in some particular market, we would want to understand why corporations in general have managed to become fixed parts

---

26 It was largely for this reason that Galbraith suggested capitalism and socialism would converge onto one economic system. That is, the idea of convergence is not based solely on the idea that socialist countries would learn the wonders of market economies, but that market economies tend to produce large swaths of socialist-like planning entities in the form of corporations. See, eg. John Kenneth Galbraith. 1973. *Economics and the Public Purpose*. Boston: Houghton Mifflin.
of markets in general. I call this the “productivity component” of the theory of the corporation. This account of how and why the corporation is an important productive institution in capitalism is not enough, however, since a productive institution that violates core liberal democratic commitments would need to be rejected. A political theory of the corporation must then offer some story about how the existence and organization of the corporation is consonant with the liberal democratic norms we endorse or how they could be made to be so. I call this the theory’s “normative component.” I go through these in more detail below.

The Productivity Component

Again, the puzzle addressed by this component is why the corporation should be part of a capitalist system which aims to maximize efficiency by using the price mechanism. In this sense, what is needed is an account of why corporations should be seen as contributing to efficiency, despite the lack of an internal price mechanism. Yet, in providing this account, one must also explain why the market and the price mechanism continue to exist, despite the productivity gains of corporation; that is, one must give an account not only of the efficiency strengths of corporations but also of their limitations. Finally, if one is going to claim that some things can be done by corporations while others are better done by markets, one needs some conceptual understanding of which relationships constitute a corporation and which do not. All together we can state the three questions needing answers in the productivity component of the theory in this way:

The Question of Efficiency: How and why is a corporation an efficient way of organizing?

The Question of Constraint: What are the limits and constraints of corporate organizing?

The Question of Constitution: Which people and what relationships fall within the domain of the corporate and which do not?

Satisfactory answers to these questions would give us an account of why corporations (should) exist in
an efficient market, why they cannot be the sole basis of economic production, and an analytic way of understanding which relationships would be deemed constitutive of the corporation and, therefore, what precisely would be subject to the political theory of the corporation’s normative component. It should be noted that the delineation of the corporation as being composed of some relationships and not others does not necessarily entail their normative primacy. We might say that a corporation is composed of particular people but that the control of the corporation ought to be society’s and not those constituting the corporation; or, we might say that despite the many relationships which constitute the corporation, certain particular relationships amongst them deserve special treatment. The point is that in order to understand the productivity of the corporation, we must have some understanding of what the corporation is, where its boundaries lie, and how it is distinct from the market and other non-corporate institutions.

The Normative Component

With an understanding of why the corporation is seen as a part of an efficient capitalist order, we are left with the puzzle of whether such an institution is, or can be made to be, compatible with liberal democratic norms. As John Rawls eloquently put it, since justice is the first virtue of social institutions, “laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust.”27 The fact that we can understand corporations as contributing to efficiency does not address the apparent tension corporations present to liberal democratic norms. The three norms that I focused on in particular were individuality, equality, and democratic decision-making, all of which were presented as reinforcing the others.

In the interests of interdisciplinarity and avoiding wheel-reinvention, I connect these above concerns with three established fields of normative enquiry related to the organization and behaviour of

27 Rawls, Theory of Justice, 3
corporations: corporate law, corporate governance, and business ethics. Just as the conceptual component of the theory needed to answer three questions, here too we see that the normative component must account for three answers to interrelated questions. They are as follows.

**The Question of Corporate Law:** What ought to be the legal limits of corporate authority with respect to internal affairs or, put differently, what are the grounds for contesting authority and calling it to account?

**The Question of Corporate Governance:** Who ought to make decisions in the corporation, on whose behalf, and why?

**The Question of Business Ethics:** Upon what normative considerations, and according to which criteria, ought corporate decision-makers base their decisions or, how ought we appraise the executive decision-making of corporations?

An answer to the corporate governance question gives an account of the types of hierarchy that ought to characterize the organization and governance of a corporation, and in so doing, explains who ought to be viewed as the “constituency group” of a corporation and why. An answer to the corporate law question explains how such a hierarchy can be squared with the legal protection of equality and individual liberty. Finally, an answer to the business ethics question gives an account of how this hierarchy can be squared with the extra-legal, moral concerns for liberal individuality and equality. Furthermore it might give an account of how such decision-making could be seen as advancing, and not merely satisfying, the normative concerns of liberal democracy.

**Conclusion**

In this chapter I have explained the basic normative presuppositions that ground my consideration of the corporation. These normative presuppositions I take as roughly resembling the background normative assumptions of our current liberal democratic institutions. These normative assumptions are that the liberal democratic norms of individuality, equality, and democracy are preferable to other norms, and that capitalism is a desirable set of institutions, due to their ability to
maximize efficiency. I then contended that the corporation poses a *prima facie* problem for both of these sets of norms. In order to resolve this apparent problem we need a political theory of the corporation which addresses six basic questions: the question of efficiency, the question of constraint, and the question of constitution, which together constitute the productivity component of the theory; and the questions of corporate governance, corporate law, and business ethics, which comprise the normative component of the theory.

I use these questions in two ways. First, I use them as tools for analyzing extant scholarship on the corporation, articulating what I take to be the Chicago School’s implicit answers to these questions—its political theory of the corporation. Second, I attempt to answer these questions as a means of proposing my own political theory of the corporation. Put differently, I use these questions first as a means of framing my exposition and critique of the Chicago School’s theory of the corporation. I then use that frame to reconstruct a theory of the corporation that I think is more in line with the values we hold as a society.
Chapter 2. The Productivity Component of Transaction Cost Economics

Perhaps the most difficult thing involved in thinking theoretically about contemporary corporations is how used to them we are. The existence of corporations in markets and capitalist economies is so commonplace that one is usually seen in tandem with the other: private corporations are the products of a liberal capitalist society, and a modern capitalist society will produce, and be dominated by, large corporations seeking to expand and maximize profits. This is not only true for the casual observer. Put in its simplest form, microeconomics is constituted by three main theories: a theory of consumer choice and rationality (how individuals desire, act, and respond to market signals), a theory of the firm (how firms produce and respond to demand), and a theory of market equilibrium (how the interaction of consumer and firm creates an efficient allocation of goods and services). So, from the standard model’s perspective, there are individuals who interact with firms in equilibrating markets. The firm and the individual simply exist.

In the last chapter I suggested this common view is based on a misconception. In fact, the corporation is something of a puzzle when considered from the standard understanding of market efficiency. To resolve this puzzle, I argued, we need the first component of a political theory of the corporation, what I referred to as the “productivity component,” which involved answering three questions: the question of efficiency; the question of constraint; and the question of constitution. In this chapter I offer an account of this component as is found in the transaction cost tradition of economic thought. While given its most influential articulation by Oliver Williamson, this approach is built out of Ronald Coase's famous insight that transaction costs are rampant and create inefficiencies in a market economy. As we will see, Coase gave the first marginalist account of precisely why an integrated firm

would be created in a market organized by a price mechanism and, in so doing, provided the basis for the productivity component of the corporation.

The Coasian legacy in this tradition is worth emphasizing. In the next chapter I will show that Coase has been interpreted in different ways and for radically different purposes, largely because he is famous for two separate and very different arguments. The subject of this chapter is the account of the firm emanating from Coase’s first major influential article, “The Nature of the Firm.” In the next chapter I argue that Coase’s second and most famous piece, “The Problem of Social Cost,” has been used to reinterpret Coase’s writings on the firm and, in the process, to offer a particular conception of how the corporation ought to be governed. This school of thought (associated with the Chicago Law and Economics tradition) has developed a full-blown political theory of the corporation, replete with both a productivity component and a normative component, which I articulate at the end of chapter 3. In chapter 4, I argue that this theory is wrongheaded both as a matter of interpretation and from a normative perspective; by emphasizing “The Problem of Social Cost,” this tradition has not only added a normative component to the theory of the corporation that is neither desirable or viable, but has also altered the Coasian formulation of the productivity component so as to render it conceptually muddled.

Because of the stakes involved in the history of interpreting Coase, this chapter proceeds in a fairly idiosyncratic manner. While the main goal of this chapter is to articulate the transaction cost account of corporate productivity, I do this by presenting its development in the context of the history of economic thought, placing Coase at the center of this development. To show the revolutionary nature of his insight, I begin this chapter with an overview of how the modern corporation’s direct antecedent, the joint-stock company, was portrayed and viewed by the classical political economists. I show that the economic thinking about the joint-stock company was very diverse, ranging from ambivalence to excitement, but that it was also grounded in the common ontological and normative assumptions of the times. The marginalist revolution of the late 19th century ushered in a new mode of economic analysis, challenging much classical wisdom, including the theorizing on the joint-stock
company that was now left undertheorized and unexplained. In the second section I show how Coase's idea of transaction cost filled this lacuna, giving an account of the integrated firm consistent with marginalist principles while simultaneously challenging some of the key assumptions underpinning this new paradigm. In the third section I show how the development of transaction costs by Oliver Williamson served to address the issues not addressed by Coase, and how these—taken together—provide a fairly powerful answer to the three questions that comprise the productivity component of the political theory of the corporation.

The Joint-Stock Company and Classical Political Economy

An incipient form of the modern-day business corporation, the joint stock company was never a primary topic for the classical political economists of the 18th and 19th century. This silence is in many ways owing to the smaller role such institutions played in the economies of their day, particularly following the legal restrictions placed on the joint-stock company in the 1720 Bubble Act. Still, many of the questions and problems that the 20th century theory of the firm addressed can be found in the writings of Smith, Mill, and Marx. Indeed, despite the relatively low volume of writing on the subject, we see that both Mill and Marx accord great theoretical importance to the joint-stock company; Smith’s relative quiet, furthermore, is instructive in pointing out how his assumptions are connected to the bounds of his thought.

Adam Smith

The classical political economists all attended to two key features of the joint-stock company in their analyses: the nature of team production amongst laborers, and the use of stocks and shares to finance the company. Smith located the advantage of team production in the division of labor which it

---

enabled and upon which it depended. Through his famous intricate discussion of the various activities that occur within a pin factory, Smith contended that such enterprises proved more profitable and efficient because they allowed for separation and specialization of activity. While the division of labor that Smith celebrated increases productivity, it creates the problem of laborers who are only trained and equipped to supply a small number of the things life requires. A regress problem arises: a man is “supplied by the produce of other men’s labor, which he purchases with the produce of his own…but this purchase cannot be made till such time as the produce of his own labor has not only been completed, but sold.” As we move further from what Smith describes as “the ruder state of things,” we require a greater division of labor that, in turn, requires “a greater stock of materials and tools than what would have been necessary in a ruder state of things,” necessitating more capital, and so on.

What is needed is a large initial store of capital to jumpstart the process.

The legal form of the joint-stock company enables precisely this accumulation for Smith. Because shareholders were already legally empowered to transfer their shares without the approval of the company or other shareholders, and were sometimes even granted limited liability with regards to debts the company takes on, the joint-stock company is able to attract more capital than a private partnership. Furthermore, because the directors of the company have a direct interest in ensuring that the enterprise is run efficiently, the joint-stock company tends to work more effectively than its government-regulated counterparts. Seeing as these are features common to our present-day corporation, their justifications are familiar to us: by allowing individuals to buy and sell shares with ease, and by limiting potential shareholders’ losses to the extent of their shares and no further, the joint

---

5 As I have already noted in the introduction, limited liability did not become a general feature of the business corporation until the latter half of the 19th century. The institution which Smith encountered was still rather inchoate in this regard, with some charters including limited liability. See Gary M. Anderson and Robert D. Tollison. 1982. “Adam Smith’s Analysis of Joint-Stock Companies.” Journal of Political Economy 90(6): 1240f.
6 Smith, 699-700.
stock company was an innovative way of incentivizing the investment of the wealthy into the enterprises of the industrious. However, this raises a problem: by expanding the scope and industry of the initial enterprise, and by increasing the number of shareholder-owners beyond those intimately connected to the endeavor, the joint-stock company creates a monitoring problem. How do the various and disconnected owners oversee the effective use of their investment? The result would be that the managers, the “servants and dependents” of the investors, will tend to waste and embezzle the surplus after dividends are paid out to stockholders, if dividends are even paid out in the first place.7

Smith’s insights here foreshadow the principal-agent problem that would characterize much of the 20th century debate around the firm (as well as the problem of representation more broadly).8 For Smith these problems are fatal to the joint-stock company, which is why he thought so many had to be insulated from competition through monopoly privileges. The only way around these principal-agent problems for Smith was to restrict the form of the joint-stock company to enterprises where the operations are “capable of being reduced to what is called a routine, or to such a uniformity of method as admits of little or no variation.”9 Smith lists four such industries—banking, insurance, canal-digging, and water-supply—and contends that only these ought to be organized as joint-stock companies.10

From this relatively short discussion, we can see that Smith implicitly provides answers to the questions of efficiency and constraint posed earlier. How and why is a corporation an efficient way of

7 Ibid, 711.
8 For the principal-agent approach to the firm see Eugene Fama and Michael Jensen. 1983. “Separation of Ownership and Control.” Journal of Law and Economics 26:301. This is discussed in greater detail in the next chapter. For representation more broadly see the classic Hanna Pitkin. 1967. The Concept of Representation. Berkeley: University of California.
9 Smith, Wealth of Nations, 713-714.
10 It is not insignificant that these industries are domestic ones since in many ways Smith’s pessimistic view of the company fits his overall project of dismantling the mercantilist economics of his day. The majority of the companies with which he is concerned are trading companies which reflected the priority placed on balance of trade and extraction of wealth from foreign countries. This emphasis distracted from economic policy-makers from recognizing that it was labor—and more significantly for him, the division of labor—which was the source of wealth.
organizing production? Smith answers that the division of labor within manufactures allows for increases in efficiency, and that the ability to buy and sell shares with limited liability provides the necessary capital to allow for this division of labor to occur. What are the limits and constraints of corporate organization? Smith’s answer is strong, contending that the principal-agent problems will be so vast as to constrain the corporation to only those industries which can be routinized, or those given the extra advantage of monopoly. Significantly, this account of corporate efficiency and inefficiency does not rely on technological variables but on the relationships between laborers, managers, and capital providers.\(^\text{11}\) The particular advantages and disadvantages that come with the joint stock company are to be found in the institutions they create in which different classes are brought together.

Yet despite these important and groundbreaking insights Smith does not have a clear or obvious answer to the question of which relationships constitute the corporation. This lack of clarity is because Smith does not recognize a tension in how he conceives of the division of labor within a firm and that division’s relationship with attracting capital through selling stock in market exchanges. Smith famously contends that the division of labor is an outcome of man harnessing his basic interest in trucking and bartering, meaning that the extent of the division of labor is directly related to the extent of the market: “it is the power of exchanging that gives occasion to the division of labor, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market.”\(^\text{12}\) There is something intuitive about this idea: labor can be divided and specialized only if we can trade with others to get the foregone fruits of other labor; therefore, the more opportunities we

---

\(^{11}\) This appears to be part of the reason why Smith chose a pin factory, a “trifling manufacture,” to exemplify his division of labor; the pin factory had neither the advanced technology of nearby iron works factories, nor the larger numbers of the “great manufactures.” (Paul Williams. 1978. *Emergence of the Theory of the Firm from Adam Smith to Alfred Marshall*. London: Macmillan Publishing. 16) In so doing, Smith controls for two key variables—technology and size—in order to account for the productiveness of corporate institutions. While this neglect might be part of the reason why he downplayed the heightened possibility of technological innovation and economies of scale resulting from innovated enterprise, Smith’s pin factory also pointed out that the division of labor efficiencies and principal-agent inefficiencies were endemic to such institutions, regardless of size and technology.

\(^{12}\) Smith, *Wealth of Nations*, Ch. III.
have to trade the greater the division of labor can be.

And yet the pin factory betrays a more subtle and nuanced view. The division of labor within the collective enterprise is marked precisely by the absence of the market. This is in contrast with the account of why the joint-stock form enables greater capital—that is, it incentivizes investment by providing the ability for investors to make profit off of commodities being sold in a market. Smith’s pessimistic view of the firm in part results from his inability to recognize a tension between the two different ways of coordinating activity he is discussing: the competitive markets in which commodities and capital are traded on the one hand, and non-market cooperation that exists within manufactures, and enterprises more generally, on the other. Unable to explain the advantages of the latter, the principal-agent problems produced by dispersed stock-ownership appear more formidable than they might otherwise be.

This inability to understand the relationships that constitute the corporation points to the faults in Smith’s account of corporate efficiency and constraint. The division of labor enabled by capital accumulation cannot be the only reason the corporation is efficient since markets are also able to produce divisions of labor. The question that needs to be asked is what the corporation is doing that necessitates the absence of market mechanisms, specifically the absence of competitive pricing internal to the firm. Similarly, while Smith was certainly correct about principal-agent relationships, his inability to grasp the full range of efficiencies achieved by the corporation led him to grossly overestimate the constraints on corporate organizing, and thus left him unable to explain the flourishing of corporations beyond monopolies and those four industries that he enumerated.

*John Stuart Mill*

Mill’s writing on the corporation is found in his *Principles of Political Economy*, a work seeking to clarify and update Smith’s system of political economy. Mill makes this clear in his preliminary remarks where he also establishes a key difference in approach between his and Smith’s
political economy. For Mill, Smith’s division of labor is actually subsumed under a larger understanding of what constitutes production and, in turn, productive labor. This understanding is technical, not social, in nature:

[Production] has its necessary conditions. Of these, some are physical, depending on the properties of matter, and on the amount of knowledge of those properties possessed at the particular place and time….Combining with these facts of outward nature other truths relating to human nature, it attempts to trace the secondary or derivative laws, by which the production of wealth is determined.13

Unlike the laws of distribution, the manner in which riches and useful commodities are produced from raw materials is given by the natural properties of the raw materials and “truths relating to human nature.” These, plus our bounded understanding of them, are the limits of productive capacity.

To this end, the division of labor is presented as a subcategory of cooperation; whereas “complex cooperation” reflects the division of labor, the “simple cooperation” of many people doing the same task at the same time represents just as large an advance in productivity and social relations for Mill.14 This acknowledgement of simple cooperation allows him to understand that costs can be reduced by virtue of the size or scale of an enterprise, that is, the principle of economies of scale: “even when no additional subdivision of the work would follow an enlargement of the operations, there will be good economy in enlarging them to the point at which every person to whom it is convenient to assign a special occupation will have full employment in that occupation.”15 Joint-stock companies are important for precisely this reason. In order to fully capitalize (pun intended) on the economies of scale generated by employing as much labor as possible, one needs to acquire as much capital as possible.

Here Mill draws a marked contrast between his view of the firm and Smith’s. Where Smith saw the principal-agent problems as being endemic and fatal to the joint-stock company, Mill attempts a number of arguments to salvage it. First, as has already been pointed out, he believed that the

---


14 Mill makes it clear that this is a significant advance in a rather unfortunate manner, by noting that “the savages of New Holland” do not engage in this simple form of cooperation (116-117). Smith distinguished advanced and “rude society” primarily by reference to the division of labor.

15 Ibid, 132.
economies of scale outweighed these problems. Secondly, he notes that the joint-stock form introduces an important aspect of publicity into matters that would otherwise be private. He specifically cites the problems inherent with industries that depend on confidence more than capital—specifically banking and insurance—which incentivize secrecy with respect to internal issues.\footnote{This note on publicity was added in the 1865 edition of the \textit{Principles}, perhaps following the conversion of a number of private banks into joint-stock companies.}

But, more importantly, Mill believed that the issues associated with monitoring could be mitigated by appealing to the self-interest of managers, fixing their remuneration and reputation to the performance of the enterprise so as to de-incentivize managerial malfeasance. Reputation would be handled by the aforementioned public nature of the firm’s internal affairs, and salaries could be arranged to track the manager’s performance by “giving them part of their remuneration in the form of a percentage on the profits,” thus uniting the manager’s interests with that of the shareholders.\footnote{Ibid, 140-141.} Mill believed these principles would work in tandem to attract superior talent to managerial positions; because of the chances to earn great profit and gain public esteem, shareholders would be able to hire men of “superior intelligence” and “habitual rectitude” to direct their operations, thus further reducing the principle-agent problems. By increasing the “zeal” and “fidelity” with which managers do their jobs, these methods salvage the joint-stock company from Smith’s critiques and explain their ability to survive market competition in the decades following the \textit{Wealth of Nations}.

Yet Mill takes this idea of remuneration and publicity to develop a striking hypothesis. If zeal and fidelity determine the quality of the manager’s labor, and these can be cultivated through particular remuneration schemes and publicity, why not extend this to all labor? Why not extend profit sharing to the workers of a firm in the form of shareholdings and ownership? Citing prior examples of profit-
sharing schemes initiated by management which promoted “zeal” in workers\textsuperscript{18}, Mill explains that the development of limited liability with joint-stock companies has actually enabled the generalization of this practice by allowing workers to share in profits without being fully liable for losses. By introducing principles of collective ownership and limited liability, and by perfecting techniques of cooperative team production and management, the joint-stock company sets the stage for cooperative worker ownership to prove itself as a more efficient (and evidently more just) system of production.

As shareholders realize they can get more out of their managers through the remuneration schemes, more and more will tend to do so; as they realize they can do this with labor writ large, they will also tend to do so; as workers realize they can own their own enterprises on the same principles, they will gradually seek to buy the stocks and set up their own enterprises on the co-operative model; and finally, in proving themselves more efficient, these organizations will come to dominate modern industrial economies, attracting capital from investors.\textsuperscript{19} Such is the logic of Mill’s hypothesis. It must be emphasized, however, that though Mill does think this arrangement would be good, this is not exactly a prescriptive claim. Following Smith, Mill contends that the ultimate test for organizations is their ability to survive competition.\textsuperscript{20} This is why it is best understood as a hypothesis: if cooperatives succeed, then all the better for the moral health of the working class; if not, then it was not a viable model in the first place. There will more to say later about why Mill’s hypothesis has not been confirmed by historical experience. For now we can simply see the developments that Mill made on the structure built by Smith.

In contrast to Smith, we find in Mill a much fuller account of the corporation’s productivity and its distinctiveness vis-à-vis the market. For Mill the question of efficiency is answered not only with

\textsuperscript{18} Ibid, 768-769.
\textsuperscript{19} Ibid, 791.
\textsuperscript{20} Ibid, 792-793.
Smith’s division of labor, but with an account of economies of scale as well. Furthermore, Mill suggests that remuneration schemes might be a way to overcome principal-agent problems, prefiguring the work of agency theorists like Michael Jensen by over 100 years. Mill’s recognition that such schemes don’t merely play on material incentives, but affect the motivation of workers by reference to things like “reputation,” “zeal,” and “fidelity,” is also an important departure from Smith. While still appealing to self-interest, Mill believed that incentives would not just prevent fraud, but would actually increase productivity and innovation. While his belief in these psychological factors has proven to be naïve (if not downright wrong) his recognition that institutional and compensatory factors affect the intensity and quality of work (as opposed to just the quantity of work or the propensity for malfeasance) is important for understanding the multidimensional effects the corporate institution can have. Finally, and most significantly, Mill recognized that the manager of a corporation, while being a representative of capital, is still essentially a laborer, whose work can be affected in the same way as other laborers; this observation is what allows Mill to hypothesize that the capital-owned firm might be transcended someday.

This more explicit emphasis on the different types and combinations of labor leads Mill to answer the question of constitution better than Smith had: for Mill the relevant relationships within the corporation were those cooperative relationships amongst laborers (including managers) that allowed for their greater degree of productivity. However, Mill’s faith in the productive combination of labor leaves him without an effective answer to the question of constraint. While Mill argues that the success of particular corporate structures will be constrained by market competition, he does not explain why the cooperation that takes place within the corporation cannot be an effective substitute for market exchange on the whole. Given his rosy picture of what remuneration and publicity can do for the

21 In chapter 4 I will argue that this is actually a crucial aspect of corporate productivity, which Chicago economists do not attend to.
productivity of labor, it’s not evident why we need the market at all.

*Karl Marx*

Mill’s inability to answer the question of why the market is needed when the combination of labor is so productive was taken by Marx to be the refutation of the question’s underlying premise. Whereas Mill hypothesized the coming dominance of worker-ownership through the corporate form, Marx raised the bet and claimed that the corporation would not merely lead to the dominance of worker-owned firms in competitive markets, but would lead to the overcoming of markets and capitalism as we know it. For Marx, production isn't the technical process Mill claimed it was, but a specific historical mode of arranging social relations. Capitalism is beset by fundamental contradictions, the most important of which is the great discrepancy between the principles that govern exchange and circulation and those that govern the process of production. As was noted in the prior chapter, Marx described the domain of circulation and exchange as the “very Eden of the innate rights of man,” the realm of “freedom, equality, property, and Bentham,” where free individuals trade goods and services on an open market as legal equals. Setting aside the sarcastic tone, Marx was intending to draw an important contrast between the domain of exchange and that of production:

> When we leave this sphere of simple circulation or the exchange of commodities, which provides the “free-trader *vulgaris*” with his views, his concepts and the standard by which he judges the society of capital and wage-labor, a certain change takes place in the physiognomy of our *dramatis personae*. He who was previously the money-owner now strides out in front as a capitalist; the possessor of labor-power follows as his worker. The one smirks self-importantly and is intent on business; the other is timid and holds back, like someone who has brought his own hide to market and now has nothing else to expect—but a tanning.22

Cooperation, including the division of labor, develops as a way for capitalists to extract more labor-power from the laborers, thus increasing the returns on their initial money investment. This is a development of historical importance for Marx, as it creates labor as a social category, thus constituting the working class and disciplining them to be able to work in combination. What Marx noted explicitly,

---

and that Smith missed completely and Mill only recognized tangentially, was that the dynamics within
the enterprise were diametrically opposed to those of the market.

Less well-known than his account of the socialization and exploitation of labor, however, is
Marx’s account of the opposite process: the socialization and exploitation of capital, carried out by the
joint-stock company. Whereas capital appears on the historical stage after the socialization of labor and
concentration of the means of production, the joint-stock company creates “social capital,” the capital
of directly associated individuals, representing “the abolition of capital as private property,” as it is
transformed into collectively-held property.23 This process also ends up splitting the personification of
capital into two distinct characters: the manager, who has the duty of functioning as a capitalist by
overseeing the exploitation of labor, and the mere money owner who only collects profits in the form of
interest. The exploitation of labor thus falls completely onto the shoulders of a particular class of
laborer, namely the manager, while the capitalist becomes utterly estranged from his capital. The result
of this “capitalist production in its highest development” is the conversion of capital “back into the
property of the producers, though no longer as private property...but rather as their property as
associated producers, as directly social property...This is the abolition of the capitalist mode of
production within the capitalist mode of production itself. It reproduces a new financial aristocracy, a
new kind of parasite in the guise of company promoters... an entire system of swindling and
cheating.”24 For Marx the chances for managerial malfeasance that so greatly concerned Smith and
Mill were precisely the point of the joint-stock company's historical development. By introducing the
ability and possibility of supervision without the capitalist, the capitalist is in fact rendered superfluous;
to quote another of Marx’s turns of phrase, “what the bourgeoisie therefore produces, above all, are its

“social capital” here is quite distinct from how we use it. For a genealogical account which traces the trajectory of this term
24 Marx, Capital III, 568-569. Emphasis added.
own grave-diggers.”25 The cooperative factories prove this by hiring managers as wage-earners for the purposes of coordination and doing without the capital class entirely.26

Marx answers the question of efficiency in much the same way as Mill. However, whereas Mill emphasized the relationships between laborers as those relevant for the constitution of the corporation, Marx’s explicit recognition that productive relationships were contrary to those characterizing the market allowed him to cast a wider and categorically distinct net: it was not merely the cooperative relationships of labor, but the socialized and team-productive nature of both labor and capital that constituted the corporation, in contrast to those competitive and atomistic relationships that characterize the market. Similarly, whereas Mill did not provide an effective answer to the question of corporate constraint, Marx’s response was to turn the question on its head. There was no inherent constraint on the corporate form; instead, the fiction of liberal individual freedom presented by the market was constrained by the reality of hierarchy and sociality that could be found in joint-stock companies. Once this power was fully harnessed, the bourgeois mode of production, including the market mode of exchange, would be transcended.

Differences and Similarities

The classical political economists’ engagement with the corporation and their various formulations of its productivity component are summarized in table 2.1 below. We might note that the differences in their views of the corporation are, crudely, correlated with their views of the market; the more skeptical of markets or the more worried about the deleterious effects of industrial capitalism, the more positive their account of the corporation. This is, in part, because of what was noted in the previous chapter: the corporation represents, prima facie, precisely the opposite values and

26 Marx, Capital III, 511.
relationships to those embodied in laissez-faire capitalism. Smith’s skepticism of the corporation was
due precisely to its mercantilist and monarchist roots, whereas Mill’s (attenuated) and Marx’s
(unabashed) enthusiasm was rooted precisely in their recognition that it provided a forum for
empowering those marginalized by the logic of capitalism.

Table 2.1 The Productivity Components of Smith, Mill, and Marx

<table>
<thead>
<tr>
<th></th>
<th>Question of Efficiency</th>
<th>Question of Constraint</th>
<th>Question of Constitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith</td>
<td>Division of labor</td>
<td>Principal-Agent problems and general malfeasance</td>
<td>No Answer</td>
</tr>
<tr>
<td>Mill</td>
<td>Cooperative labor; quality of labor induced by corporate remuneration</td>
<td>No Answer</td>
<td>Relationships between various kinds of labor</td>
</tr>
<tr>
<td>Marx</td>
<td>Creation of surplus value through exploitation of labor</td>
<td>Corporation represents the overcoming of markets: it has no ahistorical constraints</td>
<td>Social labor and the social capital characterized by hierarchy and domination.</td>
</tr>
</tbody>
</table>

Despite these differences, two similarities amongst these accounts of the corporation should be
noted, as they help set the stage for the developments of the 20th century. The first is that the
corporation is interesting to all precisely because it is an institution where different social classes
interact together. In many ways classical political economy is marked precisely by this emphasis on
class,27 beginning with Smith’s observation that there were three distinct classes who lived respectively
off of rent, wages, and profit. The corporation is interesting and puzzling to all of them because we
simultaneously see the class living off of profit and the class living off of wages working together in
ways that are hard to observe. For Marx and Mill this intermingling was embodied in the manager who,
despite being a creature of capital, was fundamentally a laborer. The second similarity we see amongst
all these economists is a shared view of value as being derived ultimately from labor. This explains
why the profit-earning class must be in contact with the wage-earning class, since it is the wage-earners
who are the basis for value creation. In this sense there is more in common between Smith and Marx
than is often thought. As Meek notes:

Smith provided Marx (and of course Ricardo) with a model of the new tripartite framework of class relationships characteristic of capitalist society; he formulated a new concept of surplus in which profit was emancipated from its former dependence upon rent and ascribed to the productivity of labor in general; and he outlined a new theory of the development of society and the nature of socio-historical processes in general which, whether Marx himself was aware of it or not, set the stage for the eventual emergence of the materialist conception of history….The distinguishing feature of this methodology was that it stressed the determined role of the techniques and relationships of production—in much the same kind of way as the more developed methodologies of Ricardo and Marx were later to do.28

Marx’s contentions about the contradictions of capital, and his more specific claims about how the corporation would bring these to fruition, might in this regard be read as the reductio of Smith’s system; Marx is a critic of the tradition initiated by Smith, but the strength of Marx’s criticisms are a function of the degree to which we accept Smith’s premises.

Coase’s “The Nature of the Firm”

It is important to emphasize the elegant nature of Marx’s critique, and how it fits within the classical paradigm. If we assume both that labor is the source of value and that the economic system is composed fundamentally of different classes, then the problem of exploitation becomes structural and systemic: the only way for capitalists to increase the value of their investments and earn profit is to systematically extract more use-value from laborers than they are paying for. Furthermore, in order to continue the process of growth, this extraction will eventually involve gaining more capital, requiring an institution like the joint-stock company.

But what happens to this understanding of the corporation if the classical paradigm is no longer viable? This is how, I propose, we can understand Ronald Coase’s contribution. Coase offered an idea of the firm built on different foundations from Smith’s, thus defusing Marx’s critical account. Yet he also adapted Marx’s key insight—that the logic of production was distinct from the logic of

28Meek, 16. It would be more precise to say that Marx was building off of Ricardo than of Smith. This is partially because while Smith does have a dimension of class analysis, he also provides the precursor to modern equilibrium theory in his “invisible hand.” Insofar as Marx built on the first and critiqued the latter he was building off of the system which Ricardo built off the former.
exchange—to the new paradigm, thereby also correcting the blind spots of Smith and Mill.

_The First Paragraphs: Beyond Classical and Marginal_

Ronald Coase’s seminal 1937 essay “The Nature of the Firm,” begins with the stated ambition to develop a theory of the firm “tractable by two of the most powerful instruments of economic analysis…the idea of the margin and that of substitution.”29 The “margin” and “substitution” are the legacy of Marshall’s synthesis and professional institutionalization of what is known as the marginal revolution,30 the sudden development and acceptance of principles developed by Jevons, Menger, Walras, and others. This intellectual development stood the classical tradition of political economy on its head. As Blaug notes, for the classical economists “the function of economic analysis was to reveal the effects of changes in the quantity and quality of the labor force on the rate of growth of aggregate output.” However, after 1870 “economists typically posited some given supply of productive factors, determined independently by elements outside the purview of analysis. The essence of the economic problem was to search for the conditions under which given productive services were allocated with optimal results among competing uses, optimal in the sense of maximizing consumers’ satisfactions.”31

As the core economic problem ceased to be the way in which labor affects output, and instead became a question of why certain goods and services are allocated in the way they are, the classical approaches to the corporation with their emphases on class and the organization of labor, were rendered no longer viable.

To understand the nature of Coase’s contribution it is worth elaborating some of the main features of the new paradigm. Lionel Robbins noted that the marginal paradigm differed from the

classical paradigm in three fundamental ways: first, the primary economic subject was taken to be the individual and her subjective desires (in place of the objective existence of value); second, the primary economic condition was taken to be the scarcity of goods (as opposed to the system of production); and third, as a result, a distinct understanding of economic value was offered:

According to the classical economists, granted that commodities possessed utility, their value was determined by their cost, that is to say by the value of the factors of production which went to make them. According to the subjective theory, given the quantities of the factors and their technical substitutability, their values (and hence the cost of production of particular commodities) must be regarded as determined by the preferences of the economic subjects….Goods would have value if there were no such thing as production, provided only they were scarce. The process of production has significance only in so far as it modifies the possibilities of choice.  

In addition to these three changes offered by Robbins, we might add a fourth: in place of the classical idea of economic development, the marginal revolution offered the Walrasian idea of a static economic equilibrium, a set of prices which, by making equal the marginal utility gained on both sides of an exchange, renders demand equal to supply.

This dealt a serious blow not only to the classical tradition of economic theorizing, but to the Marxian critique of that tradition. If value is produced by the interaction of subjective evaluation and material scarcity rather than by labor power, then an account of enterprise that sets its sights solely on exploitation does not appear quite as viable. Profit can be created by hiring relatively less scarce labor, buying relatively less scarce raw and capital goods, in order to produce a good that is marginally more scarce and therefore worth more. While the existence or prevalence of exploitation is not totally denied under marginalist premises, it ceases to be a structural requisite of capitalism. Therefore, in order to render the firm compatible with marginalist principles, one must go beyond a structural account of how

33 Blaug, Economic Theory in Retrospect, 178, 549
34 For instance, a marginalist account of exploitation focuses on the range of possible wages determined by the bargaining context and the elasticity of the goods being produced; exploitation occurs when these factors yield wages less than the value of their marginal net product. See Paul Plateau. 2001. “Some Reflections on the ‘Pigou-Robinson’ Theory of Exploitation.” History of Economics Review 33:1-16.
classes interact with each other within the firm, and instead offer a theory that explains why certain goods and services are better allocated within such an institution.

Yet being “tractable” by marginalist principles was not the only concern. Following the methodological precepts of Joan Robinson, Coase claimed that we must go beyond blackboard economics; our theoretical assumptions should not only be sound theoretically, but should correspond with the real world understandings and usages of these terms. This is not an insignificant desideratum; while the marginalist revolution provided a fresh conceptual framework that eschewed the faulty assumptions of the classical paradigm, it also left the conceptual understanding of the firm quite muddled. Marx’s important economic insight—that the relationships and norms governing production directly contradict those of the market—still remained. In fact it was amplified by marginal precepts: the advent of equilibrium theory had revived faith in Adam Smith’s invisible hand, making the absence of market mechanisms within the firm even more puzzling. In this regard, Coase is positioning his argument in contrast to the tradition of Marshall and, most importantly, to the heir to Marshall’s professorship, Pigou, in whose work the firm is posited simply as a production function. Put differently, the neo-classical synthesis of Marshall and Pigou happily accepted the firm as a black box, which simply produced goods according to certain technological variables and responded to market signals in predictable ways. Coase wants to claim that this understanding does not correspond to real-world experience of what the firm actually is, namely a grouping of actors and transactions. As will become clear, bringing the theoretical understanding of the firm in line with its real world meaning actually involves disrupting a number of assumptions underlying the tradition of welfare economics which comes from Pigou.

Coase therefore establishes two main goals for understanding the firm. On the one hand one wants to explain the existence and pervasiveness of the firm’s internal structure without falling back on the classical/Marxian idea that production was utterly distinct from other economic phenomena. On the other hand, one must also avoid assuming the problem away by sweeping it into the black box of a
production function. Coase’s attempt to do precisely this is implicit in his description of his project as providing a realistic account of the firm which is tractable by the key idea of “substitution at the margin.”

**Suspending the Price Mechanism**

Coase recast Marx’s observation of the differences between relationships inside and outside of the firm by emphasizing the differences between them with respect to coordinating function, as opposed to the differences between them with respect to purported norms. The puzzle, as Coase puts it, is that “outside the firm, price movement directs production,” but “within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-coordinator who directs production.”35 Outside of a firm, a person switches from task A to task B because he will be given a higher wage at B (or, more specifically, because he will derive greater marginal utility doing B); within the firm a person switches from A to B because he is told to do B. Coase was, of course, not the first to note this difference. Robertson, Coase’s contemporary who he quotes, noted in rather elegant prose that firms were “islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk.” What makes Coase’s formulation original was the observation that, although we can explain the reason why butter coagulates, there was no explanation of why these islands of conscious power develop. Economics had focused on the justification for the “ocean of unconscious cooperation” with equilibrium theory and substitution at the margin, while simply accepting the anomaly of the firm without inquiring as to why such an “island of conscious power” would develop in the first place. To explain the latter, one needed to understand the firm primarily in terms of its supersession of the price mechanism. Therefore, Coase’s first foray into the productivity component of the firm was to establish

---

that the relevant relationships of the corporation were those economic transactions that were not coordinated by the price-mechanism.

Yet why was the firm organizing itself in this manner? Just as Coase explained what made the firm distinct by first considering what characterized the market, Coase answered what the firm was able to accomplish by first looking at what the market wasn’t able to accomplish, namely an efficient allocation of goods and services. More precisely, the market underproduces, or does not produce, certain goods and services because of costs inherent to using the pricing mechanism itself. Because of inefficiencies inherent to the price mechanism, there are efficiency-gains to be had that can be realized by suspending the price mechanism. Coase cites a number of these costs that firms are able to economize on. The first and main sets of costs are those associated with discovering and agreeing upon the prices themselves: “The most obvious cost of organizing production through the price mechanism is that of discovering what the relevant prices are…The costs of negotiating and concluding a separate contact for each exchange transaction which takes place on a market must also be taken into account.”

In the internet age, concern with discovery of prices seems more than a little outmoded. However, the example is worth mentioning as it is here that the divergence between Coase and Pigou is made clearer. As Medema notes: “Pigouvian analysis, from Pigou on down, did not contemplate the notion of transaction costs, and so implicitly assumed that transaction costs are zero.” By making the costs of conducting transactions a key part of the firm’s efficiency, Coase pointed to the unrealistic assumptions underlying the Pigouvian system of economic analysis. Coase’s recognition that using the price mechanism to organize production actually requires a number of different tasks—discovering the prices available, determining the correct price, and bargaining over on the contractual terms—brings the idea of transaction costs to the fore.

36 Ibid, 391.
37 Medema, The Hesitant Hand, 114.
If securing efficient transactions will also entail certain costs, economic analysis must concern itself with the relative costliness of various transactions, and how the contractual relationships created will differ depending on these costs. This allows us to understand what makes the relationships within the firm unique:

A factor of production (or the owner thereof) does not have to make a series of contracts with the factors with whom he is cooperating within the firm, as would be necessary, of course, if this cooperation were a direct result of the working of the price mechanisms. For this series of contracts is substituted one…The contract is one whereby the factor, for a certain remuneration…agrees to obey the directions of an entrepreneur within certain limits. The essence of the contract is that it should only state the limits to the powers of the entrepreneur. Within these limits, he can therefore direct the other factors of production.38

Whereas contracts are generally taken to be the mutual agreement of equals, Coase here notes that the contractual nature of the firm is in the establishment of hierarchy. The juridical device of the liberal individual par excellence is posited as the basis for the hierarchy and sociality within firms; Coase plays a tune not dissimilar to Marx’s, but in a far less critical key. Here hierarchy is not a device for domination but an order that obviates the need to go through the costly process of writing and bargaining over multiple contracts.

There are a number of conceptual innovations inherent in this that should be flagged. The first is the posing of intra-firm activity as a series of economic transactions not dissimilar to transactions that take place on the market. By focusing on the supersession of the price mechanism within the firm, Coase points out that buying and selling is only one form of economic coordination; the coordinated and cooperative exchange of goods without sale is, counterintuitively, posed as another form of transaction, one that takes place within the firm. Although economics after the marginal revolution is concerned with the exchange of scarce goods, economists had unwittingly been neglecting a whole genus of exchanges because they don’t come with prices. That is, while the firm replaces a series of contracts with one open-ended contract that empowers an entrepreneur, it would be a mistake to see this as the only transaction occurring. While this may be the only transaction that the firm as an entity

engages in, *within* the firm, workers, raw goods suppliers, and all the other factors of production all transact, only they do so at the direction of the entrepreneur. What we have are different methods for coordinating transactions, methods proving to be more or less efficient depending on various factors. The question then isn’t just whether there will be a transaction, but how that transaction will be organized.

The second major conceptual innovation we see here is the portrayal of the firm as a particular and distinct type of contractual relationship. When using the price mechanism one will desire a contract that is most specific and clear about what is expected from both contracting parties. Yet what Coase notes is that when we suspend the price mechanism we do so by creating an open-ended contract that is fuzzy on the details *by design*. Far from being a problem, this vagueness is actually an ingenious alternative to the price mechanism. Doing this allows for future flexibility while avoiding the costs inherent in making a new contract for every transaction. What this draws our attention to is that contract is actually a large category containing important and distinct qualitative types. Just as Coase asks us to note a set of ignored transactions so as to force the question “which kind of transaction?” we are also asked to note a distinct class of contracts so as to force the question “which kind of contract?”

Whereas classical economic theory saw firms as being in competition with each other, Coase’s analysis suggests something more nuanced: the individual transactions that constitute the firm are to be analyzed in comparison with their possible market manifestations. That is to say, *firms don’t merely compete with other firms in the market, but with the market itself*:

The operation of a market costs something…by forming an organization and allowing some authority (an “entrepreneur”) to direct the resources, certain marketing costs are saved. The entrepreneur has to carry out his function at less cost, taking into account the fact that he may get factors of production at a lower price than the market transactions which he supersedes, because it is always possible to revert to the open market if he fails to do this.39

39 Ibid, 392.
The entrepreneur is important and interesting from an economic perspective only insofar as his presence contributes to the organization of activity that outperform the market. Once this ceases to be true, people will resort to transacting and contracting using the market’s price mechanism. On the other hand, as the range of goods and services that the entrepreneur is able to further economize on expands, then the size and operation of the firm too will expand. This leads to a rather elegant claim regarding what the firm is and what it is composed of: “a firm…consists of the system of relationships which comes into existence when the direction of resources is dependent on an entrepreneur,” with the entrepreneur previously defined as the person to whom authority has been contracted to direct resources.

It is important to emphasize the degree to which the entrepreneur’s function is in competition with the market’s price mechanism, since it serves to round out the theory. If using the price mechanism can be so costly, and if entrepreneurial authority can be so economizing, we might well ask Coase, as we asked Mill, why we need a market at all. Why not, from the perspective of efficiency, simply have a planned economy where we can just skip the costs associated with specific contracting and market transacting? Coase suggests three different constraints on the firm’s size and growth: 1) the costs entailed in organizing, 2) the proclivity for an entrepreneur to err in her allocation of productive resources, and 3) changes in the cost of factors of production as they are found in the market. These costs are not static since the costs of organizing and the tendency for entrepreneurial error tend to rise as the number of transactions being organized within the firm rise, hence the diminishing returns of management. All of this is to say that the firm is constrained by virtue of the entrepreneur’s competition with the market; when it is cheaper to do a particular transaction though the price

40 I specify this because Hayek and Schumpeter accord the entrepreneur an additional social and intellectual importance which is not required for this argument (nor desired from my own perspective)
41 Ibid, 393.
42 Ibid, 394-395.
mechanism, or when the entrepreneur is managing poorly, then the firm will witness contraction and outsourcing. So while the planning within firms bears an important similarity to the planning of socialist economies, Coase claims this crucial difference: “economic planning is imposed on industry, while firms arise voluntarily because they represent a more efficient method of organizing production. In a competitive system, there is an optimum amount of planning.”

The Limits of Coase’s Account

We can see that Coase provides a fairly compelling and parsimonious argument to explain the productivity of the firm. To the question of constitution, Coase contends that the firm is constituted by those relationships that rely on the services of an entrepreneur instead of the price mechanism. The question of efficiency is answered by recourse to the inefficiencies associated with using markets; the firm is an efficient mode of production because it avoids the costs of the price mechanism by substituting the authority of an entrepreneur for the market. Finally, the question of constraint is answered by claiming that there are also costs associated with hierarchical organization, costs that increase as the scope of such organizing increases; the firm ceases to grow when the cost of doing a particular extra transaction is equal to doing the same through the price mechanism.

Yet Coase’s elegance can only get us so far. As has been noted by many since its publication, Coase’s argument is somewhat tautological. That is, Coase does not entertain the idea that firms might be inefficient forms of organizing, nor does he argue against such a claim. Despite Coase’s implicit critique of welfare economics, he here still seems bound up in Walrasian and Paretian economics, using

general equilibrium theory as a presupposition that the market produces the efficient allocation of all goods and services; Coase is able to assert that the market produces the optimal level of planning, the costs of discovering market prices notwithstanding, because it can be assumed that planning wouldn’t be allocated were it not efficient. Coase’s claim that a firm represents a nexus of efficiency gains simply follows from the initial assumptions. By assuming that a given institutional configuration is in equilibrium, it is trivial to then find that a particular institution is in fact efficient. What is needed, and what Coase does not give directly, is an explanation of which transactions, and under what circumstances, will tend to be economized (or not) using hierarchical organization, and why (or why not). Absent this, what we have is a conceptual reframing oriented toward (and not a full-blown theory explaining) the comparative efficiencies involved in hierarchical organization.

Even if we excuse this tautology, Coase’s theory is incomplete for our own purposes because of the general subject that it explores. Coase’s understandable concern is with firms in general, which would include a range of business types. While these are not completely outside the focus of this project, it must be kept in mind that we are concerned with a theory of the corporation in particular, not the firm in general. Thus, even if we excuse the vagueness of Coase, we would want a further explanation of why the corporation comes about, preferably by a parsimonious reference to the same idea of costs being associated with particular forms of organizing transactions.

To sum up, Coase’s argument made at least four important conceptual steps toward a theory of the corporation. First, it is one of the first forays into the internal nature of the firm to be conducted within the post-marginal revolution framework. In so doing, Coase is responsible for the conceptual innovation of seeing the activities within the firm as being types of transactions, making the transaction itself an important variable. Second, because of this conceptual innovation, Coase enabled the shift of the primary economic question from “what price will enable the most efficient transaction?” to “what

type of characteristics and inefficiencies does a given transaction display, and what consequences follow from it?” Third, Coase also put forward the type of contract as an economic variable, forcing the economic analyst to attend to the open-endedness of some contracts compared to others. In so doing, Coase posited a crucial link between the juridical contract and the economic transaction.

Finally, we can see this argument as taking the theory of general equilibrium to its logical conclusions with surprising results; by noting that the firm is characterized by the supersession of the price mechanism while still existing within a market, Coase was able to claim that there is an optimal level of planning or, put different, that there is an optimal level of non-market activity. Thus, in contrast to Smith’s famous dictum, Coase points out that the extent of the market might sometimes constrain the extent of the division of labor. Instead, the extent of the division of labor is limited by the extent of institutional variety that includes, but is decidedly not exclusive to, markets with scarcity pricing. Philosophically, the efficiency-orientation of this insight provides the basis for the welfarist justification of non-market institutions, as well as the basis for rational choice institutionalism.46 This also pointed to the conceptual inadequacy of the reigning Pigouvian paradigm that assumed the absence of transaction costs (I explain this in greater detail in chapter 4). By showing transaction costs to be precisely the reason why firms exist, Coase showed the Pigouvian model to be seriously flawed, since its assumptions would render unnecessary what appears to be an enduring feature of market activity.

Yet before we can affirm the account of corporate productivity implied by Coase’s argument we need two more pieces of the puzzle. First we require an account of precisely why and when the firm’s supersession of the price mechanism makes for an efficient form of organizing (and why/when it doesn’t). Put in social science terms, we need an argument that considers the type of transaction as a dependent variable so as to explain what factors account for its difference (and in turn, the presence or

---

absence of the firm). Second, we need an argument that shows how this account can be used to explain the specific efficiency gains achieved by the corporation. In the next section I show how both of these are accomplished by Oliver Williamson in his development of “transaction cost economics.”

**Transaction Cost Economics**

In extending this Coasian approach into a fuller economic system of inquiry, Oliver Williamson has developed a school of thought known as “transaction cost economics.” Taking the Coasian idea that the transaction was a qualitative variable and not just a quantitative one—that is, the economic question need not only be concerned with how many transactions would occur, but what kinds—Williamson took the step of putting forth the transaction as the primary unit of analysis for economic inquiry, departing from neoclassical economics which takes the individual consumer or supplier as the basic unit of analysis. While this in itself isn’t a large advance over what we see in Coase, Williamson specified the precise features of human behavior and the relevant technical dimensions of transactions in order to explain more exactly when a transaction would transform from market-coordinated to hierarchically-governed (or vice versa).

**Environmental and Behavioral Dimensions**

Williamson cites John Commons as an antecedent who proposed the transaction as a basic unit of analysis at roughly the same time when Coase was penning his piece on the firm. According to Commons, using the transaction as the primary unit of analysis allows the investigator to better grasp how institutions determine which forms of economic activity are valid and which ones are not.\(^47\) Commons’ main point, which Williamson picks up on, is that the shape and meaning of a transaction

cannot be reduced to the desires of the transacting parties or the commodities being exchanged since these are influenced by larger environmental factors; political economy must take account of the “collective control of individual transactions”\(^{48}\) exerted by institutions.\(^{49}\)

Williamson posits two psychological-behavioral traits as central for understanding transaction costs: bounded rationality and opportunism. Citing Herbert Simon, bounded rationality is defined as the assumption that economic actors are “intendedly rational, but only limitedly so.”\(^{50}\) That is to say, while the general maximizing orientation of neoclassical understandings of rationality is granted, Williamson contends that the orientation toward maximization does not imply behavior which actually achieves it. This is because human rationality is bounded by neurophysiological limitations (the finite capacity to attain all relevant information, and finite ability to perform the necessary calculations based on attained information) and linguistic limitations (the finite ability to effectively and efficiently convey information to others), making the actual achievement of all possible maximizing behavior improbable. While this appears to be a plausible enough position which might simply be asserted, this is actually a rather subtle positioning of the transaction cost view vis-à-vis neoclassical economic modelling, implying a rather damning criticism.

Williamson notes that the Arrow-Debreu model of general equilibrium—what we might reasonably think of as the technical proof of Adam Smith’s invisible hand, since it implies that markets will reach the Pareto-frontier in all possible markets—requires that contingency-claims markets be complete and in equilibrium. This is necessitated by their first assumption: “each commodity may be bought or sold for delivery at one of a finite number of distinct locations and one of a finite number of future time points. For the present purposes, the same commodity at two different locations or two

\(^{48}\) Williamson, *Economic Institutions of Capitalism*, 45.

\(^{49}\) As I discuss in greater detail in chapter 5, this is also similar to Durkheim’s view of contracts.

\(^{50}\) Ibid, 45.
different points of time will be regarded as two different commodities.” General equilibrium, then, presupposes an ability not merely to make contracts that specify the terms of a transaction at this place and at this time, but to make contracts that specify the terms for any future state of the world. Insofar as we accept the Arrow-Debreu theorem, we must also accept the assumption of full rationality, since the ability to write contracts in all markets for all possible states of the world requires a herculean ability to retain, process, forecast, and communicate a great amount of information. However, insofar as we have reason to doubt our ability to be fully rational—and the mere existence of a firm gives us such reason, since it is an institution defined by the absence of fully-specified contracts—we have reason to doubt the relevance of the general equilibrium model.

Put differently, whereas Coase critiques Pigou for being unrealistic with his assumptions when assuming the absence of transaction costs, Williamson point out that this unrealistic assumption stems from a more basic misunderstanding of the economic actor itself. Transaction costs exist precisely because man is only so capable of attaining and processing information, and only so capable of communicating information. The question is thus not whether transaction costs exist or whether man is fully rational—they do, and he isn’t—but what factors make the boundedness of rationality important from an economic perspective so as to make for costly transactions. When does the economic actor’s recognition of his own inability to choose fully maximizing behavior lead him to create alternative contracts which account for his neurophysiological and linguistic limits?

While bounded rationality is an important feature of the transaction cost paradigm, Williamson downplayed it in his later writings. The second, and more significant, behavioral assumption is opportunism. Opportunism is defined by Williamson as the propensity to be strategic or, more

forcefully, to exhibit “self-interest seeking with guile,” which involves either explicit or subtle forms of deceit. It is helpful to understand this assumption as the inverse of bounded rationality. Whereas bounded rationality assumes that people are utility maximizers with finite computational and communicative capacity, opportunism suggests that the information people do possess will be used strategically and communicated selectively so as to maximize their interests. This is most relevant when it comes to information about oneself, for instance what one’s true preferences are or how hard one plans to work in the future. Opportunism is merely the “maximizing orientation” assumption of full rationality taken to its logical conclusion, with communicating and signalling recognized as potential opportunities for strategic behavior.

Again, this is a subtle indictment of neoclassical economics. Williamson notes that “neoclassical man” is presumed to be motivated by “simple self-interest”: while economic agents approach bargains with the intent to maximize their utility as much as possible based on the initial positions of everyone’s utility functions and preference schedules, the initial positions themselves “will be fully and candidly disclosed upon inquiry, state of the world declarations will be accurate, and execution is oath- or rule-bound.” To put this in philosophical terms which Williamson does not, the neoclassical economic agent is presumed to have consequentialist motivations when assessing all information and determining activity. Except when it comes to the information he knows best: information about himself. When it comes to himself, the neoclassical agent is assumed to communicate and act upon information involving himself in the most transparent and honest manner. Such an assumption is necessary for the model; given that information is perfect and contracts are held to be complete across future markets, we must assume that present attitudes and future predilections are known and made known to all. Yet no account is given for this startling change in moral psychology.

52 Williamson, Economic Institutions of Capitalism, 47.
53 Ibid, 49.
The opportunistic assumption in a theoretical sense merely jettisons this inconsistency, assuming the economic agent to be consequentialist all the way down, and acting out of his self-interest to the greatest extent possible. This may appear to be an overly pessimistic view of human nature. However, the opportunistic assumption does not mean that we ignore the possibility of people portraying themselves in an honest fashion. Opportunism does not require the assumption that all people are in fact always opportunistic; instead it assumes that “some individuals are opportunistic some of the time” but “that differential trustworthiness is rarely transparent *ex ante.*” We need not actually be opportunistic in order to take opportunism into account when contracting. The question, again, is what circumstances bring opportunism to the fore and affect the types of contracts which are created.

*Transactional Features*

In addition to these behavioral assumptions, Williamson identified three dimensions of the transaction that are relevant in explaining how a transaction will be organized: uncertainty/complexity, frequency, and asset-specificity. On the first count, all transactions will contain some degree of uncertainty because, as the bounded rationality assumption suggests, we are incapable of foreseeing all possible contingencies and because promise-keepers and promise-enforcers can never be fully relied upon. That said, uncertainty is not constant; therefore transactions can be conducted with more or less certainty depending on the predictability of contingencies, forecasting of future markets, and reliability of other parties. Frequency as an important trait is directly tied to uncertainty; the more frequently a given transaction occurs, the less uncertainty accompanies it. Furthermore, frequency also allows temporal elements to enter into the transaction. A transaction that tends to repeat itself between the

---

54 Ibid, 64.
55 Ibid, 52
same two parties on a regular basis can be economized by trust, allowing for looser terms of contracting. Furthermore, frequency not only creates economic gain, it also can outweigh economic losses. If start-up costs are required to enable a particular transaction to occur, the more frequent that transaction will be, the more economical it is to make the initial outlay. Therefore the more frequent the transaction, the greater the incentive to economize through the costly setup of governance mechanisms.

Finally, and most importantly, is asset-specificity. The idea here is that certain transactions will require investments in specific assets that can only be used for relatively few goods, or in certain skills for which only relatively few customers would demand. For transaction-specific investments the value of the product produced by an investment of time or resources has less value in any use other than the specific transaction for which it was produced. This puts the would-be investor in a precarious position, as he is committing himself to a shrunken market of potential buyers in the future:

Parties to a transaction commonly have a choice between special purpose and general purpose investments. Assuming that contracts go to completion as intended, the former will often permit cost savings to be realized. But such investments are also risky, in that specialized assets cannot be redeployed without sacrifice of productive value if contracts should be interrupted or prematurely terminated. 57

The incentives to engage in these types of investment are very low, as the investor is put in a vulnerable position vis-à-vis potential customers who would now be able to exercise their market power as monopsonist for the particular good. Thus asset-specific investing is rendered problematic because of our inability to be assured of continuous future transactions; the opportunistic nature of people implies that one might display a willingness to cooperate on good terms in the future, only to defect once the initial investment was made.

It is worth stressing asset-specificity as a problematic dimension of transactions since it is so pervasive. Williamson notes four types: site specificity (the locating of production or services in a specific place); physical asset specificity (the development of machinery or physical capital which produces a specific kind of good); human asset specificity (the development of knowledge or

57 Williamson, Economic Institutions of Capitalism, 54.
relationships specific to a particular transaction); and dedicated assets (the locking-in of physical or financial capital towards a particular purpose). Their pervasiveness means that even transactions that don’t begin with asset-specific investments might acquire that character as the economic relationship continues. For example,

Human capital investments…evolve during contract execution. Specialized training and learning-by-doing economies in production operations are illustrations. Except when such investments are transferable to alternative suppliers at low cost, which is rare, the benefits can be realized only so long as the relationship between the buyer and seller is maintained.58

Winners of initial contracts develop an advantage over their competitors allowing them to extract more out of their transacting party; on the other hand, they have also spent time acquiring information and skills tailored for one transacting party, making their position precarious. For both sides then, the existence of asset-specificity creates the need to develop safeguards against future free-riding off of initial outlays or investments specific to the transaction.

Again, the contrast here is with the neoclassical model of equilibrium that assumes that the identity of buyers and sellers doesn’t matter. What Williamson shows is that even generic transactions between anonymous parties might very well tend to produce asset-specificities that make the identity of the party a relevant feature of the transaction, creating a “party-relevant transaction.”59 Once granted, the reasons why the market will be inefficient at coordinating these transactions becomes clear: absent the competition amongst large numbers of people as postulated by the neoclassical model, certain parties are granted asymmetrical positioning in relation to their transacting partners. This may result in rent-seeking behavior, price gouging, the underdevelopment of goods (since the price will be inflated), or, most importantly, the foregoing of the initial investment to begin with, resulting in missing markets. Therefore even if we grant the unrealistic assumption of anonymity and party-irrelevance in the market this would mean that the only transactions being coordinated are those which require no party-

58 Ibid, 62.
59 Ibid, 54.
relevance, leaving a large set of transactions foregone and, in turn, a set of efficiency gains unrealized.

*The Fundamental Transformation and the Corporation*

The underlying Coasian idea that firms mitigate transaction costs is captured in Williamson by the idea of the “fundamental transformation” of a contract, when “governance structures that attenuate opportunism and otherwise infuse confidence are evidently needed.” 60 It is “fundamental” because it represents the point at which the inefficiencies and inconveniences of hierarchy and direction are outweighed by those accompanying the transaction were it to be carried out in the market. It is a transformation because the tradeoff between bureaucratic costs and adaptive capacity “switches from net negative to positive as the condition of asset specificity deepens,” rendering the development of governance structures the efficient contractual arrangement. 61 Coase pointed out that the entrepreneur clearly plays a role in promoting efficiencies that the market is unable to; Williamson demonstrated why by noting that the growth in complexity and uncertainty combines with our bounded rationality—and a growth in frequency and asset specificity combines with our concern for others’ opportunistic behavior—to make for a class of transactions that require better assurance than one-time market contracts can give.

As a result, the firm can be rationally reconstructed as a form of private ordering—“efforts by the immediate parties to a transaction to align incentives and to craft governance structures that are better attuned to their exchange needs” 62—used to contractually establish those mechanisms which will allow such transactions to occur. While the mechanisms are diverse and too technical to go into here, the basic point is that by establishing governance mechanisms through long-term open-ended contracts, the firm 1) establishes uniform communication patterns to overcome limits of communication between

---

60 Ibid, 63.
parties; 2) allows for processes to repeat, enabling the better understanding and processing of information amongst transacting parties; and most importantly 3) mitigates the problem of opportunistic behavior by bringing a large number of future transactions within one contract, removing the temporal element necessary for the opportunistic use of information.

The canonical example of this discussed by Williamson, though originally reported by Klein, Crawford, and Alchian, is the 1928 merger of Fisher Body and General Motors. In 1919, Fisher Body and General Motors signed a ten year contract, according to which Fisher Body would produce state of the art closed auto bodies for GM’s car plants. The vulnerable positions created by such a deal, however, were not lost on either party:

In order to encourage Fisher Body to make the required specific investment, this contract had an exclusive dealing clause whereby General Motors agreed to buy substantially all its closed bodies from Fisher. This exclusive dealing arrangements significantly reduced the possibility of General Motors acting opportunistically by demanding a lower price for the bodies after Fisher made the specific investment in production capacity.

By writing the exclusive dealing clause into the contract, Fisher was effectively secured against GM’s use of monopsonistic market power to create a hold-up and demand lower prices. To protect GM from Fisher using its position to raise prices once both parties were committed, a formula for determining a fair price was written into the contract. This formula, however, proved insufficient. Unable to forecast the sudden rise in demand for the new closed body cars, GM found itself overpaying for bodies that were being made in greater quantities than either party had thought at the outset. Furthermore, GM wanted Fisher to locate its plant in closer proximity to its own manufacture in order to increase efficiency, a move that Fisher balked at because of the site-specificity involved. To make the deal happen, GM eventually bought the remainder of Fisher’s stock and simply merged Fisher Body into General Motors. As Williamson describes it: “faced with the prospect that both operating and

64 Ibid, 308-309
investment decisions would be out of alignment during much of the rapid growth stage of development, bilateral governance eventually gave way to unified governance.”

The Fisher Body example helps elucidate how information constraints and the concern about opportunism led two parties from standard contracting to bilateral governance structures, to vertical integration. It therefore illustrates how hierarchy and planning contracts can supplant market transactions because doing so benefits all parties involved, giving them assurance that their initial investments will not be exploited. This also allows them to engage in more frequent transactions, thereby capitalizing on the routinization that Smith affirmed, and to work in large scales with less concern for free-riding, thereby capitalizing on Mill’s celebrated simple and complex cooperation. Where the requisite degrees of asset specificity, complexity, frequency, and uncertainty do not obtain we can expect markets to function more efficiently than hierarchies, and simpler one-off contracts to be written; this was the case before GM needed Fisher to start making particular closed auto bodies for its plants. Furthermore, even where such dimensions are relevantly present, their costs might be outweighed by the costs associated with diminished incentives that tend to attenuate hierarchy and bureaucracy.

While this is convincing as an account of why firms in general have efficient properties, we still have not answered why the corporation in particular is efficient. To understand the particular efficiency gains afforded by the corporation, we need to apply such an analysis to particular features of the corporation: for example, in what way does the transferability and tradability of shares economize on other forms of financing? While the general points that Smith made originally still hold—that is, the corporation sells shares in order to acquire larger stores of capital—we might wonder why these particular methods are used to raise capital as opposed to others. Recall that Smith argued that banks won’t willingly fund the amount of capital that can be raised through stocks and shares. The question,

however, is why this is the case. Here we see that the account given above can be applied to the question of corporate finance as well: debt should be understood as the market-like mode of financing a firm as it involves complete and specific contracting, and equity should be understood as importing administration for financial purposes as it gives its holders rights to decision-making and residual claims. It therefore follows that the degree of asset-specificity and complexity will determine the type of financing used for a firm: “debt is a governance structure that works out of rules and is well-suited to projects where the assets are highly redeployable. Equity [“the financial instrument of last resort”] is a governance structure that allows discretion and is used for projects where assets are less deployable.”67

In short: some firms will be able to finance their operations through debts when their assets might be used for other purposes in the event of forfeiture. Absent this redeployability—or in the presence of asset-specificity—financing won’t be as readily available through debt and so other means are necessary; the use of stocks and shares, limited liability, and shareholder primacy are precisely the mechanisms through which investment is induced for such enterprises. The result is a view of the corporation where all of its features are explained by virtue of the asset-specific investments they enable, and the transactions which constitute it—both financial and productive—are mutually reinforcing.

**Conclusion**

With the development of transaction cost economics we get a full account of the corporation’s productivity. Following Coase, the corporation is constituted by those transactions where the price mechanism has been suspended in favor of entrepreneurial planning. It is efficient because markets will underproduce goods and services that require great amounts of information (thereby taxing our bounded rationality) or that shrink the number of suppliers capable of producing a good (thereby

making us susceptible to opportunistic behavior). The firm creates governance structures that regularize information and communication and function as protections against free-riding and defection in future transactions; these governance structures provide assurance both for workers and suppliers (in the form of employment relationships and vertical integration) and for capital (in the form of tradable shares and rights to dividends and voting). Finally, the corporation is constrained by the costs accompanying the installation of bureaucratic mechanisms and the limiting of competition, which are inherent to administering transactions through hierarchy as opposed to prices.

The transaction cost account of the corporation’s productivity component is shown in table 2.2 below in comparison with its historical forerunners reviewed earlier. I have presented this theory in historical fashion. However, the goal has not been to give a conceptual history of the theory of the firm per se, but to give an account of the transaction cost theory’s strengths by showing how it

Table 2.2. The Productivity Components of the Transaction Cost Approach in Comparison

<table>
<thead>
<tr>
<th></th>
<th>Question of Efficiency</th>
<th>Question of Constraint</th>
<th>Question of Constitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Smith</strong></td>
<td>Division of labor</td>
<td>Principal-Agent problems and general malfeasance</td>
<td>No Answer</td>
</tr>
<tr>
<td><strong>Mill</strong></td>
<td>Cooperative labor; quality of labor induced by corporate remuneration</td>
<td>No Answer</td>
<td>Relationships between various kinds of labor</td>
</tr>
<tr>
<td><strong>Marx</strong></td>
<td>Creation of surplus value through exploitation of labor</td>
<td>Corporation represents the overcoming of markets: it has no ahistorical constraints</td>
<td>Social labor and the social capital characterized by hierarchy and domination.</td>
</tr>
<tr>
<td><strong>Pigouvian</strong></td>
<td>Presumed as Production Function – “Lumps of butter coagulating in a pail of buttermilk”</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transaction Costs</strong></td>
<td>Facilitates party-relevant transactions</td>
<td>Costs of bureaucracy and lack of competition</td>
<td>Transactions coordinated by entrepreneur</td>
</tr>
</tbody>
</table>

68 That has involved some blurring of the actual historical record: I have painted the classical economists as being much more concerned with the corporation than they were, drawn implications regarding the firm from the marginal revolution which were not explicitly stated by the marginal revolutionaries, and I have portrayed Coase as responding to Marx (when his sights were clearly aimed at Pigou). If I was primarily interested in intellectual history this would be something requiring more justification. However, since I am not, I will just note that this approach seems to mesh well with what Lakatos refers to as a rational reconstruction of history, where one gives a stylized account of intellectual history to better understand a current practices or idea. In so doing, one relates the reconstructed internal history of a discipline in the text “and indicate in the footnotes how actual history misbehaved.” See Imre Lakatos, 1971. “History of Science and its Rational Reconstructions.” In PSA 1970: Boston Studies on the Philosophy of Science 8: 107.
overcomes other theories’ weaknesses. By putting Coase in dialog with Marx we are able to see that, in contrast to the classical and Marxian approaches to the firm, the transaction costs approach eschews the labor theory of value and methodological classism in favor of a subjectivist approach both with regards to value (as embodied in the principle of marginal utility) and the methodological individualism implied by the transactional emphasis. In so doing a more coherent understanding of why hierarchy exists within the employment relationship is unearthed, as well as an explanation for why tradable shares are a more efficient way of raising capital than debt. In so doing, transaction cost economics rebuilds an account of how and why capital, labor, and raw goods cooperate together within a corporation, but without falling back on the classical strategy of ontologically assuming classes into existence. Furthermore, it is more parsimonious, drawing out the efficiency and the constraints of the corporation from the same principle (the dimensions of a transaction and its interaction with human behavior) as opposed to drawing on different assumptions to explain different features of the productivity component (the benefits of cooperation for efficiency, and principal-agent problems for constraint).

In contrast to Pigouvian welfare economics, the transaction approach denies the idea of the firm as merely a production function relying on technology for efficiency. Instead it recognizes the costs incurred in the very process of transacting. In showing how a transaction will take on a party-relevant form, it shows why assumptions of perfect competition and perfect knowledge of markets are unrealistic, and why party-relevance can lead to fear of opportunistic behavior which complicates the basic model. In so doing, the transaction cost approach makes a more fundamental critique in demonstrating the incoherence of the neoclassical subject: to recognize the reality of friction in the process of transaction requires recognizing both the bounded rationality of subjects as well as the strategic potential of signalling/communicating information. Put differently, the presumed maximizing orientation of the economic subject is in direct tension with the information assumptions necessary to produce a model of efficient markets, since the strategic rationality of the former will result in a
distortion of the latter. In that sense, the corporation as a phenomenon is not merely a puzzle from the standpoint of efficiency; it is an indictment of the most basic assumptions undergirding the faith in the invisible hand of the free market. Once one recognizes the efficiency of the corporation one needs to recognize that strategic rationality only contributes to efficient allocation for particular types of transactions. In other situations, strategic rationality will create numerous inefficiencies that require open-ended contracts to create space for forms of behavior structured by norms other than competition and strategy. The fact that so much production gets carried out by corporations (let alone the larger category of firms in general) points to just how pervasive such transactions are.
Chapter 3. The Political Theory of the Corporation of the Chicago School

In the last chapter I argued that the advent of transaction cost theory marked a watershed moment in the history of economic thought. For the first time, economists were able to offer an account of why integrated business firms exist in a competitive market without classical political economy’s ontological baggage of methodological classism or the labor theory of value. The transaction cost account of the firm challenges certain assumptions inherent to neoclassical economics, but is still given within the marginalist terms of 20th century mainstream economics.

This theory opened the black box of the firm and enabled answers to the productivity component of the theory of the corporation. However, despite its strength in explaining corporate efficiency, it does not offer an account of how the corporation ought to be structured, nor does it pretend to. Transaction cost theory does not attempt to explain if and how the corporation can be understood as consonant with the overarching social and political values of a liberal society, but simply how it can be explained in terms of market efficiency. Or, in institutional terms, we might say that transaction costs explain how the firm fits within the normative logic of the market, but not of the state. What remains to be seen is how the contemporary economic theory of the firm answers the normativity component of a theory of the corporation: how should a corporation be governed, and how should corporate actors be constrained legally and ethically?

In this chapter I argue that the Chicago school tradition of economic thought has produced the most influential normative account of the corporation. The Chicago school has been incredibly influential in large part due to what we might think of as its intellectual moxy and its willingness to expand economic analysis. Like Foucault in his now canonical critique of “neoliberalism,” I take the Chicago project as being fundamentally an attempt to bring non-economic subjects under the scrutiny of economic analysis. In many ways this is in the spirit of Coase’s original contention that one could account for both the market and the firm with a common economic logic. Coase wanted to show that
the firm was a distinct beast from the market, but that both were explainable by recourse to economic efficiency. And yet, as we will see, the Chicago School program has resulted in a slight but significant perversion of the Coase-Williamson line; whereas transaction cost theory emphasizes the institutional distinction between firms and markets, the Chicago program was and is primarily dedicated to showing how the logic of the market can be used to understand almost anything. This chapter unpacks how the Chicago school has attempted to reduce the firm to a series of market transactions, and the normative consequences of this reframing effort.

I begin this chapter with some remarks on what I mean by the “Chicago School” or the “Chicago approach.”¹ I then articulate the managerialist view of the corporation that the Chicago School was countering. I lay this view out primarily through an overview and analysis of the most exemplary and influential text from the managerialist tradition, Berle and Means’ *The Modern Corporation and Private Property*. In particular I attempt to show how Berle and Means articulated a theory primarily aimed at criticizing extant corporate practices, how this criticism tied into particular normative recommendations, and the broader political commitments such a view of the corporation implies. In the rest of the chapter I show how Chicago economists responded to the managerial account. In section two, I show how Milton Friedman used the Chicago logic to produce a theory of business ethics that challenged the managerialist emphasis on property in favor of a view of the corporation that centers on the principal-agent relationship between management and shareholders. In so doing he articulated the classic account of business ethics as the maximization of profit. In the third section, I demonstrate how the question of corporate governance was addressed by agency theorists and property rights theorists who showed how pervasive corporate principal-agent relationships were, and how these imply a view of the corporation that is entirely grounded in market exchange. In so doing they effectively blurred the distinction between a firm and a market, suggesting that corporate

¹ I use terms like “Chicago School,” “Chicago approach,” or simply “Chicago” interchangeably throughout.
governance is and should be conducted through securities markets. In the fourth section, I demonstrate how these ideas were developed by Chicago law and economics scholars into a theory of corporate law that ties these disparate threads together into a coherent normative vision largely oriented to deferring to corporate practice as it is found. I conclude by summarizing the whole of Chicago’s political theory of the corporation. As a coda, I offer some reflections on how this theory helps us understand the nature of the Chicago school and “economic imperialism” more generally.

What is the Chicago School?

Chicago’s Most Famous Critic

The most important work of political theory criticizing the Chicago economics project is likely Michel Foucault’s *The Birth of Biopolitics*. Presciently delivering his lectures on the rise of neoliberalism in 1979, just before the election of Thatcher and Reagan, Foucault argued that Chicago economics had radically recast the political subject, simultaneously depoliticizing her while also making her more susceptible to governance by economic logic. In the process, Foucault made the compelling case that this project had altered the received liberal doctrine by making the market the primary mode of social control, and generator of social truths. Given the sheer influence of this text it is worth saying a few words about the how my account of the corporation relates to Foucault’s analysis. Though I take his contribution to be very important, I also think it has significant limits, which I try to avoid.

Like Foucault, I take the Chicago project fundamentally to be an attempt to bring non-economic subjects under the scrutiny of economic analysis. Our current structures of corporate law and corporate governance are the result of Chicago scholars who wanted to show the relationships within the

---

2 Foucault used the term “neoliberal” to describe this tradition. This term has now come to encompass much more than simply Chicago style economic thinking. For that reason I avoid using this term here.
corporation to be governed by the market’s price mechanism. The main differences between Foucault’s account and my own are methodological. Most obviously, Foucault does not mention the corporation, either as a development in 20th century capitalist economies, or as a focus of Chicago economics. For Foucault, Chicago economics was fundamentally responding to the development of Keynesian economics and the growth of the American welfare state, while I suggest a different area of economic intervention--corporate law and corporate governance--that Foucault does not touch upon. More substantively, there is a subtle difference between my understanding of the project of Chicago economics and Foucault’s that is worth emphasizing, as it makes clear how I view the connection between the Chicago program, the economic theory of the corporation, and its significant normative political theoretic implications.

Foucault contends that the hallmark of the Chicago project is taking phenomena left untouched or un-theorized by classical political economy and bringing them under the scope of economic analysis. In a telling passage, Foucault describes the core features of Chicago economic analysis: “for the neoliberals economic analysis should not consist in the study of [production, consumption and distribution], but in the nature and consequences of what they call substitutable choices, that is to say, the study and analysis of the way in which scarce means are allocated to competing ends, that is to say, to alternative ends which cannot be superimposed on each other.”

In other words, Foucault takes Chicago economics to be defined by its use of principles associated with the Marshallian system of political economy reviewed in the last chapter. There are a couple of peculiar things about this. The first is that the analytical concepts to which he is referring were developed roughly half a century prior to the emergence of the Chicago school; as we saw, the labor theory of value was more-or-less debunked by the marginal theory of utility by the end of the 19th century, transferring the attention of

---


4 Foucault, 222.
economics away from population control and toward the competing use of scarce resources. Certainly we would not consider Pigou to be a neoliberal, despite his systematic use of marginal utility theory.

The other peculiar thing is that Foucault doesn’t explain in non-arbitrary or non-trivial terms what the distinction is between the economic and the non-economic, such that using the analysis from the former to engage the latter would be noteworthy. To a certain degree Foucault doesn’t need to do this: insofar as Foucault is largely concerned with how Chicago challenged previous modes of economic analysis, all he really needs to do is show how such phenomena were previously not under the scrutiny of economists, which he does. However, most have taken his account to be an attempt to show something insidious about the Chicago paradigm, and its imperial attempt to extend the economic into the social. Insofar as the analysis is supposed to have any kind of critical bite a more precise understanding of the distinction between the economic and the non-economic is necessary to make the radicalism of Chicago clear.

While Foucault is right to point out that Chicago economics emphasizes the margin, substitution, and scarcity, it is more helpful to understand their analyses in terms of the presumed outcomes of markets. Another way of putting this is that Chicago does not simply apply economic tools of analysis to non-market behavior, but applies a particular brand of price theory to such behavior, which departs from its Marshallian origins in a number of ways. First, whereas Marshall suggested that scarcity pricing determined by supply and demand has the potential to allocate goods in a way superior to alternatives, Chicagoans took that conclusion as the basic operating assumption; the normative benefits of the market’s price mechanism are therefore foregrounded by the theory. It is telling, to this end, that the first economics class that Milton Friedman taught in Chicago featured only

one text by Hayek, which was “The Use of Knowledge in Society.” The emphasis is on the informational gains that come with market prices, and has less to do with the things that go into constituting that price. Put differently, the economic tools that Chicago wields have less to do with the fact that they translate peoples’ choices and actions in terms of scarcity, substitution, and marginalism, and more to do with the underlying idea that such decisions, when not interfered with, will result in an optimal allocation.

The second major distinction between Chicago’s price theory and earlier variants can be found in Stigler’s canonical work on the subject, where the rational utility maximizer is made central to the theory. Such an assumption had not been nearly as central to previous economic theory; Smith, Marshal, and even Friedman (as we will see), all made a point of not stipulating as a fact that all actors are economically rational. In contrast, Stigler made this assumption a foundational premise. Stigler reread the history of political economy going all the way back to Smith as being “a stupendous palace erected upon the granite of self-interest,” and made this the behavioral premise that could explain, as opposed to merely predict, the function of markets. This was the key shift in method that allowed Becker's analyses, which so fascinated Foucault.

This allows us a more precise understanding of how Chicago attempts to extend the domain of the economic: the goal is to show that price theory holds in areas where it was thought not to previously apply. This can be thought of in two different ways. On the one hand, there are attempts to show how the allocative side of price theory holds in contexts where prices were previously thought to malfunction and misallocate. We can think of this as using economic analysis to vindicate the economic, to show how phenomena thought of as irrational aspects of markets are in fact rational. On

---


the other hand, Chicago projects also aim to show how positive outcomes previously thought to be the result of non-market practices are, in fact, the result of markets, or something akin to them, which explains their optimality.

As a result, Chicago is largely concerned with showing how superfluous non-market mechanisms are. This is tied to the familiar political project of limited government in a straightforward way; because misallocation through the price mechanisms constitutes a “prima facie case” for intervention (to use Pigou’s language⁹), to show that the market is in fact working as it ought to implies a charge against intervention. Similarly, to show that “non-markets” are only functioning well because their incentives are functioning like prices implies the case for the extension of markets. Where markets exist, Chicagoans attempt to show the results to be optimal; where results are optimal, Chicagoans attempt to show that markets exist. With a slight perversion of Hegel’s dictum, we can encapsulate the idea in this way: What is market-based is optimal; and what is optimal is market-based.

An Instructive Example

Perhaps the best and most famous example of “what is market-based is optimal” is what is known as the “Coase Theorem,” or the argument derived from Ronald Coase’s “The Problem of Social Cost.”¹⁰ This is not only a helpful example for understanding Chicago analysis; it helps clarify things about Coase’s place within this tradition. The Coase Theorem holds that it does not make a difference whether one deals with externalities through private contracting or through the legal assignment of liability; either will yield the same result, the efficient result, given the assumption of zero transaction costs. To understand the logic of this, imagine a hot sauce company whose factory gives off pepper

---

fumes that disrupts the lives of its neighbors. This is a clear-cut case of negative externality. Now, imagine that the neighbors sue and the judge rules that the hot sauce manufacture was violating the rights of its neighbors. Let us say as well that the hot sauce company has a huge interest in keeping its location because of its proximity to a pepper patch. If it is that profitable for them to stay—more profitable for them to stay than it is costly for the neighbors to leave—then even despite the ruling, the hot sauce company will pay its neighbors to compensate them for the inconvenience of the pepper fumes, or to offset the costs of their moving away, and will remain in business. If the court had ruled the other way—that the hot sauce company had the right to manufacture hot sauce where it was—then, assuming nothing else changed, the hot sauce company would still keep operating since the neighbors would not be willing to pay the amount necessary to get the hot sauce company to close shop or move.

The point is that, given the value of the hot sauce company location relative to the displeasure of its neighbors, the end-result will be the same from the perspective of efficiency, regardless of who a court of law finds liable. Parties will bargain and compensate others to the extent that they are willing; legal assignment of liability will only be conclusive insofar as it approximates the assignment of payment that would have been reached through a bargaining procedure. From this perspective, it is not obvious what the state is contributing by intervening. Here, the normative payoff is that the existence of negative externalities produced by markets is not in fact so negative. What look like externalities are actually the optimal allocation of costs that is produced by a market for ownership rights.

This is a slight but important perversion of Coase’s argument, because Coase was actually more concerned with showing the pervasiveness of transaction costs. Still, understanding this interpretation of Coase helps us to understand something about the Chicago approach to the corporation; it is based on this strand of “Coasian” thought, and is a deliberate departure from earlier Coasian ideas about the
nature of the firm. Coase’s “The Nature of the Firm”\textsuperscript{11} contended that the efficiency of the firm was due to its difference from the market—it’s supersession of the price mechanism in favor of entrepreneurial planning. As we shall see, the later generation of Chicago scholars sought to show that the firm was itself nothing but a market. Thinking about Chicago in terms of the “two Coases”\textsuperscript{12} – 1937 and 1960 – where Chicago scholars emphasized the latter and read the former through that lens, is helpful for understanding the broader development of the theory of the corporation. The problems of market failure that Coase’s 1937 article articulated were washed away by Chicago scholars using his 1960 article. They took the Coasian idea that the firm is an efficiency-improvement over private contracting, and turned it around by asserting that there is no distinction between the firm and private contracting. If the firm is an improvement in efficiency, then it must be an example of price theory at work—“What is optimal is market-based.”

However, Chicagoans did not generally take their analyses as being a departure from Coase, since they saw Coase as the Godfather of their style of inquiry. Instead, the target of Chicago analyses was what is known as managerialism.

**Critical Managerialism**

*Corporate Power and the Chicago School*

The most famous accomplishment of the Chicago school of economics is likely the development of the field of law and economics—so much so, that law and economics is sometimes used as a synonym for the Chicago school. The law and economics tradition, developed in the 1950s and 1960s, uses the tools and assumptions about markets and prices noted above as a way to analyze,


and normatively assess, the law. If the economics department at the University of Chicago made price
theory popular, it was in the law school where its applicability was first truly tested.

In this regard, it is perhaps telling that what is roundly considered a pivotal moment for the
development of the Chicago school was the hiring of Aaron Director (who was trained as an
economist) to teach a course on antitrust law as a member of the law school faculty. Antitrust is the
specific domain of law where Chicago price theory cut its teeth on legal analysis. Medema has referred
to antitrust law as the quintessential Chicago project, both in terms of illustrating its mission, and in
terms of its success in influencing public policy.\(^\text{13}\) The first issue of the *Journal of Law and Economics*
featured a criticism of the influential *Standard Oil* decision, arguing that Standard Oil’s monopoly had
not resulted from predatory pricing, but through buying out its rivals.\(^\text{14}\) George Priest amusingly recalls
joining the faculty at the University of Chicago Law School, and, upon receiving a cool reception from
Ronald Coase to his work in antitrust scholarship, learning of the mistakes he had made:

> First, I had not focused on puncturing any prominent Supreme Court antitrust opinions…Second, what I had
> concluded was that the Supreme Court decisions in the cases I looked at were mistaken because they
> exonerated practices that were probably anti-competitive. But this conclusion got the exercise entirely
> backwards. Director and Coase wanted to show, through the careful study of prominent cases, that the
> antitrust prosecutions were idiotic and, ultimately, pernicious interferences with the operation of the market.\(^\text{15}\)

Chicago law and economics was born out of a rejection of prevailing legal and economic approaches to
the corporation. For anti-trust policy, the prevailing approach was what is known as the Harvard School
of industrial organization. For the theory of corporate governance, corporate law, and business ethics,
the prevailing approach to the corporation was managerialism, exemplified by Berle, Means, Galbraith,
and Keynes.

\(^{13}\) Steven Medema. 2010. “Chicago Law and Economics” in *The Elgar Companion to the Chicago School of Economics.*


\(^{15}\) George Priest. 2010. “The Limits of Antitrust and the Chicago School tradition.” *Journal of Competition Law and
While the Harvard school and managerialism are distinct in many ways—the Harvard school was not terribly concerned with the internal dynamics of the corporation, and managerialism took monopoly to be largely a solved problem—they represented similar targets for the Chicago school because of their common emphasis on corporate power. For the Harvard School, a decentralized competitive market was meant to check private power; the large corporation compromised this function and therefore needed to be regulated and checked. Pre-Chicago Harvard antitrust scholarship, therefore, saw curtailing the size of the corporation as being in service of decentralizing power across a competitive market. In a similar vein, pre-Chicago managerial approaches to the corporation saw the internal dynamics of the corporation as being largely about the nature and use of corporate power to direct its various resources (human, financial, and otherwise). By understanding how the managerialist view of the corporation cashed out in a normative account of how the corporation ought to be organized, we can better understand the response of the Chicago school.

Managerialism, ownership, and control

Managerialism refers to theoretical approaches that highlight the structural nature of corporations that empower managers, giving them control over other members of the corporation. Managerialism might be thought of as a largely ontological position that starts from the assumption that power is one of the more important facts of the economy. The most famous example of this methodological emphasis on power in 20th century economics is to be found in the work of Galbraith in

16 Legally this found grounding in Learned Hand's famous interpretation of the Sherman Act in his *US. v. Aluminum Company of America* decision. The legal question was not whether a monopoly was achieved for good or bad reasons; instead, because of market power's “indirect social or moral effect,” Sherman was to be interpreted “to prefer a system of small producers, each dependent for his success upon his own skill and character, to one in which the great mass of those engaged must accept the direction of a few.” (Learned Hand. 1945. *United States v. Aluminum Company of America*. 148 F.2d 416)

17 Or as Kaysen put it: “the chief virtue of a competitive market in practice is not necessarily that it leads to economic efficiency but that it constrains private economic power…. In the evolving giant corporation, managers possess great scope for decision making unconstrained by market forces—nowhere more so than in their decisions with respect to future growth and change. (Carl Kaysen.1957. “The Social Significance of the Modern Corporation.” *The American Economic Review* 47[2]: 316 [311-319]).
his idea of “countervailing power”—where the concentration of power in industry spurs on the concentration of power in response (for example, in unions). From these stipulated facts about the corporations—that it is a structure and that managers possess power within it—there of course arose many variants. Some managerialists construed these facts in a positive light, contending that managerial power was justified by managers’ expertise and merit. The more influential strand of managerialism was critical, contending that managerial power was illegitimate and harmful on its face.

The locus classicus for critical managerialism is Berle and Means’ *The Modern Corporation and Private Property*. While sometimes thought of as simply a critique of contemporary corporate practice, Berle and Means’ project is better seen in more grandiose terms: their argument is that the development of the corporation represents a radical and new system of property and production. Sounding remarkably similar to Marx, Berle and Means contend that the development of the factory system in the industrial revolution, which placed large amounts of labor under the direction of a few managers, has now been coupled with the corporate structure that places large amounts of capital under the direction of the same few people. This second movement has effected a large change in the nature of property: first, in that the emphasis on corporate securities markets has made liquid property more abundant than non-liquid or fixed property, and second, as a result, property itself has been decomposed into “nominal ownership” and control. With ownership and control utterly severed (“the

---

18 Less well-known is Galbraith’s concern for the type of power that was to be found within the modern corporation. To take one example, in discussing the growth in size and complexity of the modern corporation, Galbraith argues that “power passed beyond the intellectual reach of the nonparticipant and thus beyond his or her capacity to intervene effectively. And increasingly within the enterprise, decisions emerged not from the single competence of any one individual but from the several contributions of specialists meeting in committee or close daily association.” (John Kenneth Galbraith. 1983. *The Anatomy of Power*. Boston: Houghton Mifflin, 134.)


21 Ibid, 249-252.
dissolution of the atom of property”) the corporation takes on a “quasi-public” status, resulting in a transformation of the economic system:

This dissolution of the atom of property destroys the very foundation on which the economic order of the past three centuries has rested. Private enterprise, which has molded economic life since the close of the middle ages, has been rooted in the institution of private property. … In the quasi-public corporation, such an assumption no longer holds. As we have seen, it is no longer the individual himself who uses his wealth. Those in control of that wealth, and therefore in a position to secure industrial efficiency and produce profits, are no longer, as owners, entitled to the bulk of such profits. Those who control the destinies of the typical modern corporation own so insignificant a fraction of the company’s stock that the returns from running the corporation profitably accrue to them in only a very minor degree. The stockholders, on the other hand, to whom the profits of the corporation go, cannot be motivated by those profits to a more efficient use of the property, since they have surrendered all disposition of it to those in control of the enterprise. … [this] destroys the basis of the old assumption that the quest for profits will spur the owner of industrial property to its effective use. It consequently challenges the fundamental economic principle of individual initiative in industrial enterprise.22

The corporation’s dominant status in American capitalism therefore represents nothing short of a revolution in political economy.

Although the language used to describe this transformation is quite clearly negative, the basis of the critique is not as straightforward. At times it appears that Berle and Means are concerned with the corporation as a new institution inimical to market efficiency. That is, if we assume that Adam Smith’s invisible hand works to transform the individual’s pursuit of private gain into a contribution toward the public good, then the corporate revolution’s effect on efficiency must be substantial. Because owning property no longer implies control of property, the owner of property cannot be assumed to direct the use of that property to his benefit. The power that has fallen into the hands of a few corporate managers is therefore problematic on the basis of one’s concern for the efficient allocation of resources.

While there are numerous parts of Berle and Means that dwell on the inefficiencies of the corporate system, this is not their only concern, nor is it their strongest claim: as we have seen in the previous chapter, the Coase-Williamson account of the corporation explains why a hierarchical structure like the corporation might still be efficient despite deviating from the market’s invisible hand. More compelling is Berle and Means’ critique of the corporation on behalf of normative political and

22 Ibid, 8-9
legal ideas. This critique is implied by the political language used, where corporate managers are the new “princes” of industry, and corporations are new “economic empires.” There are also many references to how this change in property results in a change in the property-holder, destroying the virtue, responsibility, and community that comes with the ownership and control of real fixed property. These suggest the more potent criticism that corporate and managerial power is problematic because of its illegitimate, illiberal, and immoral nature.

For Berle and Means, the modern publicly traded corporation represents a significant change in the way control over property works. On their account one can locate the group in control by locating the group that has the power to select directors (much as one sees the public as sovereign in a republic by virtue of their ability to elect government officials). In the closed corporation the owning party has full power to direct, or hire directors of, the enterprise as it sees fit and can therefore be seen as the locus of control. In cases of majority control, while the party with the most shares still has great control, certain unilateral powers are ceded, like the power to alter or suspend the corporate charter, or disband the enterprise, all of which generally involve some form of supermajoritarian procedure. These are largely in line with the older, traditional “atom of property” upon which classical economics is based. The prevalence of “minority control” and “management control” in the modern corporation represent crucial developments for Berle and Means because they are forms of de facto control; the legal right to control is rendered nominal and another party holds factual control. Berle and Means’ example of minority control is Rockefeller, who controlled Standard Oil despite owning less than 15% of the stock; by virtue of being the largest shareholder, Rockefeller was able to exert an influence over management as if he was a majority-owner because he was able to collect the proxies of other

---

23 Ibid 114, 116.
24 See, as one example, Ibid 64-65: “The spiritual values that formerly went with ownership have been separated from it. Physical property capable of being shaped by the owner could bring to him direct satisfaction apart from the income it yielded in more concrete form. It represented an extension of his own personality. With the corporate revolution, this quality has been lost to the property owner much as it has been lost to the worker through the industrial revolution.”
shareholders that, in tandem with his substantial share of votes, gave him working control of the enterprise.\footnote{Ibid 75-76} Anachronistically, we can make the Downsian point that this results from the fact that the cost of voting tends to outweigh the advantages of voting when the electorate is sufficiently large\footnote{Anthony Downs. 1957. \textit{An Economic Theory of Democracy}. New York: Harper and Row.}; the result is that the party possessing enough resources to solve the collective action problem and mobilize voters will tend to have disproportionate influence by virtue of commanding not just its votes, but the votes of all those it has mobilized. Minority control of the corporation simply occurs when this collective action problem is solved by a minority-owner.

The more pervasive and significant tendency, according to Berle and Means, is for this collective action problem to be solved by the managers themselves, which puts management in control of the corporation. This is done by the creation of a proxy committee, which collects the proxies of willing shareholders, essentially transferring the voting control over management to management itself.\footnote{Ibid 82.} Management control is therefore the result of the corporation growing in size to the point where, though every investor has the legal right to dictate control over the enterprise, this right is hollow in practice. Management assumes control over itself because it alone possesses the power to mobilize and command proxy votes. It is worth noting that Berle and Means were not able to anticipate how the growth of institutional investors would exacerbate this problem. Because institutional investors—like banks, hedge funds, and pension funds—own a large percentage of stock in general, and own this stock in many corporations, the cost of voting becomes even greater, since a share in one corporation is only a fraction of the fund’s diverse holdings.\footnote{See Ronald Davis. 2009. \textit{Pension Fund Democracy}. Vancouver: UBC Press. I will have more to say about institutional investors in chapter 6.}

Those who have a right to control in law are, in the new corporate system, divorced from this right in practice. This raises two normative problems. The first is that the corporation has come to
undermine liberal individualism. The other normative problem is that this rendered the control of the corporation illegitimate: “the recognition that industry has come to be dominated by these economic autocrats must bring with it a realization of the hollowness of the familiar statement that economic enterprise in America is a matter of individual initiative…Their activity is group activity on a scale so large that the individual…has dropped into relative insignificance.” This is, in many ways, just a negative reformulation of what Marx lauded in the joint-stock company, namely that capital becomes socialized, just as labor does. The other normative problem (with which Marx did not and could not concern himself) was that this rendered the control of the corporation illegitimate. As Berle describes elsewhere, the wielding of (and subjection to) power is always accompanied by an implicit or explicit claim that “power lies there because the holder is entitled to it by some test or standard”; as a corollary, this claim to legitimacy also implies that “the holder can be deprived of it if demonstration is made that there is no title or right in his possession of it.” Because the test or standard that gives corporate power its legitimacy is the shareholder vote, and because the shareholder vote has been reduced to the self-selection of management, management rule is illegitimate. The shareholder mandate has been rendered a ritual, akin to royal coronation, since “in both cases the ceremony is intended to induce general acceptance that the power has been well and truly located in the specific individual or individuals involved.” As the shareholder vote becomes more and more nominal, the mandate and legitimacy of management becomes more and more hollow.

Berle and Means consider the possibility, however, that even if those wielding the power might not be doing so legitimately, they might still be pursuing legitimate ends. Phrased differently, there is a distinction between the legitimate possession of corporate power and the legitimate use of such

29 Berle and Means 116.
32 Ibid 104.
The use of power is legitimate if it is “appropriate to the function of the organization,” which for the business corporation is “the ratable benefit of all the shareholders as their interest appears”; corporate power, no matter how legitimate its possession, is subject to legal limitation “when the power has been exercised to the detriment of [shareholders’] interests.” 34 This is essentially an assertion of the familiar idea of fiduciary responsibility, where managers must act in trust for the benefit of their shareholders. When their actions are not in the interests of their shareholders, they are in violation of that trust and, therefore, in violation of corporate law. Berle and Means read fiduciary duty in thick terms, asserting that such legal limitations hold even if there is language in the corporate charter that might protect particular actions; to do otherwise would be “to defeat the very object and nature of the corporation itself.” 35 The legal purpose of the corporation restricts the ends of the corporate enterprise, restricting the kinds of actions corporate managers can pursue legitimately.

A key part of the managerial critique is that these legal doctrines are not enforced by courts, giving a large amount of latitude to managerial discretion; the legal enforcement of fiduciary duty has gone the way of shareholder control, nominal holdovers from a pre-corporate economic system that no longer have any real force or effect on economic activity. The modern shareholder has “surrendered a set of definite rights for a set of indefinite expectations,” and finds herself limited in the kinds of legal demands she can make on the corporation. 36 Berle and Means put this critique in political terms, drawing on an analogy between the corporate enterprise and communist government. This is a familiar trope that can be found in other institutionalist and managerialist writing. 37 Berle and Means articulate

---

33 Ibid 99; Berle and Means, 221.
34 Berle 102; Berle and Means 221
36 Ibid 244.
37 Again, the most famous example is Galbraith, whose theory of convergence contended that capitalist and socialist economies were converging toward a common form of bureaucratic planning. For example: “There is no tendency for the Soviet and Western systems to convergence by the return of the former to the market. Both have outgrown that. There is measurable convergence to the same form of planning” (John Kenneth Galbraith. 1967 (2007). The New Industrial State. Princeton: Princeton University Press. 136).
this with reference to the subordination of the individual to the collective in both systems of production:

A sovereign can and does subordinate the interest of the individual to its own purpose, though its power to do so may be limited by self-denying ordinances such as are contained in the bill of rights in the American constitution. The peculiarity of the corporate form is that it subjects economic rights, heretofore known as property rights, to such exigencies in a peculiar and drastic degree and for far more limited ends….It is an odd paradox that a corporate board of directors and a communist committee of commissars should so nearly meet in a common contention. The communist thinks of the community in terms of a state; the corporation director think of it in terms of an enterprise; and though this difference between the two may well lead to a radical divergence in results, it still remains true that the corporation director who would subordinate the interests of the individual stockholder to those of the group more nearly resembles the communist in mode of thought than he does the protagonist of private property.  

These political analogies are instructive. The shareholders are to the corporation what the public is to the state (the source of legitimacy), and corporate law is to management what a bill of rights is to the state (protections of those subject to rule); as a result, shareholders without the protection of corporate law are to the corporation what citizens of communist countries are to the state (subjects of illiberal rule). The fact of corporate law’s impotence renders the corporation not merely an instance of the illegitimate possession of power, but the development of an illiberal and anti-individualist institution. 

This failure of the corporation to abide by the principles of legitimacy has rendered it incompatible with the liberal ideas of competition and private enterprise. With this, Berle and Means conclude that liberal-democratic ideas of individuality, legitimacy, and equality are no longer apt for the modern corporation. The corporation is now, Berle and Means famously conclude, a social institution, and must be recognized as such. No longer an instrument of shareholders, the corporate purpose must now consider the “well-being of those who are subject to the organization, whether workers, investors, or consumers.” Similarly, because shareholders can no longer be plausibly viewed as having control, the corporation ought to be controlled by society until it becomes “a purely neutral technocracy, balancing a variety of claims by various groups in the community and assigning to each a

38 Berle and Means, 245.
39 Ibid 310.
portion of the income stream on the basis of public policy rather than private cupidity.”⁴⁰ As the corporate institution becomes more severed from the shareholders who formally own it, it ought to pass further into the realm of public accountability.

While this includes formal institutional and legal mechanisms of control and constraint, Berle would later highlight an important form of informal control, which he misleadingly refers to as “public consensus,” but might be better understood as the learned consensus. By being able to influence the “corporate consciousness,” independent scholars, journalists, and intellectuals function as the “real tribunal to which the American system is finally accountable” by virtue of their appeal to principle and independence from corporate interests.⁴¹ Berle and Means here seem overly sanguine about the potential for discursive control over corporate activity from what we now think of as the public sphere. While we have seen little reason to believe that corporate managers are “finally accountable” to these members of the chattering class, it is worth pointing out that this is essentially a precursor to what we now think of as business ethics and corporate social responsibility. No longer a creature of shareholder endeavor, the corporation’s social and ethical dimensions should not be determined by the internal debates of a particular constituency’s subjective interests, but must instead be the subject of a broader debate grounded in the objective principles of society at large. In this regard, those who occupy important and influential places within ongoing public debates also, for Berle and Means, have an important role in affecting and controlling corporate practice.

Managerialism, the Corporation, and Political Theory

More can be said about critical managerialism in general, and about Berle and Means’ project in particular. However, we can conclude here by noting that Berle and Means take managerialism’s

---

⁴⁰ Ibid 312-313.
⁴¹ Berle, 90-91, 113, 115.
empirical claims—the corporation is a structure that vests hierarchical power within a select group of managers—and join them with familiar ideas of legitimacy, thereby providing answers to the three normative questions posed in the first chapter—the question of corporate governance (who ought to govern on whose behalf?), the question of corporate law (what ought to be the legal constraints of this governance?), and the question of business ethics (by what standards ought corporate decision-making be appraised?)—both according to how such questions should be answered for a liberal society prior to the development of managerial control (the “liberal” regime), and how they should now be answered in the context of the new corporate system (the “corporate” regime). Their answers are shown in Table 3.1 below.

Table 3.1 Berle and Means’ Normative Component

<table>
<thead>
<tr>
<th></th>
<th>Question of Corporate Governance</th>
<th>Question of Corporate Law</th>
<th>Question of Business Ethics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Liberal</strong></td>
<td>Shareholder-controlled manager-governance</td>
<td>Fiduciary duty to serve shareholder interests</td>
<td>Orientation toward shareholder interest</td>
</tr>
<tr>
<td><strong>Corporate</strong></td>
<td>Formal and informal social control</td>
<td>Fiduciary duty to serve the social benefit</td>
<td>Social responsibility to serve the public</td>
</tr>
</tbody>
</table>

Because the corporation is no longer controlled by those who own, but rather by those who act in unaccountable ways, the normative prescription is for statutory and public governance of the corporation. The legal rights of shareholders are no longer effective or even existent. The normative prescription is to bring corporate law within the domain of public law, as this better reflects the individual shareholder’s relationship with and vulnerability to the large enterprise. Although Berle and Means appear to be advocating public regulation from the perspective of business investors, their impact went way beyond the business community. For instance, their study was taken up by economically sophisticated socialists, particularly Abba Lerner and Oskar Lange. While best known for
their model of market socialism, which showed that a competitive allocating price mechanism was compatible with public ownership of the economy, Lerner and Lange’s argument also made the contention that the inefficiencies of bureaucracy were not unique to socialist economies but, citing The Modern Corporation, quite present in capitalist firms as well. Thus, even if we take Berle and Means as offering a criticism of modern corporate inefficiency, the fact of this inefficiency has potentially radical results.

But as I have tried to show, efficiency was not their singular or primary concern. Indeed, Berle and Means are perhaps best understood as being a part of a pluralist tradition of social criticism encompassing thinkers like Harold Laski and John Dewey. Derived from Hegelian ideas about the ethical importance of associational life, the pluralist tradition argued that there is no single entity possessing political sovereignty and that political norms are rooted “in a variety of forms of human association not confined to the state”; the state has always been but one powerful institution among many competing for power and control over social functions, and thus to confine our political theories to the state is to miss important categories of political life. Read in this light, Berle and Means’ main concern is not simply the “separation of ownership from control” in the corporation, but the fact of growing corporate power, which was taken to be a threat to a free society. In an article on freedom of speech, Berle cites Hegel’s philosophy of freedom and argues that modern corporate power stifles the

---

43 Lange 1938, 110.
44 This is not to be confused with the “pluralism” associated with American political scientist Robert Dahl. From my understanding the two traditions are utterly distinct, if not downright antagonistic toward one another. For Dahlian pluralists, there is presumed to be a seat of power, and the pluralism exists in the different groups who compete for that power. For pluralists like Laski and Tawney, the pluralism exists in the various seats of power within society. Sartori comes to a similar distinction in his discussion of pluralism. (Giovanni Sartori. 1997. “Understanding Pluralism. The Journal of Democracy 8(4): 58-69.)
ability of people to use rights in a way that might allow for the flourishing of “personality.”47 This fits into a thread running through pluralist and progressive thinking of the time, where the corporation was considered a central concern for political theory.48 The rise of the business corporation fit into a long history of social power being held by meso-level institutions. This also helps us understand why Berle and Means contend that the modern corporation had to be addressed by social control; it was not enough for businesspeople to attempt to address this from within the corporation, since the problem was in the corporation itself, given its size and reach. Like the state, it was a social institution exerting power over the norms that shape political society, and therefore was necessarily subject to social control. While perhaps appearing as an appeal to investors and denizens of the business-world, Berle and Means’ project was actually fundamentally about placing economic institutions like the corporation within the purview of political theory and social policy.

Berle and Means’ argument should be understood as a manifestation of a larger tradition of critical thought that presented a significant challenge to both advocates of business management and free marketers. The continuity that pluralists saw between the modern corporation and pre-liberal feudal fractured sovereignty was taken by critical managerialists and claimed to represent a transformation in the liberal capitalist system of property as a whole. Classical liberal norms and institutions, therefore, were no longer appropriate for organizing modern society. As a result, the gauntlet was thrown down for those who wished to adhere to such classical liberal principles: how could the modern business corporation—powerful and unaccountable as it was painted to be—be reconciled with the libertarian emphasis on individualism and non-interference? How could the

48 Take this example from Laski: “Corporations [have] a curious habit of attempting perpetually to escape from the rigid bonds in which they have been encased. May we not say that, like some Frankenstein, they show ingratitude to their creators?...A corporation will possess itself of an empire, and resent interference with its domain. ...Truly the supposed sovereignty of the state is not apparent in the relations thus discovered.” (Harold Laski. 1916. “The Personality of Associations.” Harvard Law Review 29: 407 (404-428).
corporation be conceived as anything other than a social institution that overcame and dominated individuals? For libertarianism or classical liberalism to have any coherence, the problem of the corporation and its power had to be resolved.

**Chicago Business Ethics: Friedman’s Opening Salvo**

Friedman’s classic 1970 article, “The Social Responsibility of Business is to Increase its Profits” (and the chapter in his famous 1962 book, *Capitalism and Freedom*, upon which it is based), is presented as a refutation of the concept of “corporate social responsibility.” Since then the idea of corporate social responsibility (CSR) has grown in prominence, complexity, and controversy, taking on meanings and debates that were not familiar to midcentury scholars. CSR—the idea that corporations have some ethical duty to engage in actions that directly and explicitly benefit society in some way—has a direct link to Berle and Means’ claim that the corporation is now a social institution, beholden and responsible to the society that creates it. While Friedman is certainly critical of CSR—and, by extension, of Berle and Means’ normative conclusions—focusing on this misses the more radical challenge that Friedman makes to the managerial view of the corporation.

The underlying claim that Friedman asserts—since he does not really offer an argument for it in the 1962 or 1970 pieces—is that the very idea that the corporation is a “structure” is nonsense. For Berle and Means the modern corporation represented the dissolution of liberal and individualistic ideas of property, in which to own means the ability to control, and therefore created new forms of power. For Friedman, this development was not anything new or revolutionary. The corporation was not a new system of property, but simply an arrangement where some people worked for others. Put simply, it is a principal-agent relationship. The fact that shareholders “own” and managers “control” the

---

corporation’s property is not laden with the normative implications managerialists want to place on it, since all sorts of relationships involve people acting on behalf of others, none of which imply some revolutionary system of political economy.

Friedman’s claim that the corporation is governed by a principal-agent relationship need not inherently be so different from the managerialists. The fact of managerial control and corporate power could be conceived of as a broken-down principal-agent relationship where the agent was now able to exploit and defraud the principal. Against this, Friedman simply asserts the contrary: “a major complaint against modern business is that it involves the separation of ownership and control—that the corporation has become a social institution that is a law unto itself, with irresponsible executives who do not serve the interests of their stockholders. This charge is not true.”

To understand properly Friedman’s provocative normative conclusions—that business has the sole ethical mandate to “use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game”—it helps to try and figure out Friedman’s reasoning (since he does not actually defend this position) for rejecting the Berle and Means thesis that corporate managers had become unaccountable to the shareholders whose interests they represent. Put differently: why was Friedman qua economist not concerned about the possibility of principal-agent problems, given the incentives and opportunities for corporate managers to act in their own interests, and not in their shareholders’?

One explanation is methodological. In one of the most influential essays in the history of economics, Friedman made the claim that the assumptions underlying a theoretical model need not be realistic or true, as long as they lend themselves to effective predictions. So, to recite his famous example, one might try to explain the density of leaves on a tree from the theory that leaves act as if they were deliberately trying to gain as much sunlight as possible. While the assumption underlying

---

51 Friedman, 1962, 133.
this theory is patently false since leaves don’t do anything deliberately, the predictions it leads to are generally confirmed by experience and investigation: leaves are crowded more densely on the southern face of trees, and tend to be less crowded on the north or in sunlight-obstructed areas. Because the latter is the class of phenomena that is meant to be explained, it is of little consequence that the assumptions are unrealistic.

This is significant for our purposes because Friedman segues from this example of leaves into the question of the profit-maximizing firm: is the model of microeconomics flawed because it assumes that all firms seek to maximize profits? The answer depends on how useful the assumption is for predictive purposes: “unless the behavior of businessmen in some way or other approximated behavior consistent with the maximization of returns, it seems unlikely that they would remain in business for long….given ‘natural selection,’ acceptance of the hypothesis can be based largely on the judgment that it summarizes appropriately the conditions for survival.” 53 Empirically speaking, we don’t need to know anything about the actual determinants or character of managerial decision-making. The only thing we need to know is that, given the context of a competitive market, businesses will tend to behave as if their managers were consciously and singularly pursuing the maximization of profit; put differently, making such an assumption is justified because doing so allows for the most accurate predictions of business behavior.

So, why is it false to claim that the separation of ownership from control has resulted in an overly-brazen corporate management? It is because doing so assumes such a separation matters in the first place. The actual micro-determinants of a given manager’s decisions are less important than the fact that managers will tend towards profit-maximization by virtue of the pressures that markets supply. That is, the fact that owners are not able to exert any sort of “control” over their managers does not mean that their interests are not being served by their managers; if we assume, along with Berle and

53 Ibid 158.
Means, that the stockholders’ interests lie in the maximization of returns, then they don’t need an effective vote to assure this. The “natural selection” of the market, according to Friedman, does this just fine.

This raises a question: if the market will tend to push executives toward shareholder interests in the first place, then why would the firm have a “social responsibility” to maximize its profits? Wouldn’t this be less of a responsibility and more of an inescapable fact? The reason why it has ethical connotations for Friedman is because doing otherwise will tend to undermine the competitive free market as a whole if pursued by all. This is why he argues that the businessmen who claim to be concerned with social ends in addition to profit—such as “providing employment, eliminating discrimination, and avoiding pollution”—are not actually bolstering the standing of free enterprise, but rather are “preaching pure and unadulterated socialism” as the “unwitting puppets of the intellectual forces that have been undermining the basis of a free society these past decades.” The competitive and free market is not simply a system of property rights and enforcements, but a mutual recognition that all have the right to pursue their own interests to the best of their legal ability. So, although Friedman wants to claim that given a market context firms will tend to maximize profits, the growing “corporate consensus” that firms have “social obligations” has the effect of thwarting the very existence of the free market, and is therefore “a step in the direction of creating a true divorce between ownership and control and of undermining the basic nature and character of our society.” The result is that shareholders will be robbed of the main institutional check that keeps managers working in their interest.

To better explain this idea we can start with the assumption that the manager is the agent in two distinct principal-agent relationships—with the shareholder on one hand and with the stakeholder (our

---

54 Friedman, 1970, 253.
55 Friedman 1962, 136.
stand-in for “social concerns”) on the other. One might read Friedman as claiming that to focus on the second relationship is to abandon one’s duties in the first. For an executive to pursue “social responsibility,” on Friedman’s account is simply to take someone’s money and spend it on a purpose they did not intend. In this sense, to pursue something other than profit-maximization would be socially irresponsible because it violates the moral idea that one’s property ought to be used as one wishes. The best check against such exploitation of this relationship is the free market, which is fundamentally about unanimity: “no individual can coerce any other, all cooperation is voluntary, all parties to such cooperation benefit or they need not participate.” If the executive does something “socially responsible,” like make a donation to a charity from corporate coffers, she “prevents the individual stockholder from himself deciding how he should dispose of his funds.” To pursue CSR is to violate the unanimous form of cooperation that the market produces.

This actually brings us to a larger and stronger thesis, which is what Friedman appears to mean: emphasizing one’s principal-agent relationship with the stakeholder over the shareholder has the effect of shirking responsibilities to both relationships! While engaging in CSR has the apparent effect of lessening a manager’s commitment to shareholder interests, Friedman argues that doing so also harms the social stakeholder in two ways: first, because engaging in this type activity trespasses on what elected officials are delegated to do; and second, because social interests are generally best served by the invisible hand of the market. Citing Smith and following Hayek’s account of the market’s virtues, Friedman contends that the conscious pursuit of the social good tends to lead to bad results, and is based on a number of flawed assumptions: first, it assumes that there is such thing as a “social good”;

---

57 Friedman 1970, 259.
58 Friedman 1962, 135.
59 Friedrich Hayek, “The Use of Knowledge in Society”
and second, even if such a good exists, pursuing CSR assumes that any one person or group of people is capable of knowing what it is. The market’s method of cooperation solves these problems, by substituting for the “social good” those ends that are produced by the intersection of various individual aims, and by pursuing it not through individual initiative but through the price mechanism’s spontaneous organization. The payoff is that market actors deliberately pursuing social ends are less likely to benefit society than those who simply seek to maximize their own interests.

For Friedman the manager’s pursuit of corporate profit is doubly ethical: it is ethical because it is in keeping with what he is hired to do and it is ethical because doing so will tend to benefit society. Similarly, the claim that the corporation is a social institution where managerial authority rules the day is wrongheaded for at least two reasons, both of which might be thought as enabling the very ills that critical managerialism is meant to address. On the one hand, thinking of the corporation as a social institution that ought to pursue social ends has the effect of undermining the market’s ability to orient managerial action to shareholder interest. Advocates of CSR concerned with managerial unaccountability are, in effect, creating the very separation of ownership and control they despise. On the other hand, thinking of the corporation as a social institution undercuts the best mechanism for putting resources toward their socially optimal end.

**Agency Theory and Property Rights Theory**

Taking the time to engage the ethical logic of Friedman’s claim is worthwhile as it brings into stark relief the normative nature of the economic theory of the firm developed by the next generation of Chicago economists. After all, while the ethical claim that Friedman makes has some bite to it, it rests on the assertion that the price mechanism works as he says it does, even when it comes to the corporation. But this is not such a simple assertion; as Coase and Williamson argued, the corporation represents precisely the absence of such a price mechanism. Friedman’s confidence in the market makes no sense if the corporation is characterized by a set of relationships that are *not* directed by the
market’s price mechanism. What was needed was a way of taking the key Coasian insight that the development and expansion of the firm is a function of efficiency, and divorcing it from its basic analytic distinction between activity that happens in the market outside the firm and that which happens within it. Agency theory and property rights theory, the two key developments in the theory of the firm in the 70s and 80s, can be understood as providing the conceptual basis that would allow partisans of the free market to obviate such a distinction. This would allow them to cash in on Friedman’s promissory note regarding the ethical nature of managerial profit-maximization, and transform the economic theory of the firm’s efficiency into a normative theory justifying the governance and legal structure of the corporation.

In this regard it is interesting that, despite their alignment with Friedman’s overarching aim, Jensen and Meckling60 actually reject a core feature of Friedman’s argument in their famous contribution to the subject. Specifically, they reject the idea that the supposed separation of ownership and control is eliminated due to the disciplining effects of competitive markets. Competition amongst many firms cannot be a force for eliminating the agency problem resulting from the separation of ownership and control, they correctly note, since this separation is characteristic of firms and public corporations generally; if all corporations feature a separation of ownership and control, then one enterprise won’t be forced out of the market by virtue of such separation. Put differently, competitive markets will actually serve to entrench such features, not nullify them. The question is why this is the case.

Instead of rejecting the managerial claim that those who control the firm have great discretionary power relative to the vast majority of owners, agency theorists took a different tack. Jensen and Meckling start by noting that most have approached the question of corporate control from

a normative perspective—asking questions like “what is the best way of taming managerial discretion?”—and then contend that theirs is a positive theory: “we assume individuals solve these normative problems…we investigate the incentives faced by each of the parties and the elements entering into the determination of the equilibrium contractual form characterizing the relationship between the manager of the firm and the outside equity and debt holders.”  

The positive nature of the theory is purchased on the assumption that the market effectively produces the optimal result; it follows that the firm, as it exists, is the optimal result of freely contracting individuals, leaving merely the positive question of which incentives lead individuals to choose such results.

Jensen and Meckling make this clear in a footnote, when discussing the compatibility of the non-maximizing firm with Pareto optimality:

> If we could establish the existence of a feasible set of alternative institutional arrangements which would yield net benefits from the reduction of these costs we could legitimately conclude the agency relationship engendered by the corporation was not Pareto optimal. However, we would then be left with the problem of explaining why these alternative institutional arrangements have not replaced the corporate form of organization.  

The not implausible assumption here is that were there a better option we should expect people to choose it (or expect an explanation for why that choice is structurally or legally blocked). The point, however, is that the theory is presented as positive only because the normative assumptions—that such institutions are good because they are preferable from the perspective of those who participate in them—are baked into the cake of market efficiency. For this reason, where agency theorists and property rights theorists contend that incentives can positively explain the existence of corporate structures as they are, what they are effectively doing is explaining how the existence of particular incentives appears to normatively justify particular corporate structures.

So, why is the separation of ownership and control preferable to other decision-making structures in most corporations? The answer, counterintuitively, is agency costs. This is

---

61 Ibid 309-310.  
62 Ibid 328 n28.
counterintuitive because most—including, as we have seen, Smith, Berle, and Means—have taken the
eexistence of agency costs as being the fundamental problem of the separation of ownership and control
in the first place. Agency theorists flip the script by beginning with a case in which agency costs are
zero, namely a firm that is owned and managed by the same person/small group or, as Fama and Jensen
refer to them, noncomplex organizations.\(^{63}\) In such organizations the principal and the agent are the
same actor, and so there is no divergence in interests and, therefore, no agency costs. Agency costs
arise when this actor needs more capital than that which is currently possessed, requiring that people
contribute funds—either through debt or equity, though here the significant instrument is equity—and
therefore turning the manager-owner into an agent who now works on behalf of various agents.

The agency relationship, put in this way, has an intuitive appeal, in that it allows someone to
become a decision-maker of an enterprise even if she is not endowed with the large sums of money
necessary to make her a manager-owner. Similarly, it allows for people with excess money to invest, to
do so without having to take on the task of management. What agency theorists contributed here was
the observation that various kinds of costs result from this type of relationship. Whereas Smith, Berle,
Means, and others focus primarily on the divergence of interests between the principal and agent,
agency theorists noted that agency costs entail not simply this divergence of interests but the costs
incurred in trying to mitigate this divergence. These costs—both of money and non-monetary resources
like time—include the costs of principals to monitor or to control their agents (“monitoring costs”), the
bonding costs that agents take on to give assurance to the principals (“binding costs”), and, finally, the
residual losses imposed on the enterprise by virtue of the principals’ inability to ensure that the agent
will fully act in accordance with the firm’s interests (“agency costs” proper). While the latter at first
appears to be a cost borne solely by the principal-investors, agency theorists contend that worries about
such costs will be taken into account by would-be investors, thereby raising the cost for the manager to

acquire more capital. The inability to fully resolve the principal-agent problem through bonding or monitoring therefore results in a cost for the manager as well as the investors.

The point of all of this is that the principal-agent relationship does not lead to the win-lose transformation Berle and Means imply that it will. Managers incur the cost of bonding and more expensive capital (while still getting the benefit of having capital being made available to them), and investors gain the benefit of equity shares in an enterprise they don’t need to manage (while incurring the cost of monitoring and managerial misdirection). The costs of agency are taken on because they represent improvements on the status quo—entrepreneurs with insufficient capital and investors with uninvested funds. Put differently, agency theorists like Jensen, Fama, and Meckling take the basic feature that concerned Berle and Means but recast it in an economic, as opposed to a political, light; absent is the language of power, and in its place we find a calculation of costs and benefits that determine the optimal arrangement. The power imbalance that Berle and Means bemoaned is thereby transformed into a series of contracts that are entered into because they tend to benefit all involved; property is assessed in this analysis less in terms of rights, and more in terms of the relative benefits they bring. The relationship between manager and investor is not a political one, but a contractual one in which all costs are taken into account: the misalignment of interests between principal and agent are reflected in costly monitoring and bonding mechanisms, and the cost of uncontrollable potential managerial indiscretion is reflected in the heightened cost of capital.

The agency theorist’s response to the Coasian view of the corporation, then, is that the entrepreneur does not really direct the firm. What Coase was actually discussing was a principal-agent relationship between entrepreneur-employer and the employee; employees are hired to do a particular thing in-house because the price mechanism is ill equipped to induce the employee to do so. However, once hired, the entrepreneur is then in a position where he must deal with the fact that there is a divergence of interests between him and his employees; he must therefore expend costs monitoring them or employees will have to bear the costs of depressed wages because of their inability to fully
assure the employer that they will not shirk in their duties. Alchian and Demsetz refer to this as a “metering problem,” and it is precisely the same dilemma faced by manager and shareholder.

The result is a view of the firm transformed from Coase and from Berle and Means, into one that is merely a nexus of contracts. The voluntarism implied by such a view radically recasts the received understanding of the employer-employee relationship, utterly removing any trace of “authority” that one might have thought existed:

It is common to see the firm characterized by the power to settle issues by fiat, by authority, or by disciplinary action superior to that available in the conventional market. This is delusion... It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people. I can "punish" you only by withholding future business or by seeking redress in the courts for any failure to honor our exchange agreement. That is exactly all that any employer can do. He can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty products. What then is the content of the presumed power to manage and assign workers to various tasks?...To speak of managing, directing, or assigning workers to various tasks is a deceptive way of noting that the employer continually is involved in renegotiation of contracts on terms that must be acceptable to both parties. Telling an employee to type this letter rather than to file that document is like my telling a grocer to sell me this brand of tuna rather than that brand of bread.

What agency theorists like Fama, Jensen, and Meckling, and property rights theorists like Alchian and Demsetz, share is the anti-Coasian view that authority in the firm is a chimera. What goes on within the firm is the same thing that goes on outside the firm, namely agents contracting with other agents for goods and services. The only difference is that there is one fictional agent—the corporation—who is a common contractor with all other parties.

This allows for a specialization of knowledge and information that then helps all parties achieve their shared goals. Indeed, instead of seeing the employer as hiring the employee, this leads to the view that what the employer is really doing is selling his information to the employees who “are using him [the employer] to discern superior combinations of inputs.” The principal-agent nature of this is

---


65 Ibid 151-152.

66 Ibid 174.
simply part of the cost of doing business, but does not imply any particularly distinct form of organization or organizing. Again, Alchian and Demsetz are worth quoting in full:

As a consequence of the flow of information to the central party (employer), the firm takes on the characteristic of an efficient market, in that information about the productive characteristics of a large set of specific inputs is now more cheaply available….Conceiving competition as the revelation and exchange of knowledge or information about qualities, potential uses of different inputs in different potential applications indicate that the firm is a device for enhancing competition among sets of input resources as well as a device for more efficiently rewarding the inputs….the firm can be considered a privately owned market; if so, we would consider the firm and the ordinary market as competing types of markets, competition between private proprietary markets and public or communal markets.67

By focusing on the contractual nature of the relationship between the various corporate patron groups, these theorists effectively turn the firm into a market. This allows for the assumption that within the firm something akin to the price mechanism is essentially working as it ought, and the Hayekian conclusion that the organizational results are to be preferred because they reflect the optimal use of knowledge and resources. The separation of ownership and control is merely the efficient arrangement of freely contracting parties, and the optimized use of available information.

Indeed, property rights theorists like Alchian go even further. Whereas agency theorists want to show that principal-agent relationships are both endemic and pervasive, and therefore that the separation of ownership and control is neither special nor scary, property rights theorists want to deny such a separation. The property rights view holds that the most important determinant of how goods and services get allocated in a society is not a result of initial endowments, but of the structure and integrity of the property rights assigned to such endowments. Important to this view is the idea that property rights are a bundle of rights, including the right to use, but also the right to profit from and the right to sell or destroy, a particular piece of property.68 When we talk about a good or service being bought or sold, what we are actually talking about is a trade of some type of right over property X for another right over property Y. Thus the principals and agents—whether we are talking about shareholders and managers, or managers and employees—are actually just holders of different property

67 Ibid 177.
rights. The shareholder owns the right to residual claims on the corporation and to sell her shares, while the manager owns the right to direct the use of corporate property; the manager possesses the right to direct the relationship amongst employees and resources, while employees possess the right to use corporate resources and a right to receive the wages they have contracted for. Far from the separation of ownership and control that obsessed Berle and Means, property rights theorists contend that the firm is simply a bundle of different types of properties, with different parties contracting for different rights over them.

Thus, even divorced from the ability to exercise concrete control of the firm, investor-owners have the exit-option that allows them to shed their shares if their investments cease to return profits. In many ways though, what the stock market provides isn’t just the threat of “exit” but the threat of “entrance”: depressed share prices present the possibility of a hostile takeover, wherein a corporate raider buys up the cheap stock, replaces management, and rearranges the corporation in a different manner in the hope of then reselling shares at a higher price. The existence of securities markets, where the demand for dividend-producing investments is met by investment opportunities competing with one another, becomes the primary vehicle for controlling the corporation on this account. The securities market imposes discipline on a firm’s managers, for if they do not act on investor interests they will find themselves without investment or being taken over and removed. Far from being a problem, the diffusion of ownership allows for an effective finance market to open, which offers the ability to control the firm through the power of the purse. The property rights paradigm argues for two bold theses: first, that managerial power simply does not exist, and second, that there is already an institution capable of directing the corporation, namely the market. The market for securities and the labor market for managers disciplines the corporate manager to act in the interests of the enterprise; his

---

actions could lead to losing out on necessary equity capital, or could lead to him being fired and replaced by a more loyal agent.

This brings us back to Friedman’s initial claims: that the separation of ownership and control does not exist, and that the social responsibility of managers is to pursue corporate profit. Thinkers like Alchian and Demsetz vindicate the first claim by recourse to a more subtle understanding of property rights; the second claim is substantiated by agency theorists and property rights theorists who argue that there is no such thing as managerial power and therefore no such thing as managerial social responsibility. There is merely a series of contracts binding together principals and agents, contracts that not only create a legal duty to uphold the contract, but that also instrumentally serve to allow for an efficient allocation of knowledge, goods, and services in a market at equilibrium. The manager’s duty is to increase the profits of the firm because that is what he has contracted to do. Furthermore as a market actor—since firms are simply markets by another name—following through on the transaction he has contracted to do contributes to a welfare-maximizing, optimal market outcome.

The Nexus-of-Contracts and Corporate Law

Managerialism’s political conception of the corporation—where managerial power presented a social challenge requiring mechanisms of social accountability to solve—is dissolved into a voluntary and contractual relationship capable of being explained solely in terms of efficiency. Friedman’s initial claim, that corporate executives are agents of their shareholders and have a sole ethical duty to pursue profit maximization on their behalf, was taken and generalized by agency theorists and property rights theorists such that all relationships within the corporation could be understood as market transactions; while some act as agents for others, all possess a different set of property rights so as to effect an efficient allocation of rights to use, profit from, and sell different aspects of the corporate enterprise. Power, authority, representation—these are not merely downplayed in favor of an emphasis on efficiency, but precluded by this contractual account.
In the process, this view also distanced itself from the Coase and Williamson view of the firm, and of transaction cost economics more broadly. While one can note a basic distinction between the two in the role of the assumed “utility maximizer” in their economic analyses, the main difference is how each of these schools conceive of the corporation. As we have seen, agency theory and property rights theory sought to erase the distinction between firm and market, and to subsume the corporation within a conception of markets. Whereas the foundational Coasian insight held that the firm is defined primarily by its superseding of the price mechanism, later Chicago economists argued that something approximating the pricing mechanism was still determining the structure and dynamics of governance relationships within the corporation. The corporation, on this view, is transformed into a private market and, following the lessons of the “Coase Theorem,” such a market is presumed to be optimal.

Its identification as positive theory notwithstanding, I have tried to show how this branch of Chicago economics ought to be seen in normative terms, particularly as being in continuity with Friedman’s ethical writings on corporate governance. The normative aspects of this theory were given their most straightforward and concrete articulation by legal scholars, who used this theory to argue for particular practices of corporate law, and therefore principles and guidelines for how corporations ought to be governed and how business executives ought to see their ethical responsibilities. While there are a number of works—running the gamut from journal articles, newspaper editorials, and judicial rulings—that articulate these normative visions, the canonical work, which we might think of as the modern Chicago version of Berle and Means’ book, is Easterbrook and Fischel’s *The Economic Structure of Corporate Law*. Frank Easterbrook in particular is the touchstone for contemporary corporate law, being himself a judge famous for his activist interpretations of statute and precedent in favor of the Chicago economics project.

---

Informed by the foregoing economic theories, Easterbrook and Fischel contend that the corporate form as we know it is the end-result of years of testing and adaptation in the service of efficient organization. This leads them to the view that corporate law should be conceived of as an extension of contract law: since the corporation is composed of voluntary relationships, we assume that these relationships are entered into after explicit or assumed terms have been negotiated, or proposed and accepted. The corporation is a legal fiction that stands for the nexus of contracts with which each corporate party transacts: when one invests in, or works for, the corporation, what one is saying is that they are merely contracting with everyone else who is contracting with “the corporation.”

The result is a principled ambivalence toward how internal relationships within the corporation are organized: “just as there is no right amount of paint in a car, there is no right relation among managers, investors, and other corporate participants. The relation must be worked out one firm at a time.”\(^\text{72}\) The result is that corporate law is not to be seen as prescribing some perfectionist ideal of organization, but as serving to facilitate the most efficient contractual arrangements for a given organization by reducing transaction costs: “the normative thesis…is that corporate law should contain the terms people would have negotiated, were the costs of negotiating at arm’s length for every contingency sufficiently low…it is enabling rather than directive.”\(^\text{73}\) Corporate law, we might say, serves as a store of solved problems, a pre-made set of “off-the-shelf” contractual rules, justified from the perspective of economic efficiency, that save corporate parties from having to draw up contracts each time.

There is a subtle ambiguity here. On the one hand corporate law views the corporation as a set of contracts that people have actually negotiated, explicitly or tacitly. On the other hand, there is an idealized aspect to their economic analysis, where we are to see existing statutes as the terms “people


\(^{73}\) Ibid 15.
would have negotiated.” That is, the corporation is the result of actual contracts but is also structured by principles derived from hypothetical contracts, reminiscent of the social contracts of liberal political philosophy. Without belaboring the point, the hypothetical nature of the contract is important because to imagine a hypothetical contract that hypothetical parties “would” agree to, we must impute certain motives to them. For Easterbrook and Fischel that motive is clear: “Money. They [investors and other participants in the corporation] therefore would agree unanimously to whatever rule maximizes the total value of the firm.”\(^7^4\) This would seem to fit in with the original Friedman argument that the sole duty of the corporation is to maximize profits (assuming, for simplicity’s sake, a simple correlation between profit and value). It would also seem to provide grist for the critics’ mill, who argue that corporate law is responsible for the psychopathic pursuit of profit that we see in modern corporate institutions.\(^7^5\)

How are Easterbrook and Fischel able to assert both the enabling, non-directive quality of corporate law, while imputing this maximand to all parties of the corporate contract? We must note that they assert hypothetical value-maximization, but not as grounds for trumping corporate contracts as we find them. Despite whatever theoretical grounds we might have to believe that a certain form of organization might be more efficient than others, Easterbrook and Fischel claim that there is little basis to be found in corporate law to trump agreed-upon arrangements in service of profit-maximization. In the contractual approach to corporate law, the maximand of corporate value might be said to be presumed as opposed to imputed: the ready-made contractual terms that corporate law provides are furnished from the presumption that agents seek to maximize the value of the corporation, since doing so will tend to reduce agency costs. To posit a more complex maximand that combines various types of objectives would only ensure that none of those objectives are met. As they eloquently put it: “a

\(^7^4\) Ibid 23.
manager told to serve two masters (a little for the equity holders, a little for the community) has been freed of both and is answerable to neither. Faced with a demand from either group, the manager can appeal to the interests of the other.”76 Maximizing value is the presumed maximand both because empirically it tends to be what parties want, and normatively because it tends to lower costs resulting from the corporate form.

However, this does not specify that corporate parties must pursue these terms. In the end, a given corporation can pursue whatever it wishes:

What is the goal of the corporation? Is it profit, and for whom? Social welfare more broadly defined? Is there anything wrong with corporate charity? Should corporations try to maximize profit over the long run or the short? Our response to such questions is: who cares?.. If a corporation is started with a promise to pay half of the profits to the employees rather than the equity investors, that too is simply a term of the contract. It will be an experiment. Professors might not expect the experiment to succeed, but such expectations by strangers to the bargain are no objection.77

Corporate law merely provides an operational assumption for corporations, absent explicit contractual terms that stipulate some other goal. The operational assumption is profit maximization but this need not be construed as mandatory. Indeed, from a normative perspective it cannot be. The perspective of Pareto efficiency is that a voluntary transaction is almost always a gain in efficiency, since it appears to signal that at least some party prefers the allocation resulting from the transaction to the pre-transaction allocation, while no party sees it as making him worse off than he was before the transaction.78 As such, if a corporation decides not to seek a profit at all, but manages to survive because people continue to invest in it, we have a good reason to believe that the corporation is actually promoting efficiency in the aggregate, despite appearances (or, perhaps, presumed intentions) to the contrary.

The nexus-of-contracts view of the corporation takes efficiency to be the normative foundation for corporate organizing. As a result it presumes that corporate parties wish to maximize value.

76 Ibid 38.
77 Ibid 35-36.
However, it does not mandate this as a goal; whatever contractual form a particular corporation takes, corporate law is to take a deferential approach and generally refrain from intervening or trumping existing contracts. While nudging corporations toward the maximand of shareholder value by making it the default operational assumption, corporate parties are free to contract around these presumptions. The contractual approach to corporate law assumes that whichever contractual arrangements parties take on are, in the long run, good for promoting efficiency, absent heinous third party costs or fraud, all of which are best handled by standard contract law or tort law.

This efficiency-rationale for corporate law is used to defend the various features of the corporation such as limited liability and corporate personhood that were reviewed in the introductory chapter. Here I will consider the two prescriptive features of corporate law most significant for our purposes: shareholder voting and managerial responsibility, bearing on the questions of corporate governance and business ethics respectively.

**Shareholder Voting**

Easterbrook and Fischel break up the question of shareholder voting into three separate questions: Why is there voting in the first place? Why do shareholders in particular get the vote? And (why) do shareholders exercise their votes? As might be expected from their emphasis on efficiency, none of the answers to these questions have anything to do with “democracy” or “autonomy.” Instead, voting rights are viewed as the rights allocated for the purpose of specifying terms that are left underspecified by existing contracts or by corporate law.79 Put differently, voting is a means of reducing the agency costs that result from subjective interpretation of vague contractual terms. To avoid the costs that would result from interpretive disagreement, one class of patrons is given the vote and the power to specify the unspecified.

79 Easterbrook and Fischel, 66.
Contrary to Berle and Means, the idea of control is severed from an ideal of what ownership of property entails. The justification for shareholder voting is not due to their being “owners” who are entitled to controlling the enterprise. Shareholders are given voting rights simply because it is generally most efficient to allocate such rights to them; they have the right incentives, receiving “most of the marginal gains and incur most of the marginal costs. They therefore have the right incentives to exercise discretion….The right to vote follows the residual claim.” Shareholders are the voting class because they tend to be the group best positioned to exercise such rights. This does not imply that they always must be: corporations are free to assign voting right to whomever they wish, and indeed doing so can tend to achieve efficiency gains in the decision-making process. But it explains why corporate law presumes that the residual claimant (that is, the shareholder) will be the one possessing the right to vote.

This account helps to explain something else about shareholder voting, namely why shareholders tend not to do it very often. Just as Berle and Means noted, shareholders tend to be numerous and dispersed, making the attempt to control the actions of management in any real way difficult and costly. For this reason, even when shareholders do vote, it is generally either to delegate decision-making ability or to affirm that which has been presented to them by those they have delegated. The difference, of course, is that the nexus-of-contracts view finds nothing wrong with this. Because the end of shareholder voting is not some ethical ideal of property ownership, but merely the further specification of contractual terms, there is no problem with a passive shareholder class; they have served their function by delegating their decision-making authority to management, which is in a better position to gather information and dedicate the necessary time and resources for making the best decisions for the corporate venture. As agency theorists have pointed out, this reduction in monitoring

80 Ibid 68-69.
81 For the definitive account of this variety of ownership assignments see Henry Hansmann. 1996. The Ownership of Enterprise. Cambridge: Harvard University Press. I discuss this in more detail in chapter 4 and chapter 6.
costs raises potential agency costs; the nexus-of-contracts view adopts the property rights perspective and argues that the most efficient way for shareholders to reduce agency costs is simply to sell their shares.\textsuperscript{82} As a result, exit is favored over voice: access to the securities market is favored by the contractual approach as the best means for shareholders to check the directives of management. Governance is best handled by the market for ownership as opposed to mechanisms of democratic control.

As a result, Easterbrook and Fischel heavily criticize attempts to bolster or vitalize shareholder democracy.\textsuperscript{83} The proxy rules implemented and upheld by the Securities and Exchange Commissions are a case in point. These rules mandate disclosure of various types of information to shareholders on the assumption that shareholders inactivity is due to their being on the losing end of an information asymmetry. This is where the presumption of voluntariness in corporate affairs, and a strong disinclination towards meddling with such affairs, is brought into stark relief. With regards to these issues of shareholder voting, Easterbrook and Fischel are unequivocal:

The proxy rules displace private arrangements with respect to both the issues on which shareholders are entitled to vote and the information they are entitled to have. Regulation is not entitled to the same presumption of efficiency as long-standing voluntary arrangements…As we have emphasized, there is no reason why those who supply capital to the firm should have interest or expertise in managing the firm’s affairs…the rational strategy for most dissatisfied shareholders is to sell rather than incur costs in attempting to bring about change through votes. There are, however, good reasons why investors would choose to limit both the scope of voting and the information supplied: rational ignorance implies delegation implies less voting; and the costs of information imply limits on disclosure to investors who won’t act on information even if they possess it.\textsuperscript{84}

Calls for more shareholder voting, mandatory inclusion of shareholder proposals, and greater access to information are just different examples of forcing onto shareholders that which they evidently don’t want, since if they wanted it they would do it. Foisting an ideal of democracy onto shareholders only

\textsuperscript{82} Easterbrook and Fischel, 84.
\textsuperscript{84} Ibid, 82-83.
serves to reduce the welfare of shareholders who already currently pursue the rational strategy of passivity.

Advocates of shareholder democracy or shareholder activism are operating on a category error for contractualists like Easterbrook and Fischel, in that they assume that shareholders’ right to vote implies that democracy is somehow a value in and of itself in the corporation. But the question of corporate governance is merely the question of how “suppliers of finance to corporations assure themselves of getting a return on their investment.”\(^{85}\) Given that the corporation is merely a nexus of voluntary contracts, contemporary corporate law should assume that such contracts represent the most efficient means for the parties to pursue their interests. Imposing greater democracy is neither helpful for their ends (because they don’t care about democracy, only efficiency), nor is it necessary for them to see management’s actions as legitimate (because their purchasing of stock at a particular price functions as their consenting to the corporate contract).

\textit{Managerial Responsibility and Fiduciary Duty}

Another mechanism in corporate law that is meant to reduce agency costs and contracting costs, from the Chicago perspective, is the principle of fiduciary duty. By stipulating managerial fiduciary duty, corporate law requires that managers act in ways neither stipulated by contract nor directed by shareholders, but according to a general principle of duty that attaches to their position. Determining the content of fiduciary duties is a matter of approximating “the bargain that investors and managers would have reached if they could have bargained at no costs.”\(^{86}\) One should note here that this is essentially the “Coase Theorem” applied to corporate law. Cast in such a matter, the conceptual domain of fiduciary duty is limited to those activities that enhance the welfare of shareholders. To see fiduciary


\(^{86}\) Easterbrook and Fischel 92.
duty as somehow either trumping existing contracts, or including values other than those expressly stated in contracts, is to miss the point that fiduciary duties serve an efficiency-function similar to voting rights: they specify that which is underspecified or would be costly to specify, in contracts. “Fiduciary duty” simply performs this function by deterrence: managers know that if they violate their fiduciary duties they can be held legally accountable.

Immediately this precludes wider views of corporate responsibility to be included within a manager’s fiduciary duties. Shoehorning a thick sense of managerial social responsibility into the fiduciary principle in the style of Berle and Means only serves to increase agency costs, and thus undercuts its purpose of economizing on explicit contractual terms. Yet this analysis suggests an even more radical result. The economic rationale of the fiduciary principle not only restricts managerial responsibility to maximizing shareholder value; it implies that even such a restricted principle should rarely be enforced. Easterbrook and Fischel come to this conclusion by putting a heavy emphasis on the “business judgment rule,” which says that courts should start with the “presumption that a firm’s directors have, in fact, met their duty of care.” In order to find a corporate manager in breach of his fiduciary duties, plaintiffs must produce straightforward proof of managerial conflict of interest, or a deliberate/negligent failure to obtain necessary information for a particular action. Put differently, even if corporate executives implement a patently absurd strategy, or make disastrous decisions, they are protected by the business judgment rule as long as there is no clear and obvious conflict of interest or gross negligence; or, as Blair and Stout have put it: “a director is more likely to be hit by lightning after leaving her board meeting than she is to pay damages.”

The business judgment rule thus serves to protect managers from suffering the consequences of violating what would otherwise be their fiduciary duty to do. Put more strongly, the business judgment

rule minimizes the content of fiduciary responsibility to the point where it only prohibits clear-cut cases of fraud or self-dealing. What could the economic rationale be for not enforcing the managerial responsibility to maximize shareholder value? According to Easterbrook and Fischel, the explanation for this doctrine is the same as the explanation for why shareholder do not (and should not be forced to) exercise their votes: for shareholders it is more costly to uphold contracts through courts and liability rules than to just defer to managers or sell one’s shares. Litigation can be expensive, time-consuming, and reach inaccurate conclusions. Furthermore, judges do not have the type of business expertise that corporate executives have and the labor market for managers will more efficiently and reliably weed out ineffective management practices than legal liability rules.

The deference to managers implied by the business judgment rule is therefore underwritten by the same principle that explains why fiduciary duty is to be interpreted as only implying maximizing shareholder wealth: to do otherwise raises the costs of doing business, which is precisely what corporate law tries to mitigate. From a contractual perspective, the substance of the fiduciary principle implies its own limited application: “allowing shareholders to challenge business decisions that they say were not informed has the effect of substituting the business judgment of some shareholders, their attorneys, and a court for that of the managers. Because the managers have the best incentives, the legal process is distinctly inferior.” This is all in fitting with the general framework: if we start with the assumption that markets and voluntary transactions tend to maximize efficiency, and if we view the corporation as simply a nexus of voluntary market transactions, then it follows that that set of arrangements we find in corporations will tend to be most efficient over the long run. Therefore, absent an explicit contract or agreement that stipulates not to do a particular action, corporate law should largely be in the business of deferring to business, or to those contractually placed as decision-makers.

89 And, even then, only rarely as Easterbrook and Fischel note approvingly (98, 104).
90 Ibid 100.
91 Ibid 108.
It again should be noted that this cuts in a number of directions. Whereas Friedman argued that corporate charity was an example of managerial irresponsibility, or even small-scale socialism, the contractual view contends that this is mistaken. If managers wish to use corporate funds for charitable reasons, for example, it is not a violation of the fiduciary duty to pursue profits, despite appearances to the contrary. Charity *might* create goodwill with the community that might reduce costs down the road; buying floor-side tickets at a basketball game *might* be the thing that lures top talent to the firm. The point is that there is no way to be sure, and so courts are to defer. As a result, and in keeping with what was said earlier, corporate executives are perfectly free to pursue whatever ends they wish under a contractual view of the corporation. They are in fact under no fiduciary duty to pursue profits or maximize value.\(^92\)

We can summarize the nexus-of-contracts view of fiduciary duty and managerial responsibility in the following way: *the only duty that management has is the duty not to do that which they said they wouldn’t do.* There is no greater or overarching ethical or legal principle that directs executive activity other than the duty to uphold contracts. Lee has therefore concluded that in the contractual view of the corporation, corporate social responsibility or ethical business is not the purview of management but of the shareholder.\(^93\) Shareholders can collectively impose a more ethically thick duty upon management. Of course, it is difficult for shareholders to pursue such an agenda since the contractual view suggests that the collective action problems rendering shareholders passive ought not to be changed by legal fiat, as it would constitute an efficiency-loss. Thus while corporate objectives other than profit are perfectly acceptable for contractualists, their view suggests a legal regime that makes it incredibly difficult for such objectives to be implemented. Despite their differences with Friedman’s principled defense of

---

92 The idea that there is a legal requirement for corporations to maximize profits or shareholder value is challenged most forcefully in Lynn Stout. 2008. “Why we Should Stop Teaching Dodge V Ford.” *Virginia Law and Business Review* 163-190.

profit-maximization, the contractualists’ view of corporate governance winds up justifying a type of corporate practice that is very much in line with Friedman’s vision.

Conclusion

The Chicago School’s Political Theory of the Corporation

In the last chapter we saw how the original insights of Coase were used to address the absence of a theory of the firm. This developed a theory of the corporation that explained its efficiency in terms of its ability to overcome transaction costs. In this chapter we have seen how the Chicago tradition developed a normative account of how the corporation ought to be structured with regards to larger social and ethical values. This was, I have maintained, largely in response to critical managerialists like Berle, Means, Keynes, and Galbraith, who contended that the rise of the modern corporation represented a change in the economy; the growth of corporate bureaucracy and organization had developed a powerful managerial class whose rule was not in sync with the liberal ideals that had grounded the market economy for a century prior. Based on these developments, managerialists advocated a shift in our understanding of how the corporation ought to be organized, namely as a social institution: in place of an unaccountable managership, the corporation ought to be governed in a public manner so that larger social interests might have a say in the enterprise’s operation; in place of a permissive or prostrate court system that nominally protects shareholder interests, corporate law ought to impose fiduciary duties upon corporate managers to serve the interests of society at large; and in place of an ethical vacuum, business ethics ought to be informed by a social responsibility that tempers managerial drive for self-interested profit-seeking.

In response to this, Chicago scholars like Friedman, Jensen, Meckling, Fama, Easterbrook, and Fischel countered that the managerialists had grossly misunderstood the nature of the corporation. Contrary to their power-centric understanding, Chicago asserted a contractual understanding of the
corporation. Managers don't possess power, they are merely agents contracted to act on behalf of their principals. Markets for securities and labor enable all participants in the corporation to leave the sphere of its “power,” and contract elsewhere. The result is that we should see whatever decision-making power that managers hold to be legitimated by the choices of all involved to be party to the corporate contract.. Put better: because people contract voluntarily, and because markets will direct people to contract in such a manner so as to optimally allocate property rights, we should see managerial decision-making as legitimate because it is preferred by all involved. Managers hire labor; but in this view workers also hires management to supply them with the means to best use their labor. As a result, the actions of management ought to be constrained legally or ethically only insofar as they contractually limit their behavior in some way.

Another way of thinking about this is that Chicagoans deny the very premises assumed by what I am calling the normative component of the political theory of the corporation. Recall that the three questions posed are premised on the idea that the corporation poses problems for liberal democratic values: the existence of the corporate hierarchy raises concerns stemming from a social preference for democratic forms of authority, individual liberty and freedoms, and the moral and legal equality of persons. Chicago's response to these concerns is that the corporation is not a contradiction of these values, but an outgrowth of them. Liberalism's emphasis on individual choice and freedom is what grounds the contracts that are the stuff of the corporation; the liberal revulsion for non-democratic forms of authority is what grounds the preference for voluntarist forms of economic organization; and the liberal assertion of moral equality is what grounds the corporation's emphasis on securities markets, which gives all potential entrepreneurs access to capital in order to finance their endeavors. The efficiency of the corporation is presented as a prima facie case for the corporation's liberal legitimacy, absent extenuating circumstances.

We are therefore in a position to summarize the normative aspect of the Chicago approach to the corporation, in comparison to the managerial approach as depicted in Table 3.2 below. We get an
approach that takes a principled non-stance toward the organization of the corporation. Because the corporation is a nexus of contracts, the resulting distribution of rights and responsibilities is presumed to be both efficient and chosen by those involved. It is therefore justified both on grounds of liberty and of efficiency; it is both productive and just.

Table 3.2 Managerial and Chicago Normative Components

<table>
<thead>
<tr>
<th></th>
<th>Question of Corporate Governance</th>
<th>Question of Corporate Law</th>
<th>Question of Business Ethics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial</td>
<td>Formal and informal social control</td>
<td>Fiduciary duty to serve the social benefit</td>
<td>Social responsibility to serve the public</td>
</tr>
<tr>
<td>Chicago Law and Economics</td>
<td>There is no authority, just contracted relationships held in check by markets. Shareholder “democracy” is about specifying contracts, not legitimating power-holders</td>
<td>Fiduciary duty is to be interpreted with a maximum of deference for managerial decisions. The corporation has no legal purpose other than what is explicitly stated in contract.</td>
<td>Managers are to pursue the maximization of shareholder value, unless shareholders explicitly have them pursue some other end.</td>
</tr>
</tbody>
</table>

As has already been hinted at, such a perspective implies a significant departure from the productivity component reviewed in the last chapter. Whereas Coase and Williamson both asserted an analytic distinction between processes within and without the corporation—leading to a view of the corporation as those activities coordinated by managerial discretion as opposed to the price mechanism—later generations of Chicago scholars asserted that what happens within the corporation is essentially a market by another name, or a privately owned market. Commands issued by an employer are merely offers that come with corresponding prices, which can be refused by the employee in the same way a customer can leave a hardware store if she doesn’t like the price of a nail gun. In contrast to the transaction cost account, law and economics maintains that the merits of price theory extend to relations within the firm. We must therefore further modify the productivity component of the corporation to fit their theory as depicted in Table 3.3 below.
Table 3.3 Transaction Costs and Law and Economics Productivity Components

<table>
<thead>
<tr>
<th></th>
<th>Question of Efficiency</th>
<th>Question of Constraint</th>
<th>Question of Constitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pigouvian</td>
<td>Presumed as Production Function—“Lumps of butter coagulating in a pail of buttermilk”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction Costs</td>
<td>Facilitates party-relevant transactions</td>
<td>Costs of bureaucracy and lack of competition</td>
<td>Transactions coordinated by entrepreneur</td>
</tr>
<tr>
<td>Law and Economics</td>
<td>Facilitates party-relevant transactions through effective use of specialized information</td>
<td>Costs of bureaucracy</td>
<td>The legal fiction standing in for the nexus of contracts with which all parties implicitly transact.</td>
</tr>
</tbody>
</table>

By erasing the distinction between those transactions that happen within the firm and those that happen outside of it, the law and economics paradigm was able to counter managerial concerns about corporate power by effectively denying its existence. The departure from the Coase-Williamson story about corporate productivity is therefore what enables the normative response to the managerial story about executive authority and corporate power. This helps us understand the normative ramifications for how corporate efficiency is taken into account. The manner in which corporate efficiency is conceived of will change the scope of how we can articulate the ethical and political responsibilities of corporations to act and organize in particular ways. If we think of the corporation as achieving efficiency because it uses hierarchy instead of market transactions, then we can ask in what ways that hierarchy ought to be attenuated or assessed on ethical grounds. If, however, we think of corporate efficiency as being just an extension of the kind of efficiency that markets achieve, then hierarchy doesn’t appear to exist at all. As a result we are deprived of resources to claim that corporations must be organized in particular ways in order to avoid problematic forms of authority and power.

*The Individual and the Institution in Chicago*

Considering the Chicago tradition from the perspective of the corporation brings to light key aspects of how it conceives the relationship between institution and individual. Underlying Chicago’s
approach to both the productivity and normative components of the corporation is the assumption that individuals’ preferences are exogenous to the economic institutions that facilitate their cooperation. Inherent in Friedman’s business ethics, agency theory and property rights theory, and Easterbrook and Fischel’s account of corporate law is the idea that individuals are seeking to satisfy their preferences through exchange and through the legal machinery that enables specialized forms of productive cooperation. This is essentially the “utility maximizing” assumption at work. What we see in addition to this assumption is the idea that individuals’ preference exist independently of, and are largely unaffected by, the institutions that coordinate their activity. Viewing the corporation as a nexus of contracts, which is justified because it represents the preferred and efficient organization of freely contracting individuals agents, commits Chicagoans to an assumption that the preference schedules of those agents are a given. This idea was given formal articulation in one of Chicago’s most famous and controversial papers, “De Gustibus Non Est Disputandum,” where Stigler and Becker argue that tastes and preferences are fixed both for an individual throughout time, and for individuals across time.94 This assumption is crucial for the Chicago project to work. To concede that individuals’ preferences may be endogenous and affected by economic institutions would render question-begging the normative claim that such institutions are preferable because they are efficient, since we would want to know why satisfying such affected preferences is desirable in the first place.

I flag the importance of this assumption here because it shows where the Chicago approach differs from other approaches to institutionalism in the social sciences. Old institutionalist economics in the tradition of Commons, Veblen, and Dewey took very seriously how individual preferences are affected by context. The neoclassical and rational choice approach to institutions differed in this respect, and instead took equilibrium as the best way to explain the existence of institutions, thereby ignoring or downplaying the way in which institutional preferences change by virtue of institutional

context. In the past twenty years various approaches to institutions coming from sociology, political science, and heterodox economics\(^\text{95}\) have shown how these rational choice approaches can be married to non-economic determinants of individual behavior—such as power dynamics, culture, and custom—to give a more robust account of how institutions work and why they change over time.\(^\text{96}\) Insofar as the Chicago account requires the assumption that institutions serve to satisfy preferences, and not that preferences are affected by institutions, Chicago occupies a very particular spot in the spectrum of institutionalisms, one that is comparatively blind to sociological or political facets of social life.

None of this necessarily means that the Chicago paradigm is wrong. Perhaps the agency theorists are right that what appears like corporate power is merely just a price that employees and stakeholders can accept or deny as they wish. The impulse to see power or authority within the firm may only be an appearance that covers the more accurate transactional nature of corporate enterprise. In the next chapter I will articulate why I don’t think the Chicago school is correct. Before doing that, though, it is worth being charitable to the underlying logic of this theory. If we take Hayek’s point seriously about the informational benefits of markets, as I think one should, we also need to take seriously the idea that the results of a particular market might actually be reflective of people’s preferences. Let us say we observe a large firm dominating a particular market; the instinct might be to see this as merely a greedy executive amassing power through a system of political economy designed to enable him to do so. However, what we might also be seeing is that consumers just vastly prefer that firm’s products to the alternatives, which is why the firm has grown to such a degree. Taking Chicago seriously implies taking that latter idea seriously. If one wants to assert that the result still ought to be changed, one needs to offer a compelling reason for it, instead of just pointing to the existence of power. If one wants to challenge the efficiency of the market outcome, then this challenge should come


with a story about a flaw in the market that privileges a particular firm, or a collective action problem that makes the outcome of a given market sub-optimal for all involved. If one wants to grant the efficiency of the outcome but challenge it on behalf of some other value, then this challenge should come with a story about why it is just or fair to ask society to subsidize business practices that don’t contribute efficiently to the satisfaction of people’s preferences—to move resources away from some preferred use to some less preferred use.

Without belaboring the point, we should note that Chicago gets some grist for its mill when one takes into account the variety of corporate forms that exist and are enabled to exist under our current legal and economic order. As Hansmann and others have demonstrated, there is no reason to see the corporation as a singular type of organization, psychopathically pursuing profit as some critics have argued. This is why we have different types of corporations. The existence of non-profit corporations (let alone co-operative corporations or benefit corporations) gives us good reason to understand corporations as being a malleable form of organization. Chicago’s approach to the corporation offers compelling explanatory and normative accounts for this variety. If we see the corporation as simply a nexus of voluntary contracts, then it follows that corporate forms will be as numerous as there are desired and contractually obtainable organizational schemes.

While such organizational variety could be argued for from a liberal commitment to individual freedom, the nexus of contracts view contends that such organizational variety is not simply defensible from the standpoint of freedom but desirable from the welfarist standpoint of efficiency. We can expect to see different ownership structures surviving in different markets based on differences in industries. This is why we see more worker cooperatives in less capital intensive industries, why

97 Bakan 2004.
various agricultural sectors tend to include a disproportionate number of supplier-owned cooperatives, and why social clubs that are incorporated are almost always organized as non-profits. From the Chicago perspective, these are not simply because of the ideological commitments of the firm’s patrons, but because these organizational forms reflect achievements in efficiency. In this regard, certain corporations where rights to vote and collect residuals are held solely by employees (worker cooperatives), or others where there is no patron who claims residual profit (non-profit corporations), are seen as the most efficient way of securing a specific class of goods or services and therefore serve to contribute maximally to social welfare. Perhaps surprisingly, this view ends up defending and justifying on moral grounds the existence of co-operatives and non-profits. After all, they too are simply products of market-based contracting.

I characterized this earlier as Chicago’s principled non-stance toward corporate form. The legal and economic regime will tend to favor profit-seeking shareholder-financed, manager-run corporations, and Chicago is fine with this. But a variety of other types of corporation will also tend to exist and Chicagoans are committed to non-interference with these firms as well. Hence the rejection of the idea that managers have any thick sense of fiduciary responsibility, even to pursue profit. To argue for a more strict or constrained regime of corporate governance and corporate law—whether on behalf of the profit-seeking corporation or of the socially responsible corporation—one needs to recognize that such an argument is also an argument against organizational variety. I actually take this to be a very strong point, one which critics of Chicago must take seriously, and which I explain more in the next chapter. To confront this criticism, one must explain why a certain set of corporate governance structures, or particular corporate objectives, ought to be discouraged or disallowed, despite appearing to be preferred by constituents.

---

100 Hansmann 1996.
One can recognize this without Chicago’s great faith in markets or zealous emphasis on efficiency. There may be many instances in which markets fail, and there may be many values that trump efficiency. But the Chicago approach helpfully forces us to articulate objections to market outcomes with a respect for the efficient allocation of goods and the role such allocation plays in people’s welfare. The onus is placed on the critic to either explain why the outcome does not actually represent people’s preferences, or explain why some other value ought to trump people’s preferences in determining the allocation of society’s resources. Most charitably, Chicago offers a compelling account for why liberal values of individuality, equality, and democracy are not simple trumps to lay down against corporate institutions.
Chapter 4 Norm-Governed Productivity

In the last chapter I argued that by reading the Chicago School tradition of economic theory, corporate law, and law and economics we can actually discern a complex and compelling political theory of the corporation which answers the six questions posed in chapter 1. While my aim is to engage in a thoroughgoing critique of this tradition, I have attempted to present the strongest possible version of this theory. Instead of dismissing the theory as a cynical attempt to ennoble the position of capital and to rationalize the exploitation of labor, or as a colonial ideology, the aim has been to use the principle of charity to understand it on its own terms as an earnest theory. In this way we can criticize Chicago not only in terms of political-economic consequences, but in terms of its own assumptions and aims.

One purpose of this chapter is to show how the theoretical commitments of this school of thought lead to conclusions that contradict its own findings. Yet this critique is not the only goal of the chapter. Indeed, were the goal merely critique, then an external critique might be the most effective way to go about it. After all, the most provocative claim advanced by this view—that authority in the firm does not exist, since the relations that constitute the firm are mere voluntary exchanges—flies in the face not only of conventional wisdom and most personal experience, but also empirical work on the subject. Sociologists who study work are generally concerned with how authority is exercised and implemented differentially depending on things like race, gender, and class, as well as the institutions (like the school and the family) that help produce and discipline people to respect the authority of their bosses. All of the classics in the study of management affirm the role that power and hierarchy play in

---

1 Which isn’t to say such criticisms are wrong, but rather that their case has already been made. See David Harvey. 2007. A Brief History of Neoliberalism. Oxford: Oxford University Press.
the process of production within enterprises: Taylor famously argued that the manager’s job was to compel or trick the worker into exerting more energy into a task than he otherwise would want to⁴; Drucker’s use of political science methodology to study the firm was inspired precisely by the analogous role power plays in both disciplines⁵; one of Morgan’s “images” of the organization is as a government⁶; and McGregor speaks of managers as primarily using “threats and coercion” to control how workers act.⁶ Even when management professionals counsel their readers to include workers in decision-making processes and justify their decisions to them, this is based on the background notion that certain people are in positions of power vis-à-vis others. Thus on its face, the view of the internal governance of the corporation offered by Chicago seems fairly hard to swallow since it distorts the object of reflection beyond any resemblance to the actual world we know.

The more interesting question is not whether or not the Chicago tradition is wrong in its understanding of intra-corporate relationships, but rather why it is wrong or, more specifically, why the view cannot be right, even on its own terms. By engaging with the Chicago theory internally, I argue, we actually gain a better understanding of the corporation’s productivity and efficiency, as well as some desiderata for what a different theory of the corporation ought to look like. While I will attempt to show how the Chicago approach is overly myopic in its view of human values and human motivation, the sophisticated and subtle insights about the nature of efficiency that have been derived from the law and economic tradition must not be ignored. Therefore, I do not attempt to offer a positive account of the firm by, say, disavowing Chicago’s emphasis on efficiency, or arguing that some other value ought to be prioritized (though I don’t disagree with the sentiment). Instead, I attempt to derive my alternative view of the corporation from ideas and principles immanent in the Chicago approach.

---

Even if we take on the view that the corporation’s primary purpose is the pursuit of efficiency, we must come to a different understanding of the corporation than what Chicago offers.

In the first section of this chapter I return to Coase. While Coase was important for understanding the nature of corporate productivity, we see that a different aspect of Coase, and a different interpretation, has influenced much of Chicago’s theorizing. In order to reconcile these different views of Coase, and therefore to better understand the theory of the firm he inspired, I bring the overlooked nuanced moral psychology that underlies his approach to economic theory to bear on the question of the corporation. With this in hand I show how this view dovetails nicely with more recent behavioral economics and psychology on economic motivation in order to develop a norm-oriented understanding of corporate efficiency. In this view, it is the existence of cooperative norms, and not merely the contractual allocation of governing rights, that allows for firms to economize on market failures.

In the second section of this chapter, I consider how radical democratic proposals of the firm stack up against this view of corporate efficiency. While the norm-oriented view of corporate productivity is a boon to radical critics of the corporation—since it implies that there need be no tension between corporate efficiency and democratic participation—it also offers medicine that must be swallowed. Most significantly, this view shows that norms of fairness and equality cannot trump concerns of efficiency since it is the very pursuit of efficiency that creates the room for norms to exist within the market in the first place. By reviewing the theories of radical and participatory democrats, I try to show that while they have reason to criticize the illegitimate authority within corporations, their proposals are actually the mirror-image of those of the Chicago school: where Chicago assumes an overly utilitarian agent, radical democrats assume an overly social agent; where Chicago conflates firms for markets and obscures their cooperative nature, radical democrats mistake firms for purposive communities and obscure their economic nature. While the democrats are right that undefeated
authority exists within firms, they are often in danger, if not downright guilty, of discounting efficiency in their proposals.

In the third section and conclusion I explore a more general way of understanding the efficiency constraints that a democratic approach to corporate governance must address. The aim here is to explain what realistic factors a progressive reform to the corporation would have to take into account. I conclude this chapter with a more exact enunciation of the productivity component of the theory of the corporation, which I hold is more correct than the alternatives reviewed in the previous two chapters.

A Sociological Coasian Approach

The Third Coase

It is worth going back to the first Coasian insights that informed this theory and try to see where things went awry. To refresh our memories, the basic Coasian insight into the productivity of the firm was the idea of the transaction cost and that markets are sometimes not as efficient as bureaucracy or hierarchy. This basic insight led to a comparative logic when approaching transactions: for any particular transaction we might arrange that transaction through the price mechanism or we might arrange it through some form of hierarchy. However, the basic logic for why such transactions are being arranged remains the same: to most efficiently allocate a scarce quantity of goods.

What the Chicago theory seizes on here is that ultimately, despite the variety of transactions, the basic impulse underlying each is still self-interest. The efficient economy is still produced by individuals seeking to acquire the goods they need to live their ideas of the good life. In this reading, although Coase himself admitted a certain critical or socialist logic to his 1937 piece, there is nothing

---


inherently critical about the idea of transaction costs: a market with defined property rights and individuals seeking their own interest will produce the optimal allocation of goods, including the optimal amount of planning and bureaucracy necessary to achieve this outcome.

There is something compelling about this way of looking at Coase, especially if we think of his “Nature of the Firm” with the insights from his famous 1960 piece in mind. As we saw, the “Coase Theorem,” says that markets will not simply allocate goods efficiently but, if unimpeded, will also most efficiently allocate the legal right to receive remedy in tort and the corresponding duties to pay such remedy. Or, as Stigler put it: “under perfect competition private and social costs will be equal.” That is, once rights are allocated, the legal system ought not to weigh in on who wronged who when it comes to externalities, since private contracting will just as easily take care of the problem. This idea easily gets read into the Coasian understanding of the firm, with the result being that whatever division of rights we find within the firm is the most efficient: “voluntary transfers of property rights should implement an efficient allocation, with residual claimancy and control of projects assigned to those who can operate them most productively.”

Yet it is actually bizarre that these two arguments should find themselves in such close proximity to one another since they operate on seemingly opposite assumptions. Whereas “The Nature of the Firm” argues that firms exist because of the existence of transaction costs, “The Problem of Social Cost” rests on the explicit assumption that there are no transaction costs. That Coase is most famous for these two seemingly irreconcilable arguments, and that biographically Coase did in fact go

---

from a young socialist to an older liberal, has led to the idea that there are two Coases\textsuperscript{12}: the Coase of 1937 whose insights led to the institutionalist tradition of social science,\textsuperscript{13} and the Coase of 1960 whose insights led to the Chicago tradition of law and economics.\textsuperscript{14} If one is inclined to believe in the importance of institutions like the welfare state, then the limits of markets implied by the 1937 article will find resonance; if one is inclined to be distrustful of government and favor free markets, then the 1960 piece will resonate.\textsuperscript{15}

Medema has offered a more compelling unified view of Coase’s work that is instructive. For Medema, while almost everyone recognizes that Coase is sparring with Pigou in 1960, the vast majority have underestimated the extent of his critique. What Coase is really doing in “The Problem of Social Cost” is showing the absurdity of the zero transaction cost assumption, which underlies the Pigouvian tradition of welfare economics:

Pigouvian analysis, from Pigou on down, did not contemplate the notion of transaction costs, and so implicitly assumed that transaction costs are zero….Yet, within such a framework inefficient externalities would not—could not—exist. Economists, Coase charged, were trying to explain divergences between private and social costs using a theoretical framework in which private and social costs were inevitably equal. This, he said, led to nonsensical conclusions.\textsuperscript{16}

In other words the critique of Pigou is not the generic free market critique of calls for government corrections of the market; instead, the critique is that in the world of Pigouvian economics, the types of government action Pigou demands would not exist. In a world of no transaction costs, there would be no externalities—since, as Coase shows, bargaining could solve these problems—and thus no need to


internalize these externalities. In this view the Coase Theorem is decidedly not a critique of any particular public policy or a call for a freer market but rather a *reductio* of the basic framework of welfare economics.

This allows us to see how there would be a unity between the two Coases. What both of these famous arguments actually show is the pervasiveness of transaction costs, as well as the problems of “blackboard economics” that place too much faith in their ability to model the real world with unreal-worldly assumptions. In 1937 the point was to show that every firm implied the existence of transaction costs; in 1960 the point was that uninternalized externalities also showed the same thing. The necessity of having organized hierarchy and bureaucracy, as well the need for tort law and taxation to deal with externalities, showed just how inescapable these transaction costs are, which are otherwise assumed away by economic modelling.

The synthesized Coase is a theorist who is not only concerned with the limits of markets, but the limits of economic thinking that, in turn, serve to reinforce economic problems and obscure potential pragmatic solutions. While this is an important point about the unity in Coase’s thinking with regards to institutions, it does not fully address our concerns. Because ours is a world where transaction costs are inescapable, there appears to be no good reason to think of corporate ownership in terms of a market that will efficiently allocate rights and liabilities. But what about the claim that the firm is still fundamentally composed of individual utility maximizers? That is, even if the Chicago school cannot claim that the outcome will be Pareto-optimal, the other part of their claim—that however efficient it actually is, the corporation and its allocation of ownership is still just the result of individuals seeking to gain their advantage—seems to remain intact. Thus even if it is not the most efficient one, the corporation still might be viewed as a nexus of contracts amongst self-interested subjects, which blurs the type of power and authority wielded in organizations.

I would like to argue that we can actually expand the unified reading of Coase not only to explain his various and diverse writings on institutions, but also to elucidate his more subtle views on
moral psychology. As has been pointed out, the “Coase Theorem” (as opposed to what Coase is actually doing in 1960) is really just an extension of Adam Smith’s invisible hand. Accordingly Coase is often read as having the same view of man’s moral psychology associated with the invisible hand, the propensity to truck and barter and our tendency to act for our own self-interest. Yet this misses Coase’s understanding of how human moral psychology actually works. In fact, it misinterprets Smith’s understanding of how moral psychology works, which Coase points out in a pair of essays on the Scottish philosopher. Coase correctly and approvingly notes that for Smith the claim is not that people are in fact only self-interested: “Man’s behavior, as the author of The Theory of Moral Sentiments knew, is influenced by feelings of benevolence—and the division of labor within a family, even an extended family, may be sustained by love and affection.”

It is not that everywhere self-interest reigns, but rather that, in order to cooperate on a grand scale with anonymous strangers, “reliance on self-interest is not simply one way in which the required division of labor is achieved; for the division of labor needed for a civilized life, it is the only way.”

Coase here makes a point about the way in which economists often read Smith that is helpful for understanding how this all relates to the corporate enterprise:

We just do not have the time to learn who the people are who gain from our labors or to learn of their circumstances, and so we cannot feel benevolence towards them even if benevolence would be justified were we to be fully informed. The fact that when discussing Adam Smith’s treatment of the division of labor economists have usually quoted his famous pinmaking example…rather than the long passage [on the role self-love must play in a cooperative society]—where the participants in the division of labor are scattered all over the world—has also helped to divert attention from the extremely limited role benevolence could play in bringing about the division of labor in a modern economy.

In chapter 2 I noted that Smith’s pinmaking example led to confusion as to why the division of labor was coordinated within the factory by hierarchy and outside the factory by the “invisible hand.” Coase

19 “…It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest…”
here draws attention to another confusion of the pinmaking example: it obscures the fact that the division of labor he wishes to celebrate is in fact distinct from that found within the pin factory. What Smith is at pains to show is that a spontaneous division of labor is created through the market without anyone realizing. What Coase shows, in noting this subtle point, is that not only is the institutional coordinating mechanism inside the factory distinct from that outside of it, the incentives and motivation for people to engage in this cooperation are different as well.

What Coase appears to glean from Smith is that man is not naturally self-interested, but that we encourage and cultivate self-interestedness within the market in order to cooperate on scales that our natural benevolence and empathy would not allow. This would be a slight point if the market encompassed all of cooperation, since regardless of its naturalness, self-interest would still be the only relevant motivational trait. Yet, as the pinmaking example shows and we have seen already, the market is not the entirety of our cooperative scheme; similarly, self-interest is not the entirety of our motivation to cooperate. The most obvious example would be the family, and cooperation amongst friends, which Coase insists are necessary to make the foundations of a market system work.\footnote{Ibid; Ronald Coase. 1976. “Adam Smith’s View of Man.” \textit{Journal of Law and Economics} 19(3): 544}

The more fundamental logic of the claim, however, is that the more anonymous and numerous the potential cooperators, the more self-interest is needed to make the scheme function, since benevolent and empathic motivations will only get us so far. Similarly, the more familiar the people are to one another within the scheme of cooperation, the less necessary self-interest is. In fact, we might find that self-interest corrodes these schemes. This is, as transaction cost economics has shown, the whole point of the corporation. In order to avoid the problems presented by bounded rationality and opportunism, we develop a governance system to enable the types of asset-specific investments and relationships which could otherwise not happen in an anonymous market. As we move from the anonymous market to the firm, from the invisible hand to planning and management, we also should
expect to see a move from self-interest to “the observance of moral codes” which “must very greatly reduce the costs of doing business with others.”

What I am proposing then is that though Coase claims the firm is a nexus of bureaucratically coordinated transactions that exist in order to economize on market transactions, it does not follow that the firm is therefore a nexus of self-interested individuals who cooperate on comparable terms as they do on the market. Instead the claim is that the comparative approach to cooperative institutions should also come with a comparative understanding of the different motivations and incentives at play in these institutions. Just as there is a unified critique of Pigou in Coase’s various writings on institutions, there is a unified understanding of the relationship between a person’s motivational profile and the institutional setting in which she is participating. The distinction between the dispersed, spontaneous cooperation of the market and the planned organization of the corporation comes with a motivational difference that, I will argue, implies a difference in moral codes: generally more adversarial in the former, and generally more cooperative in the latter.

_Firms as Bubbles of Norm-oriented Behavior: Resisting the Crowd-out_

This understanding of moral incentives, which we might refer to as the Smith-Coase view of moral psychology (the former for pioneering it, the latter for extending it to a contemporary market context), is actually tied to a more general understanding of the interrelationship between moral psychology and the economy than it is to our narrow concern with the corporation. What this view importantly brings out is that neither the market, nor the instrumental and self-interested roles that characterize it, are natural in any non-trivial sense. Instead, markets are important for mass modern society because of their capacity to facilitate cooperation on grander scales than we otherwise might; the cultivation and development of our more self-interested tendencies are both a by-product and a

---

22 Ibid, 545.
facilitator of this institution’s function. We are not innately unconcerned with things other than our own self-interest, nor are we naturally without benevolence; rather, our current modern scheme of cooperation has been produced by a long historical transition that has placed primacy on these traits in order to facilitate wider cooperation and organic solidarity. To understand how this all fits into a theory of the corporation, it is worth elaborating on this grander narrative.

This implied narrative is actually borne out by current work in anthropology. Contrary to economists, to whom the existence of non-maximizing behavior represents a puzzle, anthropologists have tended to be concerned with the opposite: how non-moral or non-normatively oriented behavior comes about. This is because ethnographic and historical research shows that the organizing principle that has structured cooperation for the vast majority of human history was reciprocity, not self-interest. As was classically stated by Marcel Mauss, cooperation has traditionally been secured through an economy of gifts as opposed to commodities and wages. In this form of cooperation trade does not occur because two people both desire it, but as a result of a wider set of obligations owed both to the individual and the wider community. As Mauss states it: “Food, women, children, possessions, charms, land, labor, services, religious offices, rank—everything is stuff to be given away and repaid. In perpetual interchange of what we may call spiritual matter, comprising men and things, these elements pass and repass between clans and individuals, ranks, sexes, and generations.”23 While cooperation might not have been facilitated by a thoroughly conscious acknowledgement that what one was doing was participating in wide-scale cooperation, it was an explicitly social and moral enterprise; to enter into commerce with somebody was to enter into a moral relationship characterized by norms and codes that must be observed, and which resulted in social sanctions when broken.

The question, then, is why this form of moral motivation could not sustain larger schemes of cooperation. The answer, to put it simply, is that as the number of cooperators grows, so does the probability that the deed being done will not be reciprocated; the chance that you will be cooperating with someone who intends to free-ride grows. Boyd and Richerson have demonstrated this with a formal model, concluding:

Reciprocal strategies must satisfy two competing desiderata to succeed. First, to persist when common, they must prevent too many defectors in the population from receiving the benefits of long-term cooperation. The threshold number of cooperators thus must be a substantial fraction of group size. Second, to increase when rare, there must be a substantial probability that the groups with the threshold number of cooperators will form. This problem is not great when pairs of individuals interact; a relatively small degree of assortative group formation will allow reciprocating strategies to increase. As groups become larger, however, this desideratum can be satisfied only if the threshold numbers of cooperators is fairly small, or the degree of assortment in the formation groups is large....The evolution of large cooperative societies normally depends more on kin selection [or cultural analogs to kin selection] than reciprocity\textsuperscript{24}

What is needed in order to facilitate cooperation on a mass-scale, according to Boyd and Richerson, is an effective institution of punishment, which functions as an analog of kin selection, by selecting out those would-be defectors.\textsuperscript{25} Whether their precise point about punishment is perfect is not as crucial as their overriding point: reciprocity as a normative principle can only work to secure cooperation when the numbers are small. Once numbers grow to a certain point some form of structure that can work on more consequence-oriented incentives must be introduced.

The market, in this view, is only the most sophisticated, large-scale, and effective—which isn’t to say “positive”—scheme of cooperation based on self-regarding incentives we have developed. Or, put differently, it is a scheme of cooperation achieved through non-explicit moral norms that has appeared to achieve great success. Habermas discusses the development of the market economy in a helpful way:


The capitalist economic system marks the breakthrough to this level of system differentiation, it owes its emergence to a new mechanism, the steering medium of money. This medium is superficially tailored to the economic function of society as a whole…it is the foundation of a subsystem that grows away from normative contexts. The capitalist economy can no longer be understood as an institutional order in the sense of the traditional state; it is the medium of exchange that is institutionalized, while the subsystem differentiated out via this medium is, as a whole, a block of more or less norm-free sociality.\(^\text{26}\)

In this sense the market is perhaps the “system” par excellence for Habermas as it represents a domain where actions are generally consequence-oriented, in contrast to the norm-orientation of action within the lifeworld. The development of the market privileges our strategic and instrumental impulses in order to sustain a system of cooperation that does not require agreement on, or compliance with, substantive norms from all those who cooperate. In a sense then, Smith’s view of the market and the place for self-interest actually understates the historical significance of the market; it is not merely the scope of cooperation that distinguishes it from prior cooperative schemes, but also the manner in which that cooperation is secured.

If we follow the logic of Coase’s assessment of Smith, we can start to see how this helps us with the puzzle of the corporation. As relationships become more familiar, so too does the possibility for moral codes to effect reciprocal and intentional forms of cooperation. In chapter 2 we noted that Williamson’s major innovation was introducing opportunism into the economic equation. This now appears a much more subtle point. The claim is not necessarily that humans are always and everywhere in fact opportunistic but rather that opportunistic stratagems are privileged by both the external incentives and the adversarial relations of the market mechanism (which I will elaborate upon shortly). But it is also precisely this opportunism that limits the extent to which markets can efficiently allocate goods. Thus the problem of asset-specific investments are not generic to human nature, but rather are created by the instrumental approach to social action cultivated by the market. While its development allows for cooperation to

become more expansive, it does so at the cost of a whole class of goods and services that require a thicker form of sociality in order to get made. Self-interested incentives enable much in the way of cooperation, but tend to stifle the potential for other motives to come into play.

This stifling of other potential motivations is referred to as “motivational crowding out” by the literature in behavioral economics. Drawing on various disciplines, the basic idea of “crowding out” is that people are motivated both extrinsically—that is, by material rewards gained by consequence of an action, for example—and intrinsically—by doing the action itself. The recognition of intrinsic motivation is not meant to deny strategic or “rational” behavior and thinking; like Smith\(^{27}\), Coase, Habermas, and others,\(^{28}\) this view affirms a diversity of moral motivation. However, what makes this literature unique is its finding that their roles in determining our action are not fixed but can be altered by different types of incentives. More specifically, as extrinsic incentives for action increase, other intrinsic incentives are “crowded out,” making them less salient in determining action. From an economic perspective, this has the very counter-intuitive result where sometimes “raising monetary incentives reduces, rather than increases, supply,” making it advisable to bring intrinsic motivation into play.\(^{29}\)

In tandem with the transaction cost argument, we might say that the problem that firms are meant to address are the inefficiencies resulting from the crowding out of intrinsic motivation, which

\(^{27}\) Lisa Herzog (2011) has actually made the connection between Adam Smith’s philosophy and motivational crowding out explicit: “In Smith’s system, the different passions have their place in different spheres, sometimes supported, sometimes unsupported, by institutions. But these spheres can intersect; wrong principles can be brought to bear on a situation, or a situation can be wrongly classified as belonging to a certain sphere. Motivation crowding out is structurally similar in that it describes phenomena in which intrinsic motivation, which should be the guiding principle in certain spheres, is ‘crowded out’ because the incentives (or even the description) of the situation make it appear as one in which selfish behavior is appropriate” (“Higher and Lower Virtues in Commercial Society: Adam Smith and Motivation Crowding Out.” \textit{Politics Philosophy Economics} 10(4): 382).

\(^{28}\) Jon Elster for instance, notes: “To accept social norms as a motivational mechanism is not to deny the importance of rational choice. One eclectic view is that some actions are rational, others are norm-guided. A more general and more adequate formulation would be that actions typically are influenced both by rationality and by norms.” (Jon Elster. 1989. “Social Norms and Economic Theory.” \textit{The Journal of Economic Perspectives} 3(4): 102).

makes cooperators vulnerable to holdup problems. As Osterloch and Frey argue: “Opportunism is a strong form of extrinsic motivation when individuals are not constrained by any rules. In the transactions cost view, the task is to establish institutional settings that mitigate the hazards and costs of opportunistic behavior.” The question is how firms as institutions mitigate these hazards. The answer is that they reintroduce the cooperative norms crowded out by the market. While competitive pricing is able to work its information-producing magic because of the external and monetary rewards it promises to individuals who heed its call, the firm is necessary precisely because such incentives also create problems. Firms suspend the price mechanism in order to bring more intrinsic motivation to bear, creating environments that can facilitate the development of trust and the procurement of asset-specific investments. In chapter 2, I cited Coase’s quotation of Robertson’s description of firms as “islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk.” Here we might state it more technically: firms are like lifeworld lumps of norm-oriented, intrinsically motivated butter coagulating in a pail of strategic, extrinsically motivated buttermilk.

This view makes sense of an empirical oddity much discussed in the economics literature, namely the existence of involuntary unemployment, which seems to indicate that firms are paying more than the market rate for services from their employees. Akerloff claimed that this was because firms do not pay employees simply for their services; instead the firm often functions on a gift economy where employees are paid for their services plus more, so that they will reciprocate by working in ways

31 See George A. Akerloff and Rachel E. Kranton. 2005. “Identity and the Economics of Organizations.” The Journal of Economic Perspectives 19(1): 9-32. For instance: “many monetary incentive schemes create opportunities for workers to game the system...People respond almost too well to monetary incentives. That is, firms get what they pay for, but since the schemes cannot be targeted well, what firms get is often not what they want. These problems indicate that if an organization is going to function well, it should not rely solely on monetary compensation schemes” (11).
32 This is implied by the definition of involuntary unemployment: “a situation where an unemployed worker is willing to work for less than the wage received by an equally skilled employed worker, yet no job offers are forthcoming” (Carl Shapiro and Joseph Stiglitz. 1984. “Equilibrium unemployment as a worker discipline device,” American Economic Review 74(3): 433).
that cannot be effectively monitored or extrinsically incentivized. In other words, members of firms are paid not only for their services but also to secure their cooperative orientation by establishing an environment of trust that encourages following norms of reciprocity. Whether or not one accepts the gift economy view, the larger point seems hard to avoid: firms are distinct from markets not just in institutional structure (management vs. price-mechanism) but in the norms and social rules that govern behavior therein (strategic/adversarial vs. normative/cooperative). Now, when I suggest that corporations are cooperative or use social norms to coordinate behavior, I do not mean to imply that these cooperative norms are necessarily good or pro-social. I adopt Elster’s distinction between rational action and social normative action, where the former is oriented toward outcomes and the latter is not. As I suggest below, the norms used to coordinate behavior are often problematic and alienating. What I mean by “cooperative” here is simply that people are consciously working with one another according to a shared script, and not through competitive and adversarial schemes.

As a result, the rational choice account of institutions—where institutions are understood in terms of the incentives and disincentives they offer—is ill-suited for the firm since the firm comes into existence precisely to address collective problems that cannot be addressed by extrinsic incentives alone. For this reason, firms are better understood in terms of the sociological institutionalist idea that

---

36 Again, I reiterate that Habermas describes the market as *more-or-less* norm free. I am not claiming that there are in fact no normative concerns in the market, or that intra-corporate relationships are without any form of competition or self-interested behavior. Instead I am suggesting that in general markets are characterized by competition and unintentional cooperation in which certain social norms are suspended for the sake of efficiency, and that the more intentional cooperation within firms is enabled by drawing on social norms. In asserting this, I do not mean to imply that the ideal typical ways I characterize these different modes of economic coordination are a factual description of the world as it exists.
“institutions influence behavior by providing the cognitive scripts, categories and models that are indispensable for action.”\textsuperscript{38} The idea of a cognitive script is a particularly useful one. As Giola and Poole wrote in an influential article, “a script is a knowledge structure that fits predictable, conventional, or frequently encountered situations…Scripts are schemas for behavior, or for understanding events and behaviors.”\textsuperscript{39} Scripts structure the way in which people process and engage with the world around them by grouping scenarios together into common events that suggest a correct mode of action. In contrast to economic theories of how institutions provide incentives and disincentives to coordinate action, this sociological view focuses on how notions of legitimacy and socialization are used to coordinate individuals in groups, often without those individuals being aware.\textsuperscript{40}

From this perspective, in contrast to the purely extrinsic motivations offered by the market, firms supplement the incentives and disincentives they offer with scripts that structure cooperation enterprises by creating recognizable, and recognizably correct, forms of action and interaction. This is captured in the idea that such forms of organization are “legitimate.” This is not meant in the sense that normative political theorist use the term, but in the sociological sense or, as March and Olsen put it, something is legitimate “not in the sense that what happened needs to be viewed as desirable or pleasant, but in the sense that what happened can be viewed as having occurred in the way things happen.”\textsuperscript{41} Now certain arrangements might be considered legitimate because they benefit the parties that partake in them; however, things might also be considered legitimate because they are grounded in

shared moral worldviews or in shared background assumptions of what is considered socially acceptable. This latter form of legitimacy, which Suchman refers to as “cognitive legitimacy,” is particularly interesting because it is so easy to overlook; myriad social institutions and practices—the way schools are run, the way families interact, linguistic norms, common religion, etc.—can structure what is taken for granted or deemed acceptable in particular situations.\textsuperscript{42} The result is that habit or “behavioral lock-in” can be extremely effective ways of coordinating activities that might not be otherwise chosen on the grounds of a strategic calculation.\textsuperscript{43} Firms and corporations use precisely these kinds of norms and expectations, these scripts of legitimacy, to structure cooperation. And this cooperation is used to overcome the dilemmas created by the market’s “norm-free sociality,” which is precisely the suspension of these norms in the service of a competitive and adversarial mode of coordination.

As a result, the manner in which people cooperate within the firm is significantly more complex than the voluntarist and contractual account that I reviewed in the last chapter. While people may join the corporation voluntarily because of the extrinsic incentives on offer, within the corporation their actions are coordinated by a set of social and moral scripts that are used to overcome potential cases of opportunism. As some commentators have put it, firms are actually able to function efficiently less because they can appeal to people’s preference schedules, but because they affect and alter those preferences through the social settings they produce.\textsuperscript{44} Against the assumption of the Chicago School that preferences are exogenous, we have good reason to believe that part of the reason why

\begin{flushright}


\end{flushright}
corporations exist at all is precisely because preferences are altered by cultural and institutional context.

Efficiency Constraints on Norms: The Problems with Workplace Democracy

The foregoing leads to a rather counterintuitive conclusion: the efficiency-mandate of corporations opens up, and actually requires, a certain space for normative values other than efficiency. Firms must offer cooperative schemes that appeal to norms and scripts other than self-interest in order to encourage the non-opportunistic investment of time and resources into firm-specific activities. There is an impulse to think that norms of equality ought, therefore, to be used to structure relationships within the firm. Though nobody puts it quite like this, this seems to be the idea underlying calls for the democratization of the workplace. After all, to embrace democratic norms is essentially to adopt a script that explains how to deal with cooperative endeavour: decision-making ought to be done with as much input as possible from all those involved, individual liberties ought to be protected, and equal status ought to be secured. So, when presented with a setting like the corporation where norms and scripts are brought into play, we might simply apply liberal democratic norms to structure the script: corporate hierarchies ought to be brought under democratic control, the differentiation of status ought to be leveled to as great a degree as possible, and individuals ought to be protected from the infringement of authority for a variety of potential trespasses. Put differently, we might think that our liberal democratic values provide us with a generic procedural script to apply when we find ourselves within a hierarchical scheme of cooperation, inviting us to democratize the corporation, just as we do with statutory institutions.

---

45 John Dewey perhaps came closest: “That the economic and industrial life is in itself ethical…is the meaning of the statement that democracy must become industrial.” (1899. “The Ethics of Democracy.”)
To illustrate the subtlety involved in the idea of norm-governed production, I here consider some of the most prominent theoretical traditions of workplace democracy. Calls for workplace participation, industrial democracy, or radical economic democracy are almost as old as industrial capitalism itself, resulting in a number of well-developed arguments. Workplace democracy, then, gives us a case of a strong theory that attempts to apply robust and demanding normative principles and cooperative scripts as a way of prescribing the governance structure of the corporation. By considering how these arguments fare, we can learn more about the complex interplay between norms of fairness and efficiency mandates that the corporation must inherently balance.

Three Argumentative Strategies for Workplace Democracy

1. Parallel Case Arguments

We might group radical or participatory theories of workplace democracy into three main argumentative strategies. The first type has been referred to by Arneson as the parallel case strategy, and most closely resembles the logic just given above.\(^{46}\) The core of the parallel case argument is a claim that the firm and the state share certain features, namely that both have authority and both involve hierarchy and submission to the decisions of others. Yet, with the state we claim that political authority must be subject to popular checks or democratic procedures in order to be legitimate. Therefore, the argument goes, the corporation cannot be legitimate unless it is similarly subject to democratic control or, as Robert Dahl put it, “if democracy is justified in governing the state, then it must also be justified in governing economic enterprises.”\(^{47}\) This view argues that in approaching the problem of cooperation within the firm we ought to apply the scripts we use in political society. There is great reflective purchase in this type of argument: by drawing attention to our own intuitions and


presuppositions about the inviolability of democracy, this theory asks us whether there is a good principled basis for not extending these presuppositions to economic enterprise. For people like Dahl, the claim is that there is none. In fact, not only is there no reason to distinguish the two when it comes to democratic control, there are good reasons to believe we ought to see both cases as comparably in need of democracy.

For Dahl, the normative similarity between the economic enterprise and the state is articulated in two principles, “the principle of affected interests” and “the elementary principle of fairness.” The principle of affected interests states that everyone who is affected by decisions should have a part in making those decisions.48 The elementary principle of fairness states that scarce and valued goods should be allocated fairly, with equal distribution being the default when there is no good claim for one party to have a greater share.49 The parallel case is made by claiming that in states and in economic enterprises there are people who are affected by decisions, and should therefore have a part in making those decisions; furthermore, because decision-making is a scarce and valued good, this decision-making power ought to be distributed equally amongst all the parties in the form of votes. Thus the rationale for democracy in the state is the same as it is for the workplace, such that if we claim a state to be only legitimate if it is democratic and grants its citizens universal suffrage, we must similarly grant employees of a firm the same rights.

Walzer arrives at similar conclusions, though he does so with a characteristic resistance to Dahl’s foundationalism. While similarly concerned with the distribution of decision-making rights, Walzer’s complex egalitarianism argues that since goods are social in nature, so too is their meaning and value. Accordingly the principle of justice that governs the distribution of a particular good is social and contextual in nature. The fact that a particular distributive principle pertains to one set of

goods does not imply that it applies to others: “when meanings are distinct, distributions must be autonomous. Every social good or set of goods constitutes, as it were, a distributive sphere within which only certain criteria and arrangements are appropriate...There is no single standard. But there are standards for every social good and every distributive sphere.”

Thus instead of the Dahlian tack where one principle can be applied to two different institutions, Walzer must explain why the corporation and the state operate in the same sphere of distribution.

Using the example of the Pullman Palace Car Company’s excessive purchasing of property in Pullman, Illinois, Walzer argues that purchasing the power of influence is clearly in violation of the political sphere’s egalitarian principle. The good of political power is therefore a distinct good from that of the market, making the distributive principles of the market different from those of political power. The democratic norm of one person, one vote is the distributive principle that governs political decision-making, and implies that the decision-making within corporations must be protected and insulated from the economic sphere. Workplace democracy is precisely the result of recognizing this distinction.

There is certainly an intuitive appeal to this type of argument. However, these arguments suffer from two major difficulties. In the first instance it is not clear that the parallel case takes all the relevant normative factors into consideration. Such factors include the pervasiveness of state power relative to that of the corporation, as well as the high costs of emigration relative to leaving the firm. However, the most important distinction seems to be the competitive market environment in which the corporation operates. If we follow Weber’s definition of the state as the organization with a monopoly on the legitimate use of violence, then the monopolistic aspect of the state offers an important contrast with the way firms have to behave and operate; insofar as corporations are not monopolies, and competitors

51 Ibid, 301.
who offer comparable services exist, it seems that the corporation is dis-analogous from the state in the most important respect. Particularly for Walzer, this is a rather devastating point. Walzer’s complex equality seems not to have been extended all the way, as he assumes rather than shows why the power within the corporation and the power within the state exist in similar spheres with the same distributive principle governing them.\(^{52}\)

This lack of competition for the state is part of a larger point about the distinction in how associations and societies can approach ends. As Rawls points out, a state, by virtue of holding the monopoly on force, must be neutral with respect to conceptions of the good and therefore cannot pursue any particular defined end. Associations like the firm, by contrast, can and do pursue particular ends; if the end is profit or value maximization or whatever bottom-line term we wish to use, this will necessitate competing against other firms in a market, and taking into account things like efficiency and scarcity. The ability to pursue these ends must be balanced and weighed against other concerns such as, for example, a concern for democracy; but it does not seem to imply that profit must be subordinate to democracy, especially if people’s primary reason for associating with the firm is material in nature.

The second major objection to the parallel case strategy is that it actually smuggles in another parallel-case that is not so easily defended. That is, not only must an analogy be drawn between state and firm, one also must claim an analogy between citizen and employee. While it is easy to explain why there ought to be some democratic control of the firm, it is not clear why the workers ought to be the patron group for whom democratic control rights are designated. We might very well point out that some form of shareholder democracy exists and is often the norm, and then respond to Dahl and Walzer by saying that the principle of affected interests, or the distributive principle of political power, is satisfied through shareholder voting. Furthermore, the distinction is not merely between workers and

\(^{52}\) As Robert Mayer puts it: “Security and welfare, money, office, hard work, free time, kinship and love, recognition—each is subject to different rules in different contexts. The only exception is political power, which for some reason is simple rather than complex.” (Robert Mayer. 2001. “Michael Walzer, Industrial Democracy, and Complex Equality.” *Political Theory* 29(2): 247.)
capital providers, but also between customers and raw goods suppliers. Why should we form worker cooperatives at the expense of customer cooperatives or supplier cooperatives? What makes the employee more like a citizen than a customer or anyone else?

This question seems to point out two crucial flaws in the principle of affected interests as it is used in making the parallel case strategy. On the one hand, this principle does not seem to accurately explain the normative bases for political democracy since we generally don’t extend suffrage rights to people who live in other countries (or non-citizens who live in our country for that matter). On this point, Walzer is at least consistent since he argues that such a restriction is in harmony with our understanding of national membership; the exclusion of non-citizens from voting is part of the nation-state’s distributive principle for voting, and so, it would seem, is the exclusion of customers or suppliers from voting in the firm. However, Dahl’s principle of affected interests cannot justify a claim to decision-making control only for citizens of a country and therefore cannot justify worker democracy. On the other hand, as a consequence, the principle would seem to imply that for the economic enterprise all patron groups would have to be given control rights in order for the hierarchy within the firm to be justified: workers, capital suppliers, customers, raw goods suppliers, stakeholders, local communities, and so forth. We thus either have to reject the parallel case argument as unsound, or accept it and follow it to its conclusion that there is no principled reason to grant workers rights over other patron groups.

2. Democratic Holism

A second major argument for the inherent fairness of the worker cooperative is what I refer to as the argument from democratic holism. This argumentative strategy differs from parallel case strategies because it does not see democracy as something practiced individually in analogous institutions, but rather as something larger than one particular practice or setting. Put differently, instead of saying that democracy exists in x therefore it should exist in y, democratic holism claims that
democracy as a larger goal requires certain substantive commitments and institutional prerequisites, which speak to the organizational structure of x, y, z, and so forth. One such institutional prerequisite, it is argued, is worker-democracy.

Michael Howard and Nien-he Hsieh have made one such argument within a Rawlsian paradigm. Their argument is that a well-ordered society governed by the principles of justice would require some form of workplace democracy. The difficulty both of these authors face, as was touched upon in the introduction, is that in the Rawlsian paradigm one must attempt to make such claims in a manner that is neutral to conceptions of the good life. For this reason simply claiming that participating in the governance of one’s workplace is a good in and of itself is a non-starter; if a worker would prefer to pass on a democratic control right in favor of higher wages or more leisure time, what would the basis be for preventing her?

This is the reason why Hsieh’s argument falters. Hsieh attempts to make a case for workplace republicanism built out of a Rawlsian commitment to the “social bases of self-respect,” arguing that such a commitment must also include a protection from arbitrary interference in institutions like the workplace.53 This implies democratic control, because simply exiting the corporation is too costly a check on arbitrary interference. But this runs into the problem of articulating a right against an institution that is outside of the basic structure, without violating liberal neutrality. In order for this argument to work one would need to draw on a moral notion of what is an acceptable cost of exit—that is, what would and would not be considered so great a burden entailed in exiting a corporation such that one could remain within its reach and still be subject to arbitrary interference. Yet how could one determine this within a political conception of justice? Rawls wants to claim that because the exit option is available for religious communities, the principles of justice need not apply there; this is an interpretation anyone concerned with religious freedom should want to retain. But surely one who

identifies as a Christian would see exiting the church as entailing very great costs (this is seen especially with homosexual Catholics and Mormons who face tremendous inequality and potential harm, but find the costs of leaving the church too great to bear). How do we distinguish the magnitude of exit costs in a way that is neutral with respect to conceptions of the good? It seems difficult to imagine.

For Howard workplace democracy is implied by the equal worth and fair value of liberties. Because some will be more subject to the decisions of hierarchical institutions outside of government, a commitment to the equal worth of liberty would demand that rights not be purely formal and directed toward the state; instead it would seem to demand that citizens possess rights against decisions made in those institutions outside of the halls of government that have a large impact on their livelihood. This would include the decision-making power of the corporation.\footnote{Michael Howard. 2000. \textit{Self-Management and the Crisis of Socialism: The Rose in the Fist of the Present}. Lanham: Rowman & Littlefield. 38.} The problem with this argument is similar to what we’ve seen with the parallel case strategy; there seems no real principled way of restricting such a right to workers. If we interpret the equal worth of liberties as implying checks against extra-statutory power (an argument that is dubious in the Rawlsian tradition because of Rawls’s emphasis on the basic structure\footnote{See my discussion of Rawls in the introduction.}), this would seem to imply a wide number of control rights given to a large number of parties against a seemingly infinite number of potential institutions.

Carol Gould’s argument for workplace democracy, despite its distance from Rawlsian theory, seems to fall into a similar problem. For Gould, the basis for workplace democracy comes from recognizing the unique nature of productive activity and the social dimensions of labor that distinguish it from other forms of acquisition (contra Nozick).\footnote{Carol Gould. 1988. \textit{Rethinking Democracy: Freedom and Social Cooperation in Politics, Economy, and Society}. Cambridge: Cambridge University Press. 138.} Gould leverages this sociological claim, in tandem with a principle of equal positive freedom that includes the right to conditions of self-development, to
claim that in a just democratic society there must be a “right to participate in decisions or choices concerning productive activities in which one engages jointly with others.” Gould recognizes that such a principle would seem to imply rights for a vast number of people, but claims a restriction of participation to those members of the firm is justified because they are the ones consciously engaged in a common productive project. Still, it is not clear why other patron groups of the firm—capital suppliers or raw-goods suppliers—are not similarly engaged in a conscious act of co-determined production and why they should be excluded from the right to participate. Gould seems to concede this point later in her career when she moves from worker democracy to an emphasis on stakeholder participation, including shareholders, suppliers, customers, financiers, local community, and perhaps even the public at large.

Even if we were to accept this problem of restriction and endorse a broader stakeholder democracy, it is not clear that we can get around the other problem we saw come to light in the parallel case strategy, that corporations exist to satisfy and pursue particular ends. That is, it’s not clear why we should be in a position to dictate the goods that employees or any other patron group should pursue once we note that people choose to associate with firms for at least partially material reasons. While we might accept that democratic control of the firm by all parties involved would be a good thing on one moral point of view, it is not obvious why the good of democratic control ought to be prioritized over, say, the good of satisfying material desires and necessities. Put differently, democratic holists emphasize the process at the cost of losing sight of the end. People don’t organize for the sake of organizing, and firms don’t simply exist for the sake of existing; if the material interests of individuals trump their interest in corporate democracy such that they are willing to allow others to make decisions, it is not obvious why they should not be allowed to do so.

57. Ibid, 143.
3. Workplace Eudemonism

The third argumentative strategy eschews the language of principles, rights, and entitlements that the other two arguments use, and instead argues for workplace democracy on pragmatic and consequentialist grounds: workplace democracy should be pursued because of the profound positive effects it has on society. This strategy goes back at least to John Stuart Mill and was given classic articulation by Carole Pateman. The avowed purpose of Pateman’s argument is to defend the participatory theory of democracy against the defaming done to it by Schumpeter. The latter contended that his now-famous definition of democracy—“that institutional arrangement to decide by means of a competitive struggle for the peoples’ vote”—was both empirical and value-neutral, and was therefore superior to classical theories of democracy that purportedly sought particular ends, tied up with something like a “general will” or “public opinion.” Whereas Schumpeter offered a minimalist and procedural idea of democracy that was dressed up like a merely factual report of what democracy actually is, Pateman attempts to offer a corrective: a theory of participatory democracy salvaged from Schumpeter’s critique, and shown to be viable both in practical implication and empirical analysis.

Pateman takes the tradition of participatory democracy to begin with Rousseau, and the importance she accords to participation is not due to the metaphysical qualities of the “general will,” but rather to the eudemonic properties associated with participating in the lawmaking process. Because institutions have a profound effect on individuals’ psychology, the goal of democratic theory is not just to articulate the institutional arrangement that will determine decisions, but also those that will help cultivate the features necessary for a free and self-determining polity. Participation is thus not a burden placed on the individual to help create a general will, or a privileging of the collective over the individual as Schumpeter implies. Instead, participating in the general will is what assures the individual’s freedom in the face of a social context that would otherwise dominate him: “the individual’s actual, as well as his sense of, freedom is increased through participation in decision
making because it gives him a very real degree of control over the course of his life and the structure of his environment.” The participatory view that Pateman offers, then, is one in which the development of citizens is not merely an input into the democratic procedure—as it is for Schumpeter and much of the political science literature that has followed his lead, including Dahl—but something that is affected, altered, and created by the democratic procedure itself.

If Rousseau’s account of democracy provides the origin for the participatory tradition, then the work of John Stuart Mill is the lynchpin in the tradition’s development into something applicable to modern society. According to Pateman, Mill’s concern for participation is similar to Rousseau’s and demonstrates the former’s distance from his utilitarian roots. “Mill distinguished two aspects of good government,” Pateman notes. “First, ‘how far it promotes the good management of the affairs of society by means of the existing faculties,’” or what we might call its utilitarian function of organizing things to take care of public business. However “the merely business aspect of government is the least important; fundamental is government in its second aspect, that of ‘a great influence acting on the human mind,’” or what we might call its virtue-cultivating or eudemonistic function. Like Rousseau, Mill’s arguments for participation are borne out of this second educative function of the government, its ability to develop character in a way a purely utilitarian approach to government cannot comprehend. What makes Mill of special importance, however, was his concern for participation in context of a modern industrial political and economic system. For this reason, participation in the actual process of lawmaking was both infeasible and undesirable. Instead, Mill tried to locate other venues for the educative function of participation to take place.

As we saw in chapter 2, one particular venue Mill was concerned with was industry, which Pateman picks up and runs with:

60 Ibid, 28.
Just as participation in the government of the collective interest in local politics educates the individual in social responsibility so participation in the management of the collective interest of an industrial organization fosters and develops the qualities in the individual that he needs for public activities. Just as Mill regarded democracy as inevitable in the modern world so he saw some form of cooperation as inevitable in industry.  

For Pateman, Mill represents an extension of the Rousseauvian democratic vision to the modern industrial world, but also an extension of it into the economic sphere. This view came as part of an entire system of political economy, thus complementing the normative force of his argument with strong empirical grounds to believe that it was possible and practical. For Pateman’s Mill, which is to say for Pateman, worker co-operatives provide an important educative function for democracy and do so by extending democracy to institutions not normally considered political, on account of the economic merits of doing so. The result would be a policy proposal calling for the transformation of authority within ‘private’ enterprises to one of more egalitarian, cooperative, and democratic structure.

Pateman’s approach to workplace democracy is strong, particularly her emphasis on how institutions affect the character of those involved in them. Considering the emphasis we have been placing on different types of incentives and cooperative schemes that exist within firms, an account of workplace democracy that focuses on this psychological malleability is an important one. However, the main problem here is that Pateman slightly misreads Mill. As we have seen from the previous discussion of Mill, for him the domination of capitalism by worker cooperatives was somewhere in between a normative claim and an empirical hypothesis. This is because while Mill thought there would be great moral value in the rise of worker control, he thought it only could and should happen if such cooperatives prove themselves to be efficient; the competitive market was ultimately the arbiter of economic institutions. While Pateman is correct to note Mill’s nuanced two-fold understanding of government’s function, her discussion of worker control strays from this insight by overstating Mill’s concern with the educative and character-inducing aspects of institutions at the expense of his concern

---

61 Ibid, 34.
for efficiency and material welfare. In chapter 6 I offer some idea of how Pateman’s eudemonistic account of workplace democracy might be put into practice legally and institutionally, while still taking into account the constraints imposed on the firm by the market.

**How, and to What Extent, Efficiency Constrains Democratic Scripts**

*Decision-Making Costs as Ownership Costs*

We can connect the problems of each of these arguments for workplace democracy with the foregoing discussion of corporate efficiency to better understand the subtle nature of corporate cooperation. For the parallel case strategist, the problems are establishing the analogy between democratic state and corporation, when the latter exists in a competitive market to pursue particular ends while the former does not, and limiting the citizen-analogy only to workers. For the democratic holist there is the problem of imposing a particular end upon which those associating with the corporation ought to pursue, as well as the problem of limiting democratic rights to workers. For eudaemonists like Pateman, the problem is being able to account for the worker cooperative’s efficiency as well as its fairness. Put together, while all these theories are sensitive in their account of the hierarchical relationships within the firm and the problems they raise, they are far less thoughtful when it comes to the economic nature of such enterprises.

In many ways these thinkers make the inverse errors as their Chicago counterparts. The Chicago theorists see the firm in overly individualistic and contractual terms, assuming that the existing organizational structures are the results of voluntary contractual arrangements, and that those people within the firm’s structure interact on purely material terms. As we have seen, however, people do not operate for purely instrumental and extrinsic reasons, and it is precisely the intrinsic motivations not offered by the market that firms capitalize on. To twist the phrase of a famous Genevan, the Chicago tradition has taken the laws as they are and imagined men as they might be.
The proponents of workplace democracy, on the other hand, seem to overlook the fact that people do have material reasons which bring them to want to associate with the firm in the first place. The fact that cooperative norms obtain within the firm does not negate the fact that people initially contract with the firm in order to enter this scheme of cooperation. It is therefore not clear why these material incentives can be so downplayed for the purposes of democratizing authority, when workers may prefer efficiency gains or extra wages in place of democratic control. This might all be stated more simply: the norm-based understanding of corporate production allows for, but does not demand, democratic decision-making to the extent that it allows for the firm to efficiently conduct business relative to a decentralized network of producers coordinating their relations through the price mechanism.

Whether or not cooperatives can be as or more efficient than their capital-owned counterparts is a debate which goes back at least to Mill’s affirmative hypothesis. Yet, as Rawls has pointed out, since the worker cooperative has not overtaken the capitalist economy, “nor does it show many signs of doing so, the question arises whether Mill was wrong about what people prefer.” Nozick’s take on this issue was, of course, that the historical record has proved Mill wrong. Nozick provides for three possibilities that could allow worker cooperatives to thrive. One is that worker cooperatives are as efficient as other firms, as Mill contended. Another is that though worker cooperatives are not as efficient as capital-owned firms, workers would be willing to work for lower wages on account of receiving the nonmonetary benefits of controlling their workplace; the cooperative would thus be able to be competitive because of the saved costs on labor. The third possibility is that cooperatives are less efficient, the laborers work at market cost, but the price of the good produced is higher than market;

This inflated price is borne by consumers who are willing to pay a premium for a good produced in a manner that accords with their sense of just labor practices.\textsuperscript{63}

The fact that cooperatives have not generally thrived would seem to dismiss all three of these possibilities for Nozick. That is, cooperatives are less efficient than their capital-owned counterparts, workers are not in general predisposed to bear the costs of this inefficiency in the form of lower wages, nor are there enough conscientious consumers willing to bear the cost in the form of higher prices. The result, according to Nozick, is that though the system is fully open to realizing worker cooperatives, workers would rather work at, and consumers would rather patronize, the more competitively priced capital-owned firm. In order to implement an economy of worker-owned firms one would have to forego gains from efficiency and contradict the preferences of most people (including workers). For Nozick it is not evident why this would be fair or just, let alone efficient, as Mill claims. In order for this argument to work, however, Nozick needs to dismiss one further possibility: that cooperatives are efficient and could survive competition but for the fact that exogenous constraints preclude them from forming. In order to stand, Nozick’s view would need to produce an empirical account of the law and markets to show that no such impediments exist, or provide some account of why worker democracy would be systematically underproductive. In the latter category two such theories present persuasive claims that must be taken into account by would-be democratic reformers.\textsuperscript{64}

The first argument has been most forcefully advanced by Henry Hansmann. The basic idea underlying Hansmann’s theory is that a firm has four main patron groups: capital-suppliers, raw goods suppliers, workers, and customers. Any one of these patron groups can interact with the firm either through the market—that is, by contracting explicitly with the firm or implicitly through law—or by

\textsuperscript{63} Nozick. \textit{Anarchy, State, and Utopia}, 250-252.

\textsuperscript{64} In chapter 6 I also review some contextual factors that make cooperatives comparatively less productive, and suggest different policy avenues that might overcome these factors.
“owning” 65 the firm, and determining the terms of engagement through rights of governance. If ownership rights are given to workers, customers, or raw-good suppliers, then it is a cooperative; if they are given to those who provide capital, then it is a standard business firm (which Hansmann suggests could just as easily be considered a “capital cooperative”); if ownership rights are assigned to nobody, then it is a non-profit firm. As we have seen, using the market comes with costs; Hansmann shows that ownership also entails costs. It is the balancing of the costs associating with the market, and the costs associated with ownership, that determines the allocation of ownership and decision-making rights.

Of the various costs and benefits he considers for determining ownership, the costs that Hansmann highlights are the costs of decision-making.66 These costs arise when there is disagreement amongst the decision-makers with regard to how certain goals should be achieved or, more seriously, what the goals for the firm should be. In general, these differences stem from individuals differing in the way they transact with the firm (for instance, a custodian and an executive in a workers’ cooperative) or in personal circumstances (for instance, differences in disposable income). When there are large differences amongst decision-makers, mechanisms for resolving difference will have to be employed. This results in two very important costs. The first is the cost of the decision-making process itself: the time, energy, and resources that are spent in resolving conflicts and executing the decision all represent costs that could have been foregone had the decision-making process been faster or less contentious. The second cost is the cost of poor decisions. This can happen either because a minority of the members were able to gain control of the decision-making process, producing an unfavorable

65 This word is put in scare quotes because, as I noted in the introduction, there are many scholars who plausibly deny the existence of ownership in the corporation, at least by any common understanding of what it means to own. I am sidestepping these questions here; whether or not the rights of residual claims or control are, strictly speaking, ownership rights, does not actually bear much on the argument in question.
66 Ibid, 39.
outcome, because better time-sensitive decisions were foregone in the lengthy decision-making process, or because quarrels amongst the owners creates the opportunity for managerial malfeasance.67

One can expect the cost of decision-making to be lowest when interests are less diverse within the group. Hansmann claims that this is why we rarely see many different patron groups jointly owning or running a firm, as it generates huge differences in interest and, consequently, an extremely inefficient decision-making process. What makes this such an important point for our purposes is that cooperatives tend to have more divergent interests within their ownership group than other firms. Hansmann observes that “[i]nvestor-owned firms have the important advantage that their owners generally share a single well-defined objective: to maximize the net present value of the firm’s earnings. The costs of collective decision making are thus relatively low for investor-owned firms.”68 Because in a worker cooperative the owners do not simply supply the firm with money, but also work in the firm, they cannot rely on value maximization as the uncontroversial metric for decision-making; owners will also be concerned with things like workplace conditions, task division, team coordination, etc. This becomes a more pronounced problem in firms where the labor force is heterogeneous. When there are workers doing a variety of tasks there ceases to be a straightforward single good which all share. The co-op needs to take into account the concerns of the machine-operators who want funds spent on safety and the concerns of the custodians who want funds spent on newer equipment; but they must also take these concerns into account equally because of the one-member one-vote rule.

This makes sense of the empirical finding that most successful cooperatives either have a workforce where tasks and work setting are fairly constant across the entire group (for instance, travel agencies) or have work rotation systems so that every owner is put in the same position as every other owner. While there are many examples of successful coops, it is important to note that these criteria

67 Ibid, 40-42.
stack the decks against cooperative ownership, particularly in enterprises that require a significant
division of labor or the development of multiple types of specialized knowledge and skill sets.
Furthermore, those who would want to follow Gould’s lead in endorsing a stakeholder conception of
the corporation face an uphill struggle against the vast decision-making costs such a norm would
introduce.

Because of all this, decision-making processes will often be more expensive for cooperatives
than they will be for other firms. However, the more serious cost that decision-making entails are the
costs of poor decisions. The time and energy, or even the money, spent on democratic decision-making
might not be deal breakers for the reasons that Nozick gave—workers might be willing to bear these
costs for the ability to democratically control the workplace. However, poor decisions, or non-
decisions, are harder to justify by a high preference for democratic control since democratic control is
the very thing causing the poor decisions to be made. When these costs make it difficult for
cooperatives to effectively make time-sensitive decisions that coordinate a work force and mobilize
specialized information, we can expect cooperatives to underperform relative to capital-owned firms.
These are the decision-making costs that are a major obstacle for worker ownership.

Perverse Incentives

The other argument against cooperative efficiency is that worker ownership creates incentives
for perverse responses to price signals. As Benjamin Ward observed\(^\text{69}\) the problem that cooperatives
face is that their productive capacity is tied to ownership shares. Keeping technology and innovation
constant, greater production requires a greater number of workers; however, a greater number of

\(^{69}\) Interestingly, Pateman cites Ward in her discussion of Yugoslavian worker councils, but does not cite this canonical
article, which is decidedly less optimistic than Pateman is: Benjamin Ward. 1953. “The Firm in Illyria: Market
Syndicalism,” American Economic Review. 3. For work which expands on and explains Ward’s observation, see J.E Meade.
workers also means that there must be a greater number of owners and, hence, more people who get a share of profits. Ward argues that because of this basic fact of worker-ownership, cooperatives tend to hire less when they ought to produce more, and hire more when they ought to be producing less. When the market is showing that the price of their good is high, the cooperative has an incentive to contract since it would allow for each of the members to reap a higher share of profits. On the other hand, when prices are low, cooperatives face the incentive to bring on more workers since that would mean a greater number of people to share in losses and, in the cases where owners have to put down capital to buy their shares or put some percentage of their earnings toward this, it would mean an injection of capital into the enterprise. Regardless, the end result is that cooperatives have a tendency to under-produce in times of relative scarcity, and overproduce in times of relative surplus.

This argument is a bit weaker than Hansmann’s argument. For starters, Ward assumes a simplistic decision-making process where all decisions are made by majoritarian worker plebiscite. In reality most worker cooperatives have various institutional checks on these types of incentives, including the hiring of managers, insulated boards of directors, and enterprise by-laws which enable the firm to check short-sighted majoritarian decision-making. Still, while this argument (or the decision-making costs argument for that matter) is not unassailable, it does point to general trends that raise issues for worker democracy. Both of these arguments imply that some limit on participation must be introduced, that a managerial staff somewhat insulated from worker control be empowered to make decisions for the long-term and with regard to short term price-information, and that developing a differentiated labor force will come at the cost of decision-making efficiency.

The more general point is that although the cooperative nature of intra-firm governance opens the possibility of a normatively thick form of cooperation, the nature of the market puts some constraints on this. This does not mean that things like cooperatives are inherently unviable. What it does mean is that some cooperatives will come at the expense of efficiency and perhaps at great expense. The result will either be frustrated preferences or simply the exiting of people from the
cooperative. As was said earlier, firms do not just compete in the market, but they also compete with the market itself. Firms can disband, contract and outsource, or wither away, and thus must test their mettle by providing something in a way that the market cannot. This is perhaps the most important dis-analogy between firms and states, more important than the point about monopoly or competition: firms do not need grand “foundings” to start nor civil wars or revolutions to destroy. Firms can disappear simply by virtue of not serving the purpose they provide for others better than some existing alternative. To increase the viability of worker cooperatives, what are needed are institutional complements that enable cooperatives to economize on market transactions as well or better than standard shareholder oriented corporations.

**Normative Possibilities within Efficiency Horizons**

What Ward and Hansmann are doing essentially is extending the Coasian insight to explain why firms are owned as they are. People will associate either through market contracting or through ownership; what they choose will be based on which allocation of ownership rights best reduces transaction costs and costs of ownership. Put differently, ownership rights are assigned to the patron group that can achieve the “least-cost assignment of ownership”:

The least-cost assignment of ownership is therefore that which minimizes the sum of all of the costs of a firm’s transactions. That is, it minimizes the sum of (1) the costs of market contracting for those classes of patrons that are not owners and (2) the costs of ownership for the class of patrons who own the firm.70

What this amounts to is the grand claim that control rights will be assigned to that patron group which best reduces transaction costs for all patron groups. In fact this amounts to an even grander claim: that such ownership assignment is most beneficial to all of society because it contributes toward an efficient economy.

This would appear to push us back to Friedman. If the market for ownership leads to the least-cost assignment of ownership, then we ought to accept the idea that the main imperative for firms is to benefit that patron group, and accept shareholder primacy for most business corporations. If other patron groups wanted some other arrangement, then they would have bargained for it. The fact that we see so many shareholder owned corporations should lead us to believe that people generally prefer the efficient ownership of shareholders, and generally prefer to associate as customers or workers through lawful contracting.

While I think Hansmann’s point about decision-making is a real challenge to workplace democracy, and illustrates the types of ways that cooperative scripts are constrained by the competitive market, this last claim is far less compelling. As Heath has pointed out, this argument assumes that all parties are able to bargain effectively and without constraint. That is, the assignment of ownership rights to one patron group can only be considered “least cost” if doing so increases the net-benefit to all patron groups more than it would under any other assignment of ownership rights. This means that assigning ownership not only decreases net costs for each patron group, but it decreases net costs for all other patron groups taken in the aggregate. If other patron groups would benefit more under a different ownership structure, Hansmann wants to say, they would have bargained for it. This only makes sense if we assume that shareholders not only outbid each other patron group individually, but that they outbid potential coalitions of competing patron groups. This, in turn, only makes sense if we think that patron groups are able to cooperate and form a coalition to bargain effectively. Yet, it is precisely this assumption that Hansmann cuts himself off from by virtue of his argument about the costs entailed in decision-making by groups with heterogeneous interests: “if Hansmann’s central argument…is to be
believed, these transaction costs will be quite high, precisely because they involve two constituency groups with radically different interests."  

Put another way, were the assignment of ownership such a straightforward bargaining situation where there were no collective action problems or transaction costs, we would want to know why the firm needs to be owned at all: why not just transact through the market? The fact that firms exist to overcome precisely these costs leads us to believe that the bargaining process that leads to ownership assignment will be imperfect. Although ownership assignment will tend to represent the least cost for a particular patron group, it does not follow that all groups would not benefit collectively under some other ownership structure. Other patron-groups might have benefited more from a different structure, but because of the costs in coordinating groups of divergent interests, they are unable to bargain in concert to bring this about. They still benefit under the arrangement—otherwise they would just become private contractors and contract through the market—but we cannot assume that this is the most beneficial arrangement possible.

Heath pushes back against Hansmann’s argument here to argue against a pure “profit-maximization” norm for business ethics and for his market failures approach. I will consider this view of business ethics in chapter 7. For now, I’m interested in this insight because it shows us something important about the space for norms and cooperative scripts within the corporation. If Hansmann’s

---


72 To the best of my knowledge Heath has not extended his analysis to do this. He has, however, used a very similar type of argument—specifically the distinction between the “core of the game” and the “feasible set”—to explain how normative restrictions can enter into insurance schemes. He hints that this logic might be extended to the firm: “There will be a region in which the welfare benefit generated by expansion of the pool outweighs the welfare loss occasioned by the monetary disadvantage imposed upon the other policy holders. As a result, it can be in the interest of everyone in the to accept new members, even when these new members are charged less than the actuarially fair premium (just as it can be in the interest of firms to hire more workers, even when doing so depresses average output. It is only when the marginal gains in net output reach zero that the firm should stop hiring).” (Joseph Heath. 2006. “Reasonable Restrictions on Underwriting.” Research in Ethical Issues in Organizations, Vol. 7, eds. Patrick Flanagan, Patrick Primeaux and William Ferguson. Amsterdam: Elsevier).
more general point was true, then there would not be much space at all for introducing values like fairness, equality, or democracy into the corporation; cooperative scripts would be introduced but only for purely strategic reasons. The nature of the industry and market would basically predetermine the structure of the corporation. What looks like cooperation organized according to certain values—worker cooperatives, for instance, as the practice of democracy or solidarity—is actually just an efficient allocation of ownership rights; to treat intra-firm cooperation as actually being concerned with thicker normative values, and trying to organize them in such a way, would be to miss the point, and ultimately lead to fatal inefficiencies.

Yet what the distinction between patron groups taken singly and patron groups taken in aggregate shows is that there are a number of possible arrangements between the least efficient a firm can be—that is, not coordinating production better than the market—and that arrangement which benefits all parties the most.\(^73\) As was noted before, we can imagine a situation in which non-shareholder patron groups might more greatly benefit under something other than shareholder ownership, but accept shareholder ownership because it still effectuates cooperation in a way that the market could not. This arrangement is perhaps not as efficient as it might be, but is still efficient enough for the firm to remain viable. There is therefore space for normative concerns to trump efficiency in certain respects while still enabling the firm to function more efficiently than the market.

This can be illustrated with a different sort of example. Intra-organizational studies find that wages tend to be more equal within firms than between firms.\(^74\) This fact plays out in practice most

---

\(^73\) Or, as Knight noted, “there is generally more than one way to structure social institutions in order to produce gains from cooperation, coordination, or exchange. And the major distinguishing feature of these different institutional forms is their distributional consequences” (Institutions and Social Conflict, 26).

noticeably in corporate mergers, when pay scales need to be refigured to reflect the higher wages of one company despite the comparable positions. The result is that when two corporations merge, two distinct corporate cultures, and metrics of fairness, need to be reconciled into one. Wages within firms do not correspond purely to what an individual with a given set of skills or training would command in a market, as we’ve already seen, but it seems that these wages are significantly influenced by interpersonal comparisons of wages within firms; wages are not determined purely through the amoral calculations of what an employee’s marginal contribution to the firm’s net product will be, but also what an employee deserves based on the nature of that particular firm. The egalitarian tendency of wages within firms is the result of intra-firm redistribution where those who would be earning less in a competitive market are subsidized by those who would be earning more, so as to secure the mutually beneficial result of cooperative team production that would elude all parties in an open market.

Now, the logic of Hansmann’s argument would suggest that what looks like equality within the firm here is just an accident of what is actually just the most efficient organizational form. Tinkering with this equalization on behalf of an explicit concern for intra-firm equality—say, trying to establish a more exact fair pay ratio—would lead to trouble. But in fact, the pursuit of equality within the firm is constrained only insofar as pursuing it does not push those bearing the brunt of the redistributing to exit the firm. Firms can employ norms of equality to govern their wage structures and internal relationships as long as individuals are not being asked to subsidize others to the point where they are benefiting less than they would from private contracting; if they were asked to do that they would simply leave the firm and revert to market contracting. Furthermore, because of the transaction costs that bring the firm about in the first place, certain arrangements that might theoretically be more preferable to the firm—

arrangements where a subset of people start their own smaller firm—will not be readily achievable because they entail startup and coordination costs. There is therefore quite a bit of room for a thicker conception of normative cooperation to enter into the relationships of the firm while still being mindful of its efficiency mandate.

We might sum all this up in this way. Markets fail to produce all possible cooperative arrangements because of the kinds of incentives the market offers. To remedy this, firms are developed. Firms effect efficiency gains by employing intrinsic incentives in the form of shared norms and scripts of cooperation, which mitigate the problem of opportunism and thereby enable various parties to cooperate with one another in ways markets don’t foster. Firms therefore ought to be seen as islands of consciously normative cooperation in a sea of “more or less norm-free sociality.” Yet, the purpose of the firm and its competitive market environment impose inherent limits on the thickness of this normative cooperation. The norms should not be so thin that they cannot overcome the problem of opportunism or provide a better alternative than the market pricing mechanism: we can think of this as the minimal normative horizon. However, the norms cannot be so thick that they impose more costs on the firm’s members than they would otherwise bear had they used the market pricing mechanism: we can think of this as the maximal normative horizon. Between these two points there is space for values and arrangements that bring normative values other than efficiency to bear. This is the space within which a normative component of a theory of the corporation ought to affect, according to this norm-based understanding of corporate productivity. Furthermore, because of preference endogeneity, path dependence, and “habitual lock-in” there is good reason to worry that patterns of arrangements may very well reproduce themselves without much reflection on the part of their participants. In order to protect against the development of problematic workplace arrangements we should want such normative principles to be consciously applied and implemented.

So, while there is nothing which makes worker cooperatives inherently inefficient, what Hansmann effectively shows is that there is good reason to think that variety of ownership and
decision-making structures is necessitated by a concern for efficiency. To assert the necessity of worker
democratic control of the firm is to assert a particular script of cooperation (in this case, one of
democratic legitimacy) without a thought to corporate efficiency. If all workplace democrats want to
argue is that we ought to *support* the development of productive worker-cooperatives, which we might
do through freeing up access to start-up capital or by developing insurance products that protect
workers against certain types of corporate bankruptcy, then this is not terribly objectionable.\(^7\) To
argue, however, that the corporation ought to be uniformly democratized by legal fiat, even at the
expense of efficiency, is to put the normative cart ahead of the conceptual horse. As we have seen, the
corporation exists precisely to overcome the efficiency losses created by transaction costs and
opportunism. To claim that democratizing the firm ought to trump all ignores that without being
concerned with efficiency there is no reason to have a firm at all: people can just contract privately in
an open market.

Advocates of this stronger argument for workplace democracy are helping themselves to the
normative possibilities opened up by corporate institutions and their use of cooperative scripts, but in
the process they lose sight of why those possibilities were opened up in the first place. The fact of
norm-oriented production and explicit cooperation does not imply the corporate reformer has carte-
blanche in organizing the corporation; whereas democratization might be a trump card for liberal
democratic scripts of political cooperation, it is not obvious it functions as a liberal trump within the
corporation. The efficiency constraints under which the firm operates—its various bottom lines, we
might say—don’t simply impose market discipline, they also impose a maximal normative horizon on

\(^7\) Malleson makes a compelling case for this softer support of worker cooperatives. Furthermore, he also provides the most
up to date and extensive review of the literature on cooperative performance. see Tom Malleson. 2013. *After Occupy:*
6.
the types of cooperative scripts that can be used to structure team production, narrowing the set to those which enable the team to efficiently operate.

Another way of phrasing it is to say that the radical democrats overstate the difference between the unchecked authoritarian boss and the democratic workplace. Distinct as these two ideas might be, they are actually more similar when compared to the anonymous market in which the firm exists. Both the egalitarianism of the democratic workplace and the authority of the hierarchical workplace reject market anonymity and make use of cooperative scripts, norms, and roles in order to create the possibility for intentional cooperation. Instead of placing hierarchy on one side of the spectrum and equality on the other side, it is better to see both of these ideas on the same side of a spectrum, the party-relevant, cooperative-normative side, in contrast to the market, which would be on the party-irrelevant, adversarial-anonymous side of the spectrum. Both are able to work because they replace instrumental modes of interaction with non-instrumental ones—they achieve the normative minimum of corporate productivity; however, the corporation itself cannot be devoid of instrumental reasoning since it still operates in the “norm-free sociality” of the market. All this means that the cooperative schemes and normative cores introduced in order to overcome market failure must still have economic salience—they cannot exceed the normative maximum of corporate productivity.

Although I have articulated a different account of corporate productivity (summarized in table 4.1 below), and I have dismissed the Chicago School’s normative component of the theory of the corporation, I have not yet offered an alternative normative component. As the case of workplace democracy has shown, the impulse to apply off-the-shelf principles of fairness and organization can inappropriately downplay relevant features of the corporation, the most important of which is the corporation’s function of achieving efficiency and its place in a competitive market. As a result, each question of the normative component will have to be answered keeping in mind the minimal and maximal normative horizons, how the efficiency of the corporation both requires and limits the forms of cooperation existing within the corporation.
Table 4.1 The Productivity Components of the Chicago School and the Norm-Based Model

<table>
<thead>
<tr>
<th></th>
<th>Question of Efficiency</th>
<th>Question of Constraint</th>
<th>Question of Constitution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chicago</strong></td>
<td>Overcoming of transaction costs by incentivizing behavior</td>
<td>Principal-agent problems resulting in bureaucratic costs</td>
<td>Nexus of contracts</td>
</tr>
<tr>
<td><strong>Norm-Based</strong></td>
<td>Overcoming transaction costs through cooperative scripts, norms, and extrinsic motivation</td>
<td>The non-scalability of cooperative schemes</td>
<td>The realm of norm-oriented cooperative production</td>
</tr>
</tbody>
</table>

However, by way of prelude to that discussion, we can conclude in this way. The proper way to view corporate efficiency from a normative perspective is as presenting both a problem and an emancipatory possibility. The problem, which the workplace democrats address head-on, is that the norms generally used to govern the enterprise are ones which entail authority, hierarchy, and domination. This is ironic of course because the relationships legitimated by these norms are precisely those most offensive to liberal capitalism. In this sense the corporation is regressive, undoing some of the good done by liberal society by undermining the liberal ideals of equality, individuality, and freedom. Thus, while these types of authority are not of the market, per se, they are abetted by a market logic, which renders such authority invisible from the perspective of economic theory.

Yet the corporation’s use of norms (as opposed to prices and consequences) also offers us the progressive possibility of coordinating our behavior through different, more egalitarian norms. Thus while workplace democrats are often thought of as critics of the corporation, calling for its radical dismantling, it is actually better to say that they are attempting to harness its normative potential for social-egalitarian aims. We might say that the norm-governed sociality of corporations can cut both ways, either toward repression or toward emancipation. If we wish to make use of this emancipatory potential, however, we cannot do so at the expense of that which offers it: the demand for the corporation to function more efficiently than the price mechanism which orients its interaction with other firms and consumers.
In figuring out how to address the normative questions, then, the questions should not be answered by thinking about how the corporation resembles other institutions that operate in certain ways. Instead we should consider how the norms being used to structure corporate production create relationships which might harm or hamper the liberal democratic subject. That is, following Smith, Coase, Pateman, and others, we recognize that institutions don’t merely constrain subjects in terms of their actions, but can also alter their motivations, behaviors, and capacities. Because institutions that use scripts can breed their own consent, we avoid the implied consent entailed in Chicago’s contractual approach. If we are concerned with how the structure of the corporation can be reconciled with the ideals of liberal democracy, however, the concern should not be how to make the corporation look more like a democratic state. As we shall see, formulating answers to the normative questions will imply drawing on theories up to the task: relational ideas of rights, law, and politics in order to answer the corporate law question; understandings of the relationship between background institutional environments and internal organizational structures to answer the corporate governance question; and philosophical theories that take seriously the way institutional constraints affect moral duties, in order to answer the business ethics question.
Chapter 5. Relational Corporate Law

As we have seen, the most compelling account of corporate efficiency is one which takes seriously the norms and scripts that structure corporate relationships. This is in contrast to the Chicago approach where corporations are markets by another name, nexuses of contracts given names and entity status so as to minimize the costs of contracting. In the latter view, as we have seen, corporate law is essentially a subset or extension of contract law. For this reason, absent particularly egregious acts of deception, coercion, or self-dealing, corporate law is to be fairly permissive with regards to corporate action and organization; if parties have contracted to the extant corporation, then we should largely see it as preferred to some other alternative; otherwise why would the parties have contracted in such a manner? Corporate law therefore prescribes a set of default forms of organization that are generally perceived to be most efficient and generally preferred; this makes corporate contracting more efficient, but does not constrain parties from arranging their corporation in other ways (with attenuated profit-motives, cooperative ownership structures, dual-stock classifications, etc.). To make corporate law more normatively thick, as it were, with a more onerous conception of fiduciary duty, greater specificity of prescribed governance structures, or disclosure requirements is, according to the Chicago school, to commit a category error by importing ideas from public law into private law, or from social and moral philosophy into a domain that must chiefly be concerned with efficiency.

The “norm-governed productivity” approach that I have articulated concedes the significance of efficiency for the corporation. However, as I suggested in the last chapter, by contending that the corporation’s efficiency is tied to its ability to facilitate particular types of relationships—relationships unattainable through market contracting because they require greater normative substance—I have opened the door to let in precisely the sort of normative considerations that the Chicago school wishes to exclude. That is, once we recognize that the corporation introduces norms and scripts that structure people’s preferences and attitudes in such a way as to facilitate particular types of cooperation, we cannot simply accept the idea that corporate contractual relationships reflect the contracting parties’
interests since these interests are affected by the very corporate institutions that are under dispute. Contractually-revealed preferences are not a strong enough basis to justify a particular corporate structure or decision. Instead, we must look at the types of relationships and norms that are used by the corporation to structure its productive relationships. Assessing these relationships requires normative values, explaining how such structures are consonant with basic principles of individual liberty, moral equality, and democracy.

As we saw in the last chapter, some have taken this on full-scale, contending that we ought to treat corporations as if they were mini-states, and structure them like our democratic and constitutionally bound governments. I have argued that they overstate the case by ignoring the firm’s efficiency mandate. One must assess the relationships of the corporation within the context of a market economy that is meant to promote efficiency and that punishes business organizations that don’t. Though corporate law can constrain corporations according to values beyond efficiency—indeed, in order to be legitimate I think it must—the corporation will only be able to embody such values in bounded ways, because of its first-order role in promoting efficiency. As I’ve argued, the “norm governed productivity” of the firm entails minimal and maximal normative frontiers; norms must come into play at some basic level in order to achieve the efficiency gains, but norms and scripts cannot be imposed in such a manner that patrons of the corporation would rather privately contract than associate with the firm. In this chapter I propose a way of assessing corporate law in this regard, and go through some of the consequences this has for the theory of the corporation.

Relational Corporate Law

The Relational Approach and Liberal Values

To be able to think through how these overarching liberal norms ought to bear on the legal and governance structures of the corporation, I take on a “relational law” approach to the questions of corporate law and governance. This approach appears particularly appropriate based on the account of
corporate efficiency given in the last chapter, because the ability to normatively structure cooperative relationships within the firm is one of the comparative advantages of corporations over the market. While there are various strands of the “relational approach to law,” the basic idea is that in legal analyses individuals should not be abstracted from their social contexts. Instead, we must understand an individual as “constituted and shaped in part by relationships important to a person’s identity.”¹ To approach questions about, for example, the autonomy or dignity of an individual, then, is to necessarily consider how relationships contribute to or detract from such values, and how different structures of relationships might alter them.

This claim about individuals and their relationships is drawn out into a descriptive and normative claim upon the law. Descriptively, relational law maintains that “what rights and law actually do, right now, is structure relations,”² which in turn affect individuals and their actions. Normatively, this leads to a claim that the law ought to be more attentive to these relationships: legal analyses should take into account how relationships affect people’s actions when considering questions of liability or negligence, and should strive to structure relationships in ways that promote individual flourishing. As Nedelsky puts it:

One should ask how existing laws and rights have helped to construct the problem being addressed. What patterns and structures of relations have shaped it, and how has law helped shape these relations? The next questions are what values are at stake in the problem and what kinds of relations promote such values? In particular, what kind of shift in the existing relations would enhance rather than undermine the values at stake?³

Put differently, the relational approach to law starts with the empirical presupposition that “patterns and structures of relations” play an important role in bringing about particular outcomes and results. We therefore start by considering what relationships are significant for understanding the particular

³ Nedelsky 74.
phenomenon in question. Then, based on the presupposition that the law structures relationships, we look into how existing legal structures affect these relationships. If it appears that the result has been contrary to a value we hold dear then we would want to consider what changes in relationships would better fit that value, and how the law might intervene to promote such changes in relationships.

The aim here, then, is to use this analysis to consider how we might restructure corporate law to better cohere with the liberal democratic values asserted earlier—individual liberty, equality, and democratic decision making. This might appear surprising. The relational approach, historically, was derived from feminist critiques of liberalism, particularly how the liberal assumption of an atomistic individual and the public-private dichotomy covered up structures of female oppression and was contrary to the experiences of women for whom atomism and the public-private dichotomy were alien.4 Without denying these criticisms of liberal theory, I want to argue that the relational approach to law need not be so at odds with liberalism. What the idea of the “relational self” shows is that if one is concerned with any particular social value—including liberal values—one cannot afford to ignore how relationships contribute to or detract from their achievement. Again, Nedelsky is worth quoting at length, as she captures this idea in helpful ways:

> It is relationships that enable individuals to flourish and develop the capacities that liberalism has long highlighted—reason, autonomy, liberty—as well as the capacities for love, play, and emotion that make life valuable and that have not received the same attention in the liberal tradition. Relationships must therefore be central rather than peripheral to legal and political thought and to the workings of the institutions that structure relations. I think a relational approach makes it more likely that the formal commitment to equality will become a reality in ways that it has not been in either the origins of liberalism or its contemporary practices. While I share the commitment to the inherent value of every individual, I think the individualism of the liberal tradition impedes practices and habits of thought that can give effect to that equal value.5

In this view, the relational approach can be conceived as an internal critique of liberal habits of thinking. The liberal tendency toward imagining abstracted and atomized individuals leads to habits of

---


5 Nedelsky, 86-87.
thinking that occlude the relationships that are so important to what liberalism seeks to achieve. To better effect the individual freedom, equality, and democracy of liberalism, then, requires concern for the relationships that make such values achievable.

Relationality and Contract

The relational approach brings into view a salient feature of the law as it relates to economics. The assumptions of rational action and revealed preferences that underlie Chicago economics, and contemporary corporate law, rest upon a conception of the abstracted individual. Because of the way this framework conceives of individual choices and actions, it ends up structuring relationships in particular ways. As we have seen, Chicago’s nexus-of-contracts approach counsels permissiveness and deference toward extant corporate structures, since it takes them to be the result of individual voluntary contracts. Descriptively, we can simply say that such a view structures relationships within the corporation. The relationship between manager and employee is rendered instrumental and hierarchical; the relationship between manager and shareholder is rendered abstracted (such that the manager need not pursue shareholders’ interests as they actually are, but as they are presumed to be—maximizing profit⁶); and the relationship between the manager and society-at-large is rendered outside the scope of his role as corporate manager. By excluding a substantive evaluation of relationships from its legal analysis, the nexus of contracts view enables all sorts of relations to develop that might not cohere with our social commitments.

Durkheim captured well the flaws with this economic view of the contract as a spontaneous and self-validating form of human interaction. Durkheim acknowledges the importance of the economist’s

insight that human behavior and relationship will tend to follow certain law-like patterns and that social life normally “arises not from arrangements imposed from without, but from its free internal nature.”

The contract, posited as being the result of such free and spontaneous action, becomes “the primary fact from which all other social facts have derived.” Despite the importance of this contribution, Durkheim claims this view is fundamentally flawed because it conceives of this freedom as being naturally given and deduced from the individual. Instead, this free domain of contractual action must be understood as itself being created and enabled by a web of social actions and relationships.

As Durkheim puts it, “in a contract not everything is contractual.” While contractual relationships entail the voluntary consent to a certain set of arrangements, responsibilities, and duties, contracts also call forth a number of non-consensual factors that are external to the contracting parties. The most obvious is the regulatory force of compulsory institutions. Contracts, like “property,” “markets,” and the wider economic system of which they are all a part, thus require social institutions like states, regulatory agencies, and courts to have any force. The result is that when we enter into a contract we are obliged not merely to do what we assented to, but also to a range of duties that are imposed by state-enforced law of contract.

The social embeddedness of contract is not limited to legislative enabling, judicial recognition, or statutory regulation. Additionally, contract is shaped by social norms and morals that both specify the meaning of contractual terms, and constrain and direct the range of viable contractual relationships. These range from social and linguistic conventions that give contracts particular meanings beyond what is expressly put in them, to the moral and professional obligations that are exerted by our colleagues.

---

9 Durkheim 1984, 158.
and peers that direct how we discharge our contractual duties. For Durkheim, this all flows from a view of contract as being a functional prerequisite for a system of organic solidarity to develop; the contract becomes a pervasive instrument for effecting cooperation only after a certain degree of differentiation and division of labor occurs. The institution of the contract exists to achieve certain social ends of mutual advantage and therefore, as a result, is subject to more than the mere mutual assent of contracting parties. Contracts, therefore, must be analyzed and assessed with these contexts and ends in mind, which, for Durkheim, is a form of solidarity grounded in social justice.

It would follow, then, that viewing the corporation as a “nexus of contracts” does not necessarily imply that it is the product of voluntary arrangements; instead, it would be a nexus of particular institutional arrangements that are on the one hand voluntarily instigated, but on the other hand brought to bear by social institutions and norms. In one sense the Chicago view does not actually differ so much in orientation from this Durkheimian account. In the economic view, contract law exists not to uphold the promises of individuals, or some other expression of individual volition; it exists to effect efficiency gains. For Chicago scholars, however, there is little difference between the individual volitional arrangements and maximizing efficiency. That is, the Chicago view agrees with Durkheim that contracts are socially instrumental; the difference is that they see contract solely in terms of efficiency instead of justice or solidarity, and they see this efficiency as best achieved through the spontaneity of elective and voluntary cooperation in the market. The social good of the contract is best achieved by keeping social action as minimal as possible, hence the libertarian nexus-of-contracts corporation.

10 Ibid, 162.
Taken together, the norm-governed productivity of the corporation and the relational approach to law lead us more toward the Durkheimian view than the Chicago view of the firm. As we have seen, the corporation is a type of organization that is able to cultivate productive relationships because it can induce norm-oriented behavior through intrinsic incentives that are precluded by the market’s competitive, consequence-oriented behavior. That is, the efficiency of the firm is achieved by bringing in other normative values and scripts to enable non-market cooperation. To assess intra-corporate relationships as efficiency-oriented contractual relations misses how the efficiency of the corporation differs from that of market transactions. The result is that to assess the relations within the corporation, efficiency should not be the sole or even the primary value we use. The legal contours of the corporation, in this view, must be assessed based on the relationships they structure and how well these match up to social values other than efficiency.

One consequence of this approach emerges if we look at the literature on what legal scholars refer to as “relational contracting.” The idea of the relational contract is distinguished from the contractual transaction in that instead of delineating responsibilities and duties in some particular instance between two discrete parties, it structures a pattern of relationships. As Macneil puts it, the distinction between discrete transactional contracts and relational contracts is that “although both involve economic exchange, only the latter include whole person relations, relatively deep and extensive communication by a variety of modes, and significant elements of non-economic personal satisfaction.”¹⁴ These are of course ideal-types: many discrete transactions entail aspects of relationships, and contracts that structure relationships will still entail particular instances of exchange.

Still, this distinction is helpful in understanding the relationship between corporate governance and the kinds of relationships that fit with social values. Even if we concede to Chicago that there is an important contractual element to the corporation, the concept of the relational contract draws our

---

attention to the different types of contracts in a given corporation. Specifically, what we see is that the shareholder contract more closely resembles a discrete transaction whereas the employment contract more closely resembles a relational contract:

The enterprise in which people work must be viewed as a firm with many members. Particularly with a publicly held corporation, the traditional “owners,” the shareholders, have no power, no control, and a bond with the company that can be broken by a call to their broker. They are not even providers of capital, except by buying stock that a thousand trades ago was issued to an actual investor whose money made its way to the company treasury….The people who do have a bond to the publicly held corporation, and who have made a real investment with it rather than in it are management and workers, as well as the localities that have grown so dependent upon it through years of symbiosis.  

Linzer draws a normative legal conclusion from this observation: “employees and others who have invested their lives or their environment also should have rights in the firm, deriving not from traditional contract, tort, or property, but from their contribution to the firm and their close relationship with it.” This has traces of republican worries over non-domination, the kind that we saw in the democratic holists of chapter 4: because certain people have made deep investments into the corporation, they are more subject to its effects and therefore ought to have rights in the corporation as a form of protection.

Now, as I have suggested, such an argument goes a bit too far. Still, it points us toward an important fact about corporate law from a relational perspective: by treating all corporate contracts as purely discrete we ignore the ways in which some parties relate more deeply and personally to the corporation. In general—though not always—this will include employees. Though this does not lead necessarily to the idea of a republican workplace or worker democracy, it does suggest that corporate law ought to take such concerns into account. While workers might not be given a privileged place in corporate law, they likely do require special consideration and concern.

16 Linzer 191.
What is the Corporation?

From Individuals or States?

Given all of this, we can ask the most basic question: in my view, what exactly is the corporation, legally speaking? This question is important to answer not just because it is standard fare in corporate law classes and popular debates on legal theory and policy. The manner in which we answer this question can have the effect of biasing our normative conclusions. Currently, debate centers largely around two answers to this question. The first is that corporations are aggregates of individuals, or what Eric Orts calls a “bottom-up” approach to the corporation. In this view corporations have no presence or reality other than their component parts, that is, the individuals incorporating with one another. As we have seen, this is the genre of theory advanced by the Chicago School in their contention that corporations are merely legal fictions standing in for the nexus of contracting parties. Also as we have seen, this view of the corporation leads to the view that the state should be permissive, generally deferring to the private orderings that individuals have contracted into.

This view has further normative upshots, one of which has become (in)famous in recent years. The majority decision of Citizens United, which found unconstitutional legislative limits on corporate donations to electoral campaigns, rests on an aggregate, bottom-up theory. To summarize Kennedy’s reasoning: individuals have freedom of speech, which protects their freedom to donate money to electoral campaigns. Such rights can only be abridged by a law that serves compelling government interest, and only if such law is narrowly tailored and minimally restrictive (that is, a law that survives “strict scrutiny”). Corporations, being aggregates of individuals who hold the right to free speech, therefore have the similar freedom to donate to electoral campaigns, absent a law that survives strict scrutiny. Thus, a corporation has the right to free speech by virtue of its composite parts (rights-bearing individuals) having the right to free speech. Put differently, the corporation in this view is merely an

---

outgrowth of individuals’ freedom to associate, which is given effect by virtue of those very same individuals’ right to speech. As Kennedy puts it: “wealthy individuals and unincorporated associations can spend unlimited amounts on independent expenditures….Yet [under then-extant campaign finance laws] certain disfavored associations of citizens—those that have taken on the corporate form—are penalized for engaging in the same political speech.”18 Just as the aggregate, bottom-up view of the corporation tends toward a permissive corporate law because individuals are understood as having the freedom to contract, such a view tends toward political and social permissiveness with respect to the corporation when we focus on those individuals’ right to free speech and association.

The other main tradition in answering the question of the corporation is “concession theory,” or what Orts calls “top-down” explanations. Taking it cue from the historical fact that corporations were originally created by governmental fiat (at first from monarchical prerogative, and then later by legislative bodies), this view sees corporations as gaining their power and privileges from governments, which concedes certain powers to them.19 In contrast to the bottom-up view, then, corporate bodies are not owed deference out of respect to their constituent right-bearing individuals; instead, because the corporation is a legal creature of the state, the state has the purview, and obligation, to order the corporation in keeping with its social and governmental purpose.

The concession theory of the corporation takes its cue from a history going back at least to the trading companies and financial corporations of the early modern period, if not all the way back to the collegium of ancient Rome.20 However, recently this view is generally used as a means of responding to the aggregate theories mentioned above. Against the laissez-faire views of the Chicago School, Joel Bakan has contended that the corporation has been legally constituted as a psychopath, ruthlessly

pursuing its own narrow interests by virtue of its government-mandated orientation. Now, we have already seen how the claim that the corporation is law-bound to pursue profit is an overstatement. Regardless of its content, what is important about this claim is its emphasis on government in the constitution and nature of the corporation. Furthermore, precisely because it is a creature of the government, Bakan contends there is hope yet to salvage our corporate system; we can legally alter its orientation through the state to be a more sociable and socially beneficial institution.\footnote{Joel Bakan. 2005. \textit{The Corporation: The Pathological Pursuit of Profit and Power}. Toronto: Viking. 149, 164.}

We also see the concession theory used to counter the logic of the \textit{Citizens United} decision. In his dissenting opinion, Justice Stevens contends that the founders could never have conceived of corporations as having rights like freedom of speech. This was because corporate charters were granted only by particular legislative acts, and were “conceptualized as quasi-public entities, designed to serve a social function of the state.” \footnote{Justice Stevens. 2010.\textit{Citizens United v. Federal Election Commission Dissent}, 130 S. Ct. 876: 36.} Thus to think of corporations as being merely an outgrowth of individuals is to misunderstand precisely what a corporation is: “While individuals might join together to exercise their speech rights, business corporations, at least, were plainly not seen as facilitating such associational or expressive ends.” \footnote{Ibid, 37.} Because they are created for social purposes, corporations must be regulated and controlled so that they will fulfill their very purpose; as a result a corporation’s participation in social and political affairs—such as donating large sums of money to presidential campaigns—is very much a legitimate concern for legislation. In contrast to “bottom-up” or aggregate approaches to the corporation, accounts in the concession “top-down” tradition present a corporation that not only \textit{can} be shaped and regulated by government, but that \textit{should} be so shaped, so as to ensure it is pursuing its social function.

\footnote{Ibid, 37.}
The relational and norm-governed approach I offer cuts a different path. As I have argued, corporations are certainly created by law and have a social function—namely to increase economic efficiency. There is thus in principle no reason why the rights of the corporation can’t be abridged or regulated. However, to treat it as merely a governmental concession misses a very important voluntary aspect of the modern corporation, namely that though the corporate form itself is established by legislation, any given corporation may be created by individuals without legislative fiat. North, Wallis, and Weingast have referred to this as the creation of an “open-access order,” an opening-up of an institution previously limited to only a select group.24 Given this, concession theory can only account for a part of what the corporation is. The fact that government creates the legal form of the corporation does not mean that individuals will opt to enter into it. Concession theory can not explain why individuals continue to incorporate nor can it specify in any detail how such institutions will be organized.

While governmental in provenance, a corporation is distinct from a state in that its founding is a prosaic and bureaucratic affair and, although a corporation going belly-up, or being bought out and dissolved, is not insignificant, it ultimately happens with a minimum of bloodshed. The creation and dissolution of a corporation is far less dramatic (or traumatic) than the revolutionary or imperial actions needed to bring a state into existence or destroy it. Whereas the state’s power and function comes from its presumed endurance, universality, and compulsory nature, individuals incorporate to pursue their own ends, and corporations can be dissolved when they are no longer useful. To alter the corporate form on behalf of some normative ideal, without considering the voluntary aspect of incorporation, is to miss the fact that people can always choose not to incorporate. We might call this the fallacy of fixity: the unwarranted presumption that a given phenomenon will continue to exist, even if radically

changed. The corporation brooks no such presumption: if changes to the corporate form prove too costly or onerous, people can always choose not to incorporate.

The Corporation as a Re(lation)al Entity

The corporation then has a component that is legal and governmental in nature and a component that is not. However, as we have seen, understanding the non-governmental aspect of the corporation in purely voluntary and contractual terms is reductionist to the point of absurdity. What the relational and norm-governed approach points toward is a revival of a fairly dormant approach to the corporation, generally referred to as the “real entity” theory of the corporation. In this view corporations are neither aggregations of individuals, nor “states writ small,” but social entities that have status in and of themselves. They are not reducible to their individual members and—though gaining particular status and privileges through law—are not purely derivative of government.

This view of the corporation was most famously articulated by Frederic Maitland in his essay “Moral Personality and Legal Personality”:

Group-personality is no purely legal phenomenon. The law-giver may say that it does not exist, where, as a matter of moral sentiment, it does exist. When that happens, he incurs the penalty ordained for those who ignorantly or willfully say the thing that is not. If he wishes to smash a group, let him smash it, send the poleman, raid the rooms, impound the minute-book, fine and imprison; but if he is going to tolerate the group, he must recognize its personality, for otherwise he will be dealing wild blows which may fall on those who stand outside the group as well as those who stand within it….if n men unite themselves in an organized body, jurisprudence, unless it wishes to pulverise the group, must see n+1 persons.  

The group, for Maitland, has personality not by legal decree, but by sociological fact. Furthermore, the corporate entity has legal standing qua corporate entity not as a fictional shorthand that stands in for the

25 As I have already stated in the introduction, it is worth noting that Ciepley claims to argue for a view of the corporation that is “in between public and private,” which sounds like what I am arguing for. However, in the end Ciepley ends up endorsing something more like a concession view, where corporations are “state colonizers of the market.” This is particularly evident in the fact that he does not draw on transaction cost economics. David Ciepley. 2013. “Between Public and Private: Toward a Political Theory of the Corporation.” American Political Science Review 107(1): 139-158.

individual group members, but out of recognition that the corporate group is more than just the sum of its parts. The idea that the corporate body is an “n+1” phenomenon is a helpful way of understanding this. Whereas aggregate theories like Chicago see the corporation only in terms of “n,” and concession theorists emphasize the “+1” of government, the real entity school contends that the corporation is an organization with its own existence, requiring both individuals and the state to manifest fully, but fully reducible to neither.

For the German scholars from whom the real entity theory was derived, accounting for the “+1”-ness of the corporation was not difficult; the corporate’s real existence was a manifestation of its underlying spiritual wholeness. Describing this older view, Gierke put it this way: “since the world is one organism, animated by one spirit, fashioned by one ordinance, the self-same principles that appear in the structure of the world will appear once more in the structure of its every part…. That part itself is a whole with a final cause of its own, itself [appearing] as a self-determining unit.”27 This is not a firm grounding for a modern view of the corporation, since these metaphysical assumptions are far too controversial to underwrite such an important aspect of public policy. The search, therefore, is for an account of the “+1” nature of the corporation that does not rely on such metaphysics.

Recently both Levy and Orts have offered helpful ways of doing this. Levy argues that Maitland’s logic is essentially sound. The +1 nature he locates in six organizational facts of long-term organizations.28 What is first needed, obviously, are (1) individuals who wish to organize. However, once we have organizing individuals we need (2) decision-making procedures: “if we wish our union to outlast the first disagreement…we will need a way to come to decisions, whether a simple majority voice vote or…a self-perpetuating episcopacy with authoritative decision-making power or anything in between.” These decision making procedures are then used to determine the (3) pooled resources that

individuals contribute to enable the group to pursue its ends. As a functional consequence, the group will create (4) offices and roles designed to use the resources in a manner as directed by the decision-making procedures. To effectively delineate between the officer’s and the group’s resources, the resources are to be understood as owned by the corporate body itself. This must be understood both (5) by its members (so that the officer knows there are ramifications for misusing resources not her own), and (6) in law, so that the group can dissuade individuals from misusing group resources and seek remedy in the event they do.

Participant theorists will see all six of these derived from (1) and concession theorists will see them all more or less derived from (6). What Levy helpfully does here is show that 2-5 can remain constant despite associating members changing; once procedures, resources, offices, and ownership are established, the organization has an existence separate from any of its members. Similarly, he shows that 2-5 can exist at a sociological level without legal recognition; it won’t be terribly effective and group endeavour will be frustrated, but it will exist. The corporation in this view is not derived from the state but enabled by it.

In a similar vein, Orts offers his institutional theory in which corporations are “socially established entities that are both authorized and recognized by government and organized and managed by individual participants.” 29 Both distinct from, and compatible with Levy’s account, Orts emphasizes the legal, normative, and pragmatic nature of corporate personality. To use his terms, the law is both “jurisgenetic” and “jurispathic” with respect to the corporation: it enables individuals to establish their own law-like norms within organizations (meaning that there would be a pluralist order of intra-corporate norms, possibly as numerous as there are corporations) but it also suppresses those norms when they conflict with the overarching normative principles of the polity. 30 What the corporation is,
for Orts, is the institutional embodiment of these norms and rules that are both generated by individuals, restricted and enabled by the state, and executed by officers representing neither individuals nor state, but the corporation itself.

The norm-governed account of efficiency that I offered builds on both of these modern accounts of the corporate real entity. We might think of this as the “Re(lation)al Entity” approach. In this view, the corporation is a social phenomenon that exists by virtue of the rules, offices, and resources dedicated to an enterprise distinct from any individual and the kinds of normative scripts (cooperation, legitimacy, authority, teamwork, democracy, etc.) used to induce such dedication. Here, my view of the +1 is an account that sees the procedures and offices that Levy focuses on as requiring the kinds of norms and rules that Orts focuses on, while also giving them effect and actuality. What distinguishes the corporation as an entity in its own right are the kinds of relationships—both normative and procedural—that enable a type of cooperation that would not be possible by individuals acting merely as individuals.

In some ways, this ties in with Blair and Stout’s “Team Production” model of the corporation.\(^{31}\) This view is articulated in opposition to the shareholder primacy doctrine we get from the Chicago School. Instead, in this view, shareholders are seen as but one patron-group among many from which the corporation attempts to secure asset-specific investments. The governing body of the corporation then—the board of directors and the executives—is not seen as representing a particular group’s interests, but the viability and sustainability of the corporate enterprise as a whole. The corporation, then, is a “mediating hierarch,” a system of authority developed to prevent any one group from

withdrawing their investment prematurely, or from taking advantage of another patron groups’ vulnerable situation that results from their making an asset-specific investment.\textsuperscript{32}

This reading is in contrast to Ciepley who reads “team production” as showing the governmental nature of the corporation: “because no natural persons own corporate assets, legitimate control of these assets not only might, but must, derive from something other than ownership of them…. it is government that authorizes management’s rule within the corporation’s jurisdiction.”\textsuperscript{33}

Yet what this view misses, again, is the important individualist element of people wanting their assets to become corporate assets. Despite his claims to be asserting a view of the corporate that is “between public and private,” Ciepley edges on the side of the concession theorists, and therefore cannot tell the full story. Instead, it seems like the Blair and Stout story about the corporation’s “team production” is better understood in terms of the re(lation)al entity status I am arguing for: the asset-specific investments that individuals make for the enterprise are the “+1,” which government enables by recognizing corporate authority over collective assets. What I highlight, based on my account of norm-governed production, is that these asset-specific investments are enabled by cultivating a sense of identity and legitimacy in the form of norms and cooperative scripts. To say that the corporation is a mediating hierarchy that solves the corporate hold-up problem begs the question as to how this hierarchy effectively does so. Some explicit or implicit claim to legitimacy, or some script of cooperation, is necessary.

In this view there is no principled slant one way or another with regards to permissiveness or restrictiveness. The corporation is both constituted by its individual-initiated rules and governed by the state. The significance of its real-entity status is only that it is a thing apart from both. Following Dewey, the question of how the law of corporations ought to be oriented is not a question


\textsuperscript{33} Ciepley, 150.
predetermined by the status of the corporation, but is rather a matter of public policy. As Dewey put it so well, “the entire discussion of personality, whether of single or corporate personality, is needlessly encumbered with a mass of traditional doctrines and remnants of old issues.”34 Because “personhood” denotes a legal agent possessing rights and duties, but does not dictate the content of those rights and duties, insistence on the real status of corporations does not settle questions about what rights or responsibilities corporations have.

At best the normative upshot of the corporation’s real status is to clear the air of the predispositions of both the aggregate and the concession theorists. The questions shift from first principles to thinking in terms of what consequences will likely result from being more permissive here and more restrictive there. To go back to the Citizens United example, we see that understanding the corporation in this way does not predetermine an answer to the question of corporate campaign spending. On the one hand, corporations are not stand-ins for the citizens that create them, and so there can be no simple presumption that corporations must have individuals’ rights. On the other hand such a view does not imply that corporations could not have them either. We might find that seeing corporations as possessing rights is necessary for individuals to exercise their rights: how effective is an individual’s right to free speech or freedom of the press if we don’t see corporate news entities as having those rights, for example? While corporations certainly don’t have an innate right to freedom of religion, it would seem that we want churches (as corporate entities) to possess that right, since precluding that will curtail individuals’ ability to practice their religion freely. Though liberal individual rights here are not dispositive, a concern for those values should lead us to think that sometimes corporate entities ought to be treated as having certain rights; other times—like in the case of campaign finance, I would argue—considering corporations as having such rights would be harmful to such liberal values, and greater regulation is warranted.

34 Dewey, 673.
To answer the question posed by *Citizens United*, then, we need more than just an understanding of the corporation: we need a theory of what we want out of social cooperation, viewing the regulation or freedom of corporations as a means of achieving this end. Similarly, taking the corporation as an entity (or a person) does not do much to help us understand the relational values to be promoted by the corporation. However, its use of normative scripts and relational schemes to achieve efficiency-gains, as well as values inherent to liberal society, does lead us toward certain conclusions about the content and scope of corporate law.

**Fiduciary Duty**

The main thing that the relational perspective of the corporation makes clear is that contemporary understandings of fiduciary duty leave much to be desired. Nedelsky touches on this briefly in her book, arguing that “the obligations of CEOs to maximize profit” tend to negatively affect the manner in which they relate to other patrons both in and out of the corporation, by privileging one set of values for one set of patrons over others.\(^{35}\) In this regard Nedelsky makes the same mistake as Bakan in not noting how the business judgment rule in practice absolves CEOs of any thick obligation to pursue profits for their shareholders. The fiduciary duty of CEOs ends up being much more attenuated, to the point that basically only straightforward self-dealing and personal misappropriation of funds is seen as a violation. Were Nedelsky (and Bakan) correct it would surely be a problem, but at least fiduciary duty would impose some sort of obligation upon executives to act in a particular way for particular people. As it is now fiduciary duty is left mealy-mouthing in the face of judicial deference to business decisions.

Still, this is helpful for us. What we see is that it is not only the specific content of the corporate executive’s fiduciary duty that structures corporate relations; the lack of such content can also structure

\(^{35}\) Nedelsky 236.
relationships insofar as it subjects some corporate patrons to the caprice of others without legal recourse to question their actions. If we want to assure properly structured intra-corporate relations by manipulating the legal scope of managerial action via “fiduciary duty”, we need to state two things. First, we would want to know to whom such an obligation is owed. Does management owe a duty to one particular patron group of the corporation? To numerous or all patron groups of the corporation? Do they have particular duties (attached specifically to their corporate office) that are owed to groups outside their corporation? And so forth. Second, we would want to know what the nature of this duty is. Does management only have a duty to maximize profits (as Friedman contended) or to take the interests of society as a whole into account (as Berle and Means famously concluded)?

As I have suggested, the relational approach leans toward a different answer altogether, namely the team production view advanced by Blair and Stout. This view suggests that the board of directors have duties neither to one particular patron group (like the shareholders) nor to all of them (as the stakeholder theorist contends). Instead, the board and executives owe their obligation to the corporation itself, as a whole: “The interests of the corporation…can be understood as a joint welfare function of all the individuals who make firm-specific investments and agree to participate in the extracontractual, internal mediation process within the firm.”36 Blair and Stout read the deference given to leadership under the business judgment rule as being a necessary part of granting it the kind of independence required to solve collection action problems inherent in team production by making and enforcing decisions that no stakeholder group singly would want. The deference owed to business under the business judgment rule, in this view, is actually a statement about the distinctness of the business enterprise as an entity from its employees, shareholders, investors, customers, or suppliers; the fiduciary duty is therefore owed to this entity, and not to anyone else per se. If we understand the

36 Blair and Stout 288.
corporation in terms of the particular relations it cultivates, this view seems apt: fiduciary duty is owed toward the collective enterprise embodied in these relationships.

But as we have seen, the business judgment rule cuts the other way too; that is, if the manager is viewed as a mediating hierarch that is given the legal room to operate independently of its stakeholders’ particular interests, it is possible to give it so much room that it doesn’t effectively look after the corporation’s interests either. This is a similar problem to what we find in Hobbes’s political philosophy, which Keohane describes as “Hobbes’ Dilemma”: “since people are rational calculators, self-interested, seeking gain and glory, and fearful of another, there is no security in anarchy….but precisely because people are self-interested and power-loving, unlimited power for the ruler implies a predatory oppressive state.”37 In Hobbes’s theory, the basic assumptions about people are what require certain institutional mechanisms (i.e. the creation of the sovereign), but then those same assumptions give cause to doubt how desirable such mechanisms will be. In the team production model the assumption is Williamsonian opportunism, which gives rise to the need for a mediating hierarch; yet, the mediating hierarch, freed from obligation toward any one group, is empowered to use its power in opportunistic ways. Just as the model of the absolute sovereign gives way to despotic practice, the team production model of the corporation gives way to a practice of managerial capture and CEO domination. As one critic from the shareholder-primacy school puts it:

There is no reason to think that independent directors care about stakeholder more than shareholders do. Entrenched managers may pay their employees more, but the excessive compensation of autocratic executives alienates workers. Consider also Enron, WorldCom, and Tyco, whose boards slept while mismanagement destroyed their companies. The losses suffered by shareholders were huge, but few had their lives destroyed as did many employees. These boards were not unusually insensible; they were quite typical. Directors often play along with auctions of financially troubled companies that favor insiders to the detriment of creditors. “Mediating” directors are less watchdogs for stakeholders than lapdogs for management.38

Earlier we saw that the contractarians raised an important point about fiduciary duty to multiple stakeholders, namely that a man obligated to two masters is obligated to none. In the team production model the problem is more acute: a man obligated to no masters is, in fact, obligated to no masters.

Can the team production model be rescued from this criticism? Is there a way to preserve the idea of a corporation as an independent entity to which directorship and management owes a fiduciary duty, without depriving that duty of any real substance? Lee suggests that the fiduciary duty to pursue the best interests of the corporation ought to be interpreted in discursive terms, as a “focal point for the justification of choices by the board and others deliberating about corporate policy.” In this view, the corporation as mediating hierarch does not imply independence from its various patron groups, but rather imposes a kind of social, quasi-public reason onto decision-makers. Like Rawls’s view of public reason and its place in a political conception of justice, directorial and managerial fiduciary duty obliges corporate decision-makers to be able to justify those decisions to their constituents by appealing to a “social reason” that all might find reasonable, without appealing to particular interests or ideological frameworks. Because what is to be done will always be indeterminate and subject to human error and disagreement, fiduciary duty cannot be applied mechanically in the way a straightforward profit-maximization duty might. Instead, actions must be deliberated upon and subjected to scrutiny; to demand that decision-makers perform their job in order to further the best interests of the corporation provides a criterion for deliberation, a basis by which such decisions can be criticized and justified.

---

40 This may seem like a bizarre claim, since Rawls’s philosophy is premised on the distinction between the state as a public institution and other non-public institutions. But in fact Rawls recognizes that non-statutory, “non-public” institution in fact have a certain kind of publicity to their reasoning, which he refers to as “social reason”: “corporate bodies…need a way of reasoning about what is to be done. This way of reasoning is public with respect to their members, but non-public with respect to political society and to citizens generally….These reasons are social, and certainly not private.” (John Rawls. 2005. Political Liberalism: Expanded Edition. Cambridge: Harvard University Press, 220, emphasis added).
41 Though even with profit maximization distinctions between short-term, medium-term, and long-term profit make such a norm subject to dispute.
Giving fiduciary duty a deliberative interpretation does not, in itself, mean much absent an institutionalized forum for deliberation or an empowered audience to whom justifications are owed. Furthermore, conceiving fiduciary duty in this way does not solve the question of what corporate directors and managers ought to be pursuing; it does not and cannot do this. What such an interpretation does do, however, is screen out particular reasons or justifications that might otherwise appear acceptable, and give judges a basis for legal intervention. A decision that can only be justified in terms of shareholder maximization without concern for long-term employees interests, in this view, is likely not a decision in keeping with leadership’s legal obligations. If there exists hard-and-fast evidence or records that suggest such kinds of decisions, these might be grounds for courts to rule that management was in violation of its fiduciary duty to the corporation. This, admittedly, is not likely to happen terribly often; smoking guns are not generally forthcoming, and the multi-interest nature of the corporation can always give a cornered manager an “out” in trying to justify herself. Still, enabling such a method for adjudicating claims is more in line with how corporate leadership ought to be assessed from a legal perspective.

What this view might more concretely suggest are different procedures by which the discharge of fiduciary duty can be assessed. As we have seen, the business judgment rule reads fiduciary duty procedurally: “in gauging whether or not directors have fulfilled their duty of care in a particular transaction, the rule permits judges to consider only the quality of the board’s decision-making procedures.” If we understand the corporation in relational terms, and we understand fiduciary duty as being owed to the corporation so understood, this suggests certain procedural changes. One of the key changes would involve the corporation’s relationship with labor. As I have suggested, the relational approach to the corporation highlights the significance of employees’ personal investments in the

enterprise. Therefore, even if we say that directors owe no particular duty to employees per se, the fact that their obligation is to the corporation imposes upon them certain care toward employees. To this end we might consider asymmetries between the ways the corporation balances the concerns of employees as opposed to shareholders and add a procedural element that might bring more fairness into the consideration.

One such change would be in the corporate duty to disclose information. Currently, the board and management are obligated to be truthful when disclosing information to shareholders and other capital-providers. This is the result of anti-fraud statutes in various jurisdictions but also because not doing so is seen as not following the necessary procedures to discharge one’s fiduciary obligations. As Greenfield notes, however, there is no such obligation for corporations to disclose the truth about the enterprise’s financial solvency or future strategic plans to its employees.43 This is perhaps an even greater problem than shareholder fraud, since shareholders have more access to information about corporations through independent analysts or even just the price of a share,44 while employees generally don’t have access to this kind of information. It seems that making the fiduciary duty of care to include properly informing all parties—not just shareholders—about corporate finances or future directions would be a necessary step if we think of fiduciary duty in discursive terms. After all, how can a particular decision be conceived of as reasonably justified if the decision-maker feels compelled to cover it up? Measures against fraud with respect to employees would instil greater fairness and respect for the various corporate parties generally, and workers particularly, and their contributions to the corporate enterprise.

44 This is actually Easterbrook and Fischel’s argument for why there should not be mandatory disclosure laws. See Easterbrook and Fischel 86-87,282-283.
Limited Liability and Tort

One of the areas of corporate law in which relationally-oriented scholars have directed their criticism is on the question of limited liability. As we saw in the introduction, limited liability is generally understood as crucial for the corporation because it facilitates investment by limiting the risk that would-be shareholders take on. That is, by buying a share in the corporation, a shareholder does not take on an unlimited share of the losses the corporation might take on, either through bankruptcy or tort suits. The most that a shareholder can lose is her initial investment; her personal assets are not on the line to cover greater corporate losses. By allowing investment with liability only for the size of the investment made, limited liability increases efficiencies in three ways: first, it encourages people to invest who might otherwise be too risk-averse to do so; second, it lowers monitoring costs by lowering the urgency for shareholders to engage in such actions; and third, it enables shareholders to minimize risk by diversifying investments (which, without limited liability, would actually increase their risk). Put differently, limited liability enables more risk-averse people to become shareholders by lowering the costs of being passive.

This is, maybe predictably, precisely the consequence to which some relational feminists object: “the concept of limited shareholder liability is anti-feminist from the outset…separation of the investor from the productive use of her assets is but another pernicious form of alienation.”\textsuperscript{45} That is, the separation of ownership and control that limited liability facilitates is generally taken to be a bad thing by these critics precisely because it encourages less activity on the part of corporate investors. By shifting the risk away from those who invest, limited liability deadens the possibility of participatory relationships of investors to their investment, and encourages offloading risk on to third parties. While passivity can’t be outlawed per se, limited liability structures corporate relationships to be passive in

ways that run counter to values of care, compassion, and participation; insofar as relational feminists
want social institutions structured by law in ways that facilitate such values, limited liability is seen as
an impediment to a society in which empowering and liberating relationships flourish. While I am
sympathetic to the spirit of these criticisms, it should be clear that I think such arguments overstate
their case. It is not obvious that getting rid of limited liability would actually promote more
participation or activity so much as it would simply discourage investment.

It is interesting to note here that such relational feminist legal scholars actually overlap to a
degree with some contractarians. Contractarians, again, see corporate law as the default setting for
contracts—they are the contractual terms that corporate parties would reach had they the time and
resources to negotiate effectively and costlessly. As a result, it should be feasible for all terms in
corporate law to be reconstructed privately through contract. Limited liability, in this regard poses a
problem for the contractarian. While it may be possible to imagine investors, managers, and creditors
contracting so as to grant shareholders limited liability with respect corporate debt and bankruptcy, it is
impossible for a contractual arrangement to give shareholders limited liability against tort. Because
torts by definition involve costs imposed on third parties, a contractually-created limited liability of
torts would require the third party to assent to such an arrangement. However, the hypothetical

The contractarian concern with limited liability and tort is characteristically twofold, encompassing claims both on behalf of libertarian rights discourse and economic efficiency. On the one hand, limited liability of tort seems to unfairly shift costs from corporate parties to outside parties and, in the process, limits third party ability to file legitimate tort suits in ways to which they never assented. In terms of efficiency what this means is that corporate shares will be mispriced. Share prices are subsidized by the cost of risk that is externalized onto a third party. From this perspective, if limited liability of debt helps facilitate contracting so that it best approximates the most efficient arrangement, limited liability of tort facilitates contracting too much, leading to over-investment. As a result some have called for the unlimited liability of shareholders for torts.\footnote{My discussion of this is derived largely from Henry Hansmann and Reinier Kraakman. 1991. “Toward Unlimited Shareholder Liability for Corporate Torts.” \textit{Yale Law Journal} 100(7):1879-1934.}

So, some contractarians have called for the elimination of some parts of limited liability, and some relational feminists call for the abolition of limited liability in full. To me, both seem a bit too boldly stated in different ways. As I have noted above, while the relational feminist is surely right that limited liability does not encourage participation and compassion, it is not obvious that eliminating limited liability will do much better to instill these values. That is, getting people to actually \textit{want} to structure their productive and financial relationships according to such values is a deeper puzzle that mere institutional configuration is unlikely to solve. The result is that getting rid of limited liability may not do much in the way of making shareholders more participatory and responsible; it will just discourage investment. In this respect the relationalists have it backwards; their argument is that because shareholders will be concerned with their investment, they will have to become more attentive if they are fully liable. This seems like another fallacy of fixity. Given full liability, people will just be deterred from investment; passivity would not give way to participation, but inactivity proper.
elimination of limited liability seems likely to produce great under-investment without much gain in relational values.

If abolition of limited liability wholesale would be radically inefficient, there seems something mildly unfair about the proposal for unlimited liability in tort. There is so much that already encourages passivity amongst the shareholder (e.g. the liquidity of the stock market, the collective action problem faced by shareholders, etc.). It seems slightly unfair to hold shareholders responsible for actions that they have been institutionally incentivized not to monitor or engage with. Absent other changes that would make, for example, voting and shareholder proposals more meaningful or necessary, to make shareholders liable for corporate tort claims is to hold people responsible for actions over which they had neither knowledge nor say, though from which they may have profited. On the other hand, one might imagine a reason to pierce the corporate veil and assign liability for tort to institutional investors that have the means and incentive to monitor management, and exercise greater influence in managerial decision-making. This would seem to be a caveat to a rule regarding limited liability as opposed to large overhaul to the basic regime; for that reason I bracket this question for now. I consider institutional investors and their specific duties a bit more in the next chapter.

Still, both the relational and contractual perspectives are surely correct that there is something off about the current regime of corporate limited liability. Rhee has captured this as being, fundamentally, a problem about equality. As he puts it, the justification for the current liability regime is something like Kaldor-Hicks efficiency: “the rule’s purpose is not to facilitate liability avoidance, though this is always the effect of shielding shareholders from excess liability….the sole justification for limited liability must be that A should suffer loss of a to make B or society wealthier by b where b > a.” 48 That is, limited liability makes it so that the tort claimant or creditor cannot claim all they otherwise would (or that more risk is imposed on them than they would otherwise take on) in order to

make for a more efficient capital market that benefits the corporate parties to such a degree that they could, in principle, compensate the injured party. Rhee asks if there is a way in which this could be changed so that potential win-loss transfers enabled by the Kaldor-Hicks criterion were converted into win-win Pareto improvements by requiring the winners to actually compensate the losers in the change.

How can this be done without simply getting rid of limited liability, whether for tort or altogether? To illustrate how the relational approach to corporate law might lead to such changes, I here briefly consider one proposal. One way of dealing with these issues is to allow the corporate veil to be pierced and hold corporate decision-makers liable, and require that they hold adequate insurance to cover such potential costs. A tort claim against a corporation that required more than the corporate assets could then extend to executives that are found culpable, with their tortfeasor insurance paying out to the claimant. Another possibility is to require shareholders to post a bond in exchange for their limited liability. These bonds would be small enough so as not to disincentivize investment greatly, but large enough that taken together the bonds would capitalize a large enough pool to cover potential obligations. The latter would likely require less by way of administrative costs, involve no veil-piercing (the corporation is still the one covering obligations, just supplemented by shareholder bond fees), and therefore keep limited liability otherwise intact. The result is a slightly higher cost imposed on shareholders, but otherwise a more or less similarly efficient outcome with a slightly more egalitarian set of relations because of the more stringent Pareto criterion.

The Employment Relationship

The question of to whom legal obligation is owed by management is structured by domains of law other than what is traditionally considered in the rubric of corporate law. There are a number of

---

49 Galbadon, 1449-1454.
50 Rhee 1450-1453.
examples of this—consumer protection laws, environmental law, international law—all of which might merit more discussion in light of the analysis I am offering. Here, I will discuss one aspect of employment law that this relational view of the corporation weighs in on. Employment law is essentially a legal delimiting of the kinds of employment arrangements an employer may impose on a worker. For our purposes what is important is that employment law provides a legal channel through which those who relate to the corporation as employees can challenge decisions—insofar as they relate to employment—in the same way that corporate law does shareholders. Now, in the standard contractual view of the corporation, employment law is viewed as having the same purpose as corporate law—namely to fill in the unstated or unstatable terms of a contract. In the same way that the nexus-of-contracts leads to a laissez-faire view of corporate law, it would seem to imply a Lochner-era style understanding of employment law. While the legal battle over the constitutionality of statutory employment regulation has largely receded into the past, it must be noted that the way in which employees are understood to relate to the firm is in tension, not in harmony, with such regulation.

This is seen most in debates over the “employment-at-will” doctrine. This doctrine, largely confined to the United States, claims that absent some explicit or implicit contractual arrangement to the contrary, employees should be seen as hired “at will”; the result is that they can be fired in a similar manner, without needing to show just cause. That is, unless there is some agreement or statement to the contrary, we should see employee and employer as simply and symmetrically contracting to exchange services for money; if the employer no longer wishes to be in this arrangement then he should be free to end it, just as Alchian and Demsetz’s customer should be free to leave the butcher if he doesn’t like the price of meat. While even in the USA this doctrine has been mitigated by other concerns throughout

the course of the 20th century, it is still an influential consideration in legal judgment, and has been in
the ascendance over the past twenty years.52

The doctrine of employment-at-will has been the subject of criticism from various perspectives.
One influential argument grounds its criticism in the equality of rights-claims; that is, while the
employer and employee both have rights to terminate their cooperative arrangement, in reality the
employer has much more power to do so than the employee because of how both parties are situated
vis-à-vis each other and the atomized labor pool.53 The law of employment should recognize this
asymmetry in adjudicating employment disputes, which should generally involve requiring more than a
view of the employment arrangement as being one entered into “at will.” On the other side, the
response to these criticisms is fairly straightforward: to introduce more substantive concerns into the
law of employment—like requiring notions of “just cause”—will have detrimental effects on the
efficiency of the contracting process and, as a result, will likely result in fewer jobs.54

Viewing the corporation as a relational entity suggests a different interpretation of the problem.
To assert that employees are mere contracting parties with the firm ignores the initial reason why there
are employment contracts at all. That is, the reason why the employer-employee relationship is
characterized by an incomplete contract with some form of authority is precisely because specifying the
terms of the contract is inefficient or impossible ex ante. Employees are often paid not merely for some
specifiable and discrete task, but to provide services that may require firm-specific learning and
investment, or that respond to uncertainty in a flexible manner; they are able to perform certain tasks
better than they would were they coordinated by the price mechanism because they are subject to

52 Katherine V.W. Stone. 2007. “Revisiting the At-Will Employment Doctrine: Imposed Terms, Implied Terms, and the
54 For a discussion of the various debates debates on the at-will doctrine see Alexei Marcoux. 2008. “Business Ethics.”
Accessed July 1, 2015.
cooperative scripts that motivate them intrinsically. Generally, this script is hierarchical in nature, where the employer is followed because she is in a position of authority. This is what the concept of relational contract reviewed above draws our attention to.

Regardless of the specifics, to view the employment contract as a simple contractual exchange is overly reductive and misses precisely the distinction between firm-based cooperation and market-based cooperation. In reducing the employment relationship to one of discrete contracting, the at-will doctrine structures the relationship in a way inimical to basic notions of equality; by abstracting away the actual relations between employer and employee, the at-will doctrine misreads the arrangement as purely voluntary and contractual, and cannot take into account the power that one party wields over the other. In this sense some notion of “just cause” is required by my view of the corporation in light of our valuing of liberal equality and liberty. In this view, the default arrangement is that employees have a legal check against managerial decision-making in the sense that they can challenge unilateral dismissals that are done capriciously or without good reason.

**Conclusion**

This concern for the workers’ vulnerable place in the corporation again raises the question of worker empowerment in corporate governance. Again, it must be restated that efficiency forces us to recognize that not all corporations can be fully governed according to the values we might want, and the partially contractual nature of corporations suggests that we should respect that corporate forms are, in part, contractual and voluntary. Still, the place of the corporation in society, as well as its relational and legal nature, suggests that we should have moral preferences for certain forms of corporate organization, and that these need not be brushed aside out of a misplaced deference to voluntary arrangements. Though democratic control of the corporation by workers or stakeholders through legal fiat may not be permissible within the framework of a liberal democratic market society, unilateral legal intervention is not the only mechanism through which change can be brought about.
As I will argue in the next chapter, there may be ways of bringing the corporation more in line with liberal democratic values without sacrificing efficiency by re-arranging the context in which corporations operate. That is, without assuming away the basic context of a market society, there are institutional and regulatory mechanisms by which more just forms of corporate organizing might be encouraged, even if we don’t make such forms mandatory. In fact, in a liberal democratic society, this might be precisely the nature of a just and efficient regime of corporate governance and corporate law: to create the incentives to organize in a way that cultivates the normative values of society without foisting such values upon corporate patrons by statutory force. This is in keeping with a tension that is at the heart of liberal democracy’s emphasis on value pluralism. Liberal society extolls values like moral equality and liberty, yet it also must tolerate people living according to values that run counter to equality or liberty, for instance according to religious custom that is misogynistic or repressive.55 Yet, as the relational approach shows us, some interpersonal relationships will be more hospitable to liberal values than others will, and so there would be a state interest in delimiting the range of relationships that individuals might enter into. The most consistent, though not romantic or idealistic, way of approaching the corporation then is grant some deference to pluralism (though, as we will see, not to the extent that the Chicago school does), while encouraging individuals to structure corporations in more ethical and politically salutary ways.

Chapter 6. The Architecture of Corporate Governance

In the last chapter I ended by recognizing a tension inherent to liberal pluralism that is necessary to deal with when assessing the corporation from the perspective of liberal democratic norms. What I have called the question of corporate governance—that is, the question of who ought to make decisions within the corporation and on whose behalf—is an area in which this tension is especially acute. As we saw in the chapter, the corporation’s relation to law is both “jurisgenetic” and “jurispathic,” to use Eric Orts’s terms: the state enables the corporation to generate its own internal laws and norms on the one hand, but also restricts those laws and norms in other instances. As a result, I contended that a theory of the corporation must navigate between contractarian and concession theories of the corporation, which see the corporation, respectively, as an extension of the individual and as an extension of the state. On the question of corporate governance, however, I must admit that it is tempting to endorse one more than the other, since it would make answers much easier to articulate. The contractarian would say, simply, that we should let corporations govern themselves, as contract law would allow; the concession theorist would see no issue in intervening and forcing the corporation to be governed according to certain state-endorsed or constitutional values.

What does a relational approach such as the one I am offering suggest? In the first instance, it says that corporate governance ought to be structured so as to cultivate the relationships that best accord with the values we wish to promote. But this just pushes the question back a step, forcing us to ask which values we should be promoting: the values that corporate parties wish to promote (as the contractarian would say) or the ones most in line with liberal values (as the concession theorist would aver)? Because both of these theories are correct in part, neither of them is satisfactory on the whole. As a result, I propose a two-pronged approach to these questions, one which sails closer to the jurisgenetic contractarian account, the other closer to the jurispathic concession theorists.
In the first instance we can look to the values that are explicitly or implicitly endorsed by a particular corporation and then ask if its organization and actions are cultivating relationships consistent with those values. This approach is pluralist in nature, prescribing deference to individual corporations to choose their “normative cores,” the scripts and values that structure how corporations ought to be governed and how managers ought to act.\textsuperscript{56} The normative strictures of law and ethics on this account are simply ways of forcing corporations to be consistent with the values they espouse and have used to attract cooperative patrons to engage in productive relationships.

The second approach starts with the liberal values implicit in society as a whole, and then asks whether corporations are structured and regulated so as to produce relationships consistent with those values. As I have already suggested, I take these values in a liberal democracy to be a sense of moral equality, individual liberty, and a belief in democratic decision-making. This way of addressing the question, then, is more restrictive and less pluralistic, prescribing a certain set of values and norms to which corporations must subscribe.

Taken together these approaches provide a spectrum of prescriptive analyses of the corporation, ranging from more deferential to more regulative, and therefore from more conservative to more aspirational. Their relationship to the efficiency of the corporation is also different: the more conservative and deferential approach will be more likely to be guilty of undercutting what I called in chapter 4 the “minimal horizon” of normative constraint, while the more regulative and aspirational approach will tend toward overstepping the maximal horizon. All of this is to say that I do not present either analysis as complete or perfect; they are more suggestions of the kinds of institutional reforms and regulations such analyses point toward. Despite their different orientations, the prescriptions resulting from both approaches are not mutually exclusive; it is possible to create arrangements such

that corporations are obliged to govern themselves according to the standards they espouse while also encouraging them to affirm particular kinds of relational values.

**The Pluralist, Deferential Approach**

If we begin with the permissive idea that corporations can choose the relational values around which to organize, this would seem not to suggest much in the way of change. As we have seen, the economic account of the corporation does not morally privilege capital-providing shareholders as the ultimate decision-makers or “demos” of the corporation. This principled non-stance on who runs things, however, skews things in favor of shareholder primacy and an empowered executive. The economic explanation for this, following Hansmann,\(^{57}\) is that capital providers tend to be the more efficient patron group to assign decision-making rights to. However, for many enterprises it might be preferable to give them to customers or to workers; if their interests are homogenous enough to reduce decision-making costs, or if their industry is not terribly capital intensive, then it may be more efficient to govern the firm by assigning some group other than shareholders such rights. A key insight from this economic perspective worth highlighting is that diversity and variety of corporate governance structures is not just desirable but important and beneficial for both the individuals involved in a given enterprise as well as the economy at large. Because particular governance structures will tend to be more efficient in different industries and market settings, the economy in which things are most efficiently done will be one that boasts a variety of governance structure. By having a diversity of legal governance structures available for organizing enterprise we enable different comparative advantages to flourish and therefore enable a more efficient economy.

---

The Two-Tiered Cooperative

From this perspective it would seem that the weaker version of the relational approach does not have much to add. If we are assuming that corporations get wide discretion in choosing their “normative cores,” and that having a variety of ownership structures is necessary for efficiency, then at first blush we have nothing to add; the corporations “may select their preferred ‘constituencies,’” and the governing-governed relationship is part and parcel of this choice. However, this shows a slight weakness in Hansmann’s account. For Hansmann the different assignment of governance rights is purely an economic decision. IBM assigns shareholders these rights, Ocean Spray to their fruit growers, and a cooperative apartment complex to its tenants, all for the same reason: because it is most efficient to do so.

This is, frankly, a fairly bizarre way of understanding cooperatives. While cooperative forms of enterprise do entail the allocation of governance rights to non-traditional patron groups like workers, customers, or suppliers, they generally do not do so solely for reasons of efficiency. They do so out of a pronounced affirmation of values like democracy, fairness, and solidarity. While Hansmann is likely correct that the reason the cooperatives succeed in the marketplace is due to the comparative advantages their ownership structures provide, one should not confuse the reasons behind their success with an explanation of their origins. For Hansmann this leads to odd results; for instance, he likens law firms that are owned by (some of) the lawyers who work for the firm to a traditional worker cooperative. Hansmann does this to draw attention to the fact that law firms, like worker cooperatives, find it favorable to assign decision-making rights to employees as opposed to investors;

60 Hansmann, 90-95. His efficiency-centric view also leads him to the claim that women are not offered partnerships on part-time bases in law firms, because to do so would cause inefficiencies in decision-making procedures by introducing heterogeneous interests. Though he does admit that “simple sexism” also plays a role, this is presented as an afterthought (94-95).
we should therefore look for commonalities between law firms and worker cooperatives to explain their assignment of governance rights (for Hansmann, these are common professional experiences and low-capitalization requirements). However, if we take the relational approach to the corporation to at the very least require corporate relationships to match the values that they espouse, than the distinction between a worker cooperative and a law firm that simply has workers as owners becomes very large. Because of its outward commitment to democratic and egalitarian governance, a cooperative must be committed not only to allocating voting rights to members of a particular patron group, but to grant such rights to all members of that patron group.  

This might seem a small point but it is a genuine source of tension for cooperatives. As we saw in chapter 4, worker cooperatives have perverse incentives not to hire when prices are high, since hiring an additional worker means an additional dividend to pay out. One of the ways to get around this is to differentiate newer members from older ones, either through lower pay, or through higher pay but with no entitlement to membership rights. Meade referred to this as the “inegalitarian cooperative,” and made the case that this was a way for cooperatives to work around the inefficiencies Ward forecasted. By differentiating newer members such that they could be reimbursed (both in terms of wages and membership rights), the cooperative could still effectively respond to changes in demand while protecting their cooperative advantages. This “hiring bias” is inegalitarian because the efficiency is achieved “at the expense of a distributional principle which may involve two workers of equal age, sex, ability, skill, etc. working side by side at the same job at the same work-bench, but receiving different shares in the product.” In other words, early members of the cooperative circle their wagons and then

\[\text{61 See the definition and values of cooperatives, adopted from the original Rochdale Principles, given by the International Cooperative Alliance } http://ica.coop/en/whats-co-op/co-operative-identity-values-principles\]


begin hiring other factors of production from members of the same patron group in a non-cooperative manner.

As Co-opLaw (a wiki in the cooperative sector dedicated to addressing legal questions) puts it, “having employees who are not members of the cooperative has positive and negative consequences. Non-member employees can provide an effective temporary workforce for a season business that cannot maintain a year-round level of employment…Members then become like bosses to the non-member employees, creating divisions between workers.”64 This also happens in cooperatives with customers or suppliers as members: cooperative apartment buildings will rent out units to non-members at rents higher than monthly member maintenance fees and supplier cooperatives will buy from non-member farmers, or issue preferred stock to newer members that come with derivative rights but no voting rights,65 which the Saskatchewan Wheat Board did in 1996.66 The question then arises: does the cooperative owe a duty to treat potential incoming members in the same way that it does established members? Is a two-tiered cooperative membership justified?

As I have suggested, the relational view would submit that it is not. The affirmed values that are used to structure and secure cooperation in these firms are explicitly built upon a commitment to equality, solidarity, and democracy. This commitment does not end with the originally incorporating cohort of workers (or suppliers, or customers, etc.), but is characteristic of the enterprise in itself. As a result the governance structure must reflect these values by granting voting rights to all eligible members of that patron group. To do otherwise would be to run contrary to values that are being used to secure cooperation from others; that is, it is not only unfair to workers who do not get a share in decision-making processes, but also to customers and suppliers who do business with the cooperative

64 http://www.co-oplaw.org/topics-2/governance/#Members_vs_Nonmembers_WithinWorker_Coopera
tives July 3, 2015
because they understand it to be run according to certain values. As a result, we have reason to believe that such arrangements are unfair.

Now this would not mean that in a worker cooperative, a worker must—as a matter of right—get a share in governance in her first day on the job. A process by which non-members become members, or a stipulated period of time after which workers would have the ability to gain membership rights, does not violate the idea given above. This is in fact generally how cooperatives work. However such policies should not be implemented selectively to discourage conversion to membership, or make it more difficult, as a means of securing the benefits of greater cooperation without extending the franchise of the cooperative. This last point is worth stressing: such a constraint will almost certainly impose efficiency-losses on the cooperative. There will generally be gains to be made by hiring non-member workers (or renting to non-member tenants, buying from non-member suppliers, etc.) and so forcing cooperatives to take their cooperative values seriously is not an insignificant demand.

This appears to be what the values avowed demand of the cooperative firm. Still, it might be asked whether this is not overly-demanding, having the effect of discouraging a form of governance that already has difficulty flourishing. This is likely true, but this presupposes the question we will address shortly: should values of corporate governance be constrained or oriented in certain ways by appealing to values endorsed by society? If so, then we will need to consider whether such a demand on cooperatives may serve to damage the types of relationships we want to cultivate more generally in society. For now we are taking the more conservative assumption that corporations can choose their values, and therefore are merely required to structure their internal relationships in light of those normative cores. In this view, the cooperative, by virtue of being a cooperative, ought to follow through on its cooperative values, even in the face of inefficiency. This does not mean that trade-offs are

---

67 Lund, 11-12.
impossible; the democratic involvement of members may be attenuated by manager prerogative or by time considerations. Still, it seems that such a basic aspect of a cooperative’s identity is its assignment of governance rights to its members; as a result, to creates tiers of membership in a cooperative is to violate fundamentally a core facet of cooperative organization, and therefore ought to be disallowed or, at least, discouraged.

*Shareholder Democracy vs. Shareholder Influence*

Does the analogy hold for standard business corporations where shareholders have voting and derivative rights? That is: are corporations with dual-class, differentiated voting stock structures violating their implicit relational norms in the same way that two-tiered cooperatives do? The answer, it seems, is no. Unless the corporation has committed itself to being a shareholder democracy of some kind, there is no reason to think that a standard business corporation has endorsed anything like equality in voting the way a cooperative has. Instead, because the presumed or asserted value of the corporation is maximizing efficiency, there is no reason to think that shareholder voting implies a concern for shareholder democracy. Differentiated stock in a business corporation is not an affront to its basic identity and values; after all, the degree of influence that shareholders have is already so differentiated based on the number of shares a shareholder owns, or the amount of investment a shareholder makes. However, one might suspect that some mechanisms that are used in standard business corporations do still imply a certain amount of shareholder input. That is, by issuing voting rights to (particular) shareholders, it would seem that the corporation is obliged to take their voice at least somewhat seriously. Even if there is no implication of shareholder *democracy*, per se, we might think there ought to be some semblance of true shareholder *control*.

This helps us understand why there are regulations in most jurisdictions requiring, say, annual shareholder meetings and director elections. The fact that the corporation allocates voting rights to shareholders does seem to imply certain relations between it and them; as such, we require corporations to have annual occasions for shareholders to exercise those votes. For Easterbrook and Fischel, this is tantamount to foisting democratic values upon a group of people that did not necessarily want it. As I have suggested, though, this is misleading in two ways. First, democratic values require more than just voting: at the very least we would expect that voting to be equal across those with the franchise and as inclusive as possible. Foisting democracy onto shareholder elections would require both the elimination of differentiated voting shares and, possibly, an institution of one-shareholder one-vote. Second, the corporation is more than just a group of individuals and therefore may involve constraints or procedures that any of those individuals taken singly would not ask for. That voting rights are given to shareholders implies that there exists a general body “vested with the authority to decide or act on behalf of the corporation as a whole.”

This, not democracy, is the basis for mandatory shareholder voting. As such, we can accept all of the economic logic behind why shareholders are generally extremely passive and still insist that such procedures are required.

This has at least one prescriptive implication. The first would be that when shareholders are not passive they are to be taken seriously, especially in the instances of their most foundational mechanism for influence, director elections. While it is rare for shareholders to be interested or mobilized enough to challenge either management’s nominations for directors or the re-election of an incumbent director, it does sometimes happen. In the face of public controversy, a “no” campaign can be mobilized (usually by an institutional investor wielding enough votes and resources to do so), which is generally the most active use of the shareholder franchise. However, even in the event that such a

---

campaign is successful, it does not guarantee that the directors will be removed. Because the shareholder vote usually just means that directors must tender their resignation, the board can always refuse to accept their resignation, leaving the voted-down director still installed in his position, what is known as a “zombie director.” A study by the Committee on Capital Markets Regulation found that in the US, this phenomenon was fairly common, and that there was little evidence that actual director resignation was influenced by shareholder votes.

This set of institutional practices appears to violate the norm inherent in issuing voting rights, that is to say, that such voting rights matter and ought to be taken seriously. While some advocates have contended that the best way to challenge this is by replacing plurality voting with majority voting, it seems the practice that makes most sense is to disallow boards from rejecting a director’s resignation or, at least, to make it extremely difficult for a board to act in this way. This would appear to give shareholder voting the type of status implied by the corporate allocation of decision-making rights. One could envision similar types of reform that would make shareholder voting more meaningful in “say-on-pay” provisions, and instances of bylaw required votes (say on buyouts or major directional changes in the corporation) and thus further strengthen the implicit idea of shareholder decision-making.

---


Institutional Investors

Taking seriously the idea of shareholder control, if not shareholder democracy, raises an interesting question regarding the duties of institutional investors. According to SEC Commissioner Luis Aguilar, in 2010 institutional investors owned roughly 67% of all publicly traded equities, and 73% of securities in the 1,000 largest corporations.\textsuperscript{75} Given institutional investors’ share of shares, they are in a unique position to wield influence on corporate governance issues through threat of exit, voting in corporate elections, or sponsoring shareholder proposals.\textsuperscript{76} This is why whenever voting campaigns are launched among shareholders to oust a particular slate of directors, there is generally an institutional investor behind it; they are in the unique place to solve the collective action problem at the heart of shareholder voting. Insofar as shareholder control should be made more viable by virtue of shareholders being granted control rights, this would imply a special place and duty for institutional investors. The problem, however, is that institutional investors hold massive portfolios that are diversified with holdings in myriad corporations. The cost involved in exercising the proxy votes for all of those shares is prohibitive. The question, then, is how to encourage, incentivize or force institutional investors to do so.

One proposal has been that institutional investors ought to be required to hire outside directors to monitor board and governance issues of the corporations in which they have interests.\textsuperscript{77} This would further enable institutional investors to maximize the value of their portfolios while also keeping them apprised of potential issues in which they ought to exercise their influence. The other proposal, which has been implemented to some extent, is to treat the corporate governance powers of institutional investors

\textsuperscript{76} Though shareholder proposals tend to be weak mechanisms for corporate influence, they are significantly more likely to be effective if sponsored by institutional investors. See Stuart L Gillan and Laura T. Starks. 2000. “Corporate Governance Proposals and Shareholder Activism: the Role of Institutional Investors.” \textit{Journal of Financial Economics} 57: 275-305.
investors’ holdings as part of their fiduciary duty to the plan-holders; institutional investors like pension funds would not only have to act responsibly with regards to their portfolios’ financial earnings, but also with regards to potential opportunities to intervene. In the USA such movements are underway: proxies have been interpreted as plan assets, which brings their exercise within the purview of investors’ duties.\(^78\) Part of the requirement has been that institutional investors devise policies governing when they exercise proxy votes and what procedures they use to do so. Proponents of “pension fund democracy” take this view further, contending that institutional investors must consult their planholders on these policies as part of their fiduciary obligations, and that such policies must have a “comply-or-explain” component in order to provide accountability to their beneficiaries.\(^79\)

**Corporate Governance II: The Societal, Regulative Approach**

I now turn to the second, more restrictive approach to the questions of corporate governance. If we radically depart from the status quo and presuppose that the relationships internal to the corporation ought to be thought of in terms of the more general values of liberal democracy, what are the result? If we ask that corporate structures cultivate relationships conducive to individual liberty, moral equality, and democratic decision-making, then how does this affect our view on corporate governance?

As we saw in chapter 4, some political theorists have argued that this should require corporations to be run as worker democracies or stakeholder democracies. These arguments were motivated by one of three different assumptions: (1) “the parallel case” -- the idea that corporations are analogous to states and therefore demand the kinds of rights and procedures we demand of states; (2) “democratic holism” -- the idea that democracy in a holistic sense requires a certain set of extra-


statutory institutional arrangements, one among them being democracy in the firm; or (3) “workplace eudemonism” -- the idea that participatory democracy in the workplace is necessary for cultivating the kind of citizens necessary for a free and democratic society. As I suggested, the last of these is most convincing, since the parallel case strategy and democratic holism assume too much congruence between corporation and state and/or underestimate how far-reaching liberal neutrality extends in the case of intra-corporate organization.\(^{80}\) The workplace eudemonist approach of Mill and Pateman, on the other hand, notes how structures outside of governmental relations affect individuals and the way they approach political life. As a result, those structures that tend to cultivate individuals conducive to democratic society should be encouraged, and those that do not ought to be discouraged.

This has the advantage of also fitting very well within a relational framework—in fact, it essentially is a relational approach though Pateman and Mill did not use such language. That is, the workplace eudemonist argues for the extension of democratic norms into the workplace based on the idea that such norms will cultivate relationships that will be good for democratic society. Thus, though workplace eudemonism does not follow from a relational approach (since, as we saw in the last section, such an approach does not fix the kinds of values that ought to be encouraged), it fits well within it. As a result we can tie the democratic theory of thinkers like Mill and Pateman to both the idea of the corporation as a relational entity and the idea of the relational contract. The corporation is given real presence because of the relational investments it enables and cultivates with certain patron groups playing a much larger part in generating such relationships. This means that within the corporation, certain members will have made investments that are less fungible than others, making their relationship more precarious than others. Furthermore, these same people are more likely to be beneficiaries were they given democratic control; contrary to shareholders who generally relate to the

corporation in instrumental and transactional ways, workers have a more pronounced relational stake. This means that if we see cultivating citizenship as a value inherent to liberal democracy, and we see increasing participatory relationships as a strategy of achieving such a value, we ought to encourage worker participation in the corporation. This helps us understand both how certain patron groups differ with respect to their contributions to the corporation (contra Alchian and Demsetz) and how such arrangements affect our endorsement of wider values in liberal democracy.

However, as I also argued earlier, partisans of workplace democracy can underestimate the kinds of efficiency constraints under which corporations necessarily operate. This is why Mill did not argue that the corporation ought to be transformed into a worker democracy by legal fiat. Instead he thought that worker cooperatives would prove themselves the most efficient form of corporate governance, a process that would then have salutary consequences for citizenship and, by extension, liberty and democracy. Because the corporation is always in need of being more efficient than pure market transactions, prescribing a particular form of governance can be in danger of making the corporation less able to fulfill its inherent efficiency mandate. Thus, even if we see the relationships that constitute the corporation as implying a preference for worker democracy, it does not appear to be sound public policy to legally force corporations to fit the mold. Instead, it seems like we can say this to the question of corporate governance: we ought generally to prefer structures of corporations where those who most closely relate to the corporation have a say in its governance. This will often be workers, but might also include customers or raw goods suppliers. Indeed, for many industries we might judge that those who most closely relate to the corporation are in fact the capital providers. In this view a more just regime of corporate governance may not be one that mandates worker control, but is perhaps one that encourages cooperative organization where decision-rights are allocated to those whose investments give them a relational stake in the corporation.

In admitting the above I run very close to Dow’s view that workers ought to have an “alienable right” to a democratic firm, though such a right can be attenuated by larger social concerns, including
those of efficiency. That is to say, despite the concession to pluralism that we cannot force cooperatives onto people, I want to argue that cooperatives ought to be encouraged, and that there may exist regulatory and institutional impediments that make the cooperative undervalued or under-chosen as a corporate governance structure. The fact that not all corporations can be turned into worker cooperatives by legal mandate does not mean that the number of cooperatives that our current system produces is just or right. It may be that many non-standard governance structures, like cooperatives, would be preferred and implemented, were it not for the manner in which current structures discourage their organization. In this view the best alterations for a more just corporate governance structure have less to do with mandating particular arrangements, and more to do with setting up the game such that particular arrangements friendly to values of liberal democracy might best flourish. This approach has the benefit of alleviating some concerns for efficiency: since corporations can always revert back to their standard shareholder-oriented form, or never make the jump to cooperatives in the first place, we need not worry that encouraging cooperatives will cause large economic inefficiencies. We can think of two broad strategies for doing this: 1) arranging the choice architecture such that it favors cooperative governance and 2) arranging the institutional environment such that it reduces impediments to cooperative survival. These strategies are complementary and may be helpfully thought of in terms of providing incorporating members with the will to choose cooperative governance and the means to sustain it.

---


82 This is only true if cooperatives are not encouraged by great amount of subsidization. Subsidies have the effect of softening an enterprise’s budget constraint, causing inefficiencies for the economy on the whole (Janos Kornai. 1980. “‘Hard’ and ‘Soft’ Budget Constraint.” Acta Oeconomica 25(3/4) 231-245.). A key assumption of this analysis is that although we want to encourage cooperatives to exist, we must also allow underperforming cooperatives to fail and to be kicked out of the market, out of a concern for social efficiency.

Choosing Cooperatives

One way to think about encouraging cooperative corporations is to think about the various stages at which such a governance structure can be chosen. For ease of exposition, I use the example of a worker cooperative though, as mentioned above, the particular relations of a corporation might demand governance rights be on offer to other patron groups. An enterprise might start out as an informal worker-run business or a more formal partnership amongst workers and then naturally transition into a worker cooperative. Alternatively, an enterprise might begin as a partnership between a worker-owner and a capital-providing “silent partner” until it incorporates, at which point the worker-partner (and several other employees) buy out the silent partner’s share and a worker cooperative is formed. Finally, a corporation may start out as a standard capital-owned, shareholder-driven company and then be bought out or converted into a worker-owned enterprise. Put differently, we can think of worker-ownership as being effected pre-incorporation, at incorporation, or post-incorporation.

Again, it is worth noting that nothing legally prevents worker enfranchisement at any of these stages under existing arrangements. The question posed to partisans of worker-ownership, then, is why such models are not chosen. One possible reason is that the way corporate law is set up, people are nudged to choose capital-centric models of incorporation. That is, both because the default setting for corporate law is shareholder-orientation, and because most people are relatively uninformed about worker cooperatives, enterprise-formation tends to follow the path that is both presented and known. One very simple way of countering this would simply be to make the cooperative option known and presented. This is hard to do at the pre-incorporation stage, but one could imagine that when a firm incorporates there would be an additional step beyond assigning ownership shares that involved deciding whether such shares would be limited to certain patron groups in the future. That is, incorporators must actively decide not to incorporate as a cooperative.

A more significant and ambitious approach would be to legally entrench the option of transitioning to a worker cooperative after incorporation. Malleson has suggested that such a legal
mechanism is a necessary prerequisite for true expansion of cooperatives, capturing it as part of the conditional licence for corporations to hire labor: “employers should be allowed to rent people’s labor in the manner that one rents machines, if those same hired workers are legally entitled to turn around and buy in to become co-owners.”

There are various ways of imagining such a procedure. Dow, for instance, has offered one particularly ambitious proposal. In this model, workers at a corporation would have the right to hold a referendum on whether or not to convert to worker ownership. The steps to this process would go something like this:

1. Workers first must meet a particular threshold (10-20% perhaps) of signatures on a petition in order to trigger a referendum.
2. The referendum begins, with workers and management weighing in.
3. If the referendum passes (the threshold could be a simple 50%+1 or some supermajoritarian percentage) then the process for transitioning to worker ownership begins. This process includes:
   a. Founding a labor trust. The trust is a distinct legal entity, funded by percentages of worker’s wages, which has the role of phasing out private shareholders in the stock market. With the oversight of a financial manager, the labor trust uses the money garnered from wages to slowly buy shares on the open market. The labor trust holds these shares, using dividends to further fund its purchasing. The labor trust is run by a board of directors elected by all workers (including management), with each employee getting only one vote regardless of wage contribution to the trust.
   b. As the labor trust buys shares in the corporation, the trust’s board places directors on the corporation’s board of directors in proportion to the number of shares the trust holds. Eventually

---

84 Malleson, 88.
85 Dow, 263-268.
this will be 100% and the board of director’s will be made up entirely of directors elected by the labor trust.

c. The corporation has then been transformed into a cooperative. The dividends that trust-held shares are earning are then distributed to employees in addition to wages, calculated in the same manner that wages are calculated.

I summarize this proposal in detail for two reason. First, it is a fairly comprehensive and serious proposal for how to facilitate a greater transition to worker cooperatives. It also proposes to do so without causing win-lose transformations; because the labor trust buys the shares, and does not expropriate them, shareholders are compensated, and the outcome can be considered win-win. It is worth noting, to this end, that what Dow essentially does is take an Employee Stock Ownership Plan (ESOP), adds the legal mechanism of the worker referendum, finances employee ownership through wages (slowly) and not through debt (quickly), and insists that the distribution of power amongst workers be equal, not proportionate to wage contribution. This would appear to bolster its practical viability, considering that ESOPs are now a known-quantity, at least in the American economy, and appear to have done fairly well; although ESOPs and worker cooperatives are distinct, the mechanism for conversion has a track record that warrants trying it for the purpose of stimulating cooperative growth.

The second reason I have gone over this proposal in detail is to make a different point: to note how quickly the discussion shifts from how we can enable workers to choose cooperatives, to what the necessary means are for such a choice to occur. Nudging incorporators toward forming cooperatives or


legally enabling workers to take over a corporation will not be very effective if the means aren’t there to acquire the enterprise or keep it capitalized. This is why Dow’s referendum can only work with a much more complicated scheme about how the transition will actually happen, involving setting up separate entities, garnering money from wages, and so forth. It is also not fool-proof. ESOPs are generally instigated by the employer and do not necessarily involve 100% of shares being purchased by labor. This difference is important for Dow’s proposal: if the existing board or management is unhappy with the results of the referendum, they may attempt to quickly liquidate assets and distribute the proceeds to shareholders—especially if they themselves own shares. Dow's plan, then, actually has the characteristics of a hostile takeover by labor that is given force by law, but which happens slower than most hostile takeovers, allowing for management to respond in kind. It's also not clear what happens if management holds a substantial portion of shares. This points us beyond the problems of whether or not workers will choose cooperatives, toward the problem of whether cooperatives are a viable form of governance structure, given various barriers in their way.

All of this is to say that altering the law to enable the choosing of cooperative governance structures can only play a limited role. The bigger need, it seems, is to alter the background institutions in such a way as to even out the advantages that standard business corporations enjoy. There are a number of advantages and disadvantages that firms have that are likely intrinsic to the way they are structured: for instance, worker cooperatives will intrinsically be better able to monitor worker performance and facilitate communication of worker preferences, while they will have a harder time making efficient and timely decisions; investor-oriented corporations, on the other hand, have the advantage of a simpler and straightforward maximand, and an easier time obtaining collateral for loans, while having to outlay more costs for monitoring. However, in addition to these inherent attributes, are advantages and disadvantages that have more to do with the institutional setting in which such
corporations operate. The thought is that a more just regime of corporate governance will have to involve some altering of these settings to make the playing field more level.

Environmental Changes

We can discern three main challenges that cooperatives face that seem to have more to do with the background structures of the economy than they do with their viability or desirability (or at the very least, which are exacerbated by background structures). The first is a lack of institutions through which worker cooperatives can get capital. That is, because the valuable asset that workers have—their labor—is inalienable, lending institutions will be less likely to offer them credit since workers lack collateral; they can always walk away from the debt and leave the creditor without much in the way of recourse. The result is that it will be more difficult and expensive for cooperatives to get necessary capital.

The second challenge is an inability for workers to diversify their investments. This is in many ways an extension of the first challenge. Because capital is alienable and human labor is not, workers are unable to sell or buy shares in labor; workers, unlike capitalists, are forced to invest in the same place they work since their work is what gives them shares in the firm (and because they will be unable to finance the enterprise through debt because of the first challenge). The result is that worker-cooperatives impose more risk on their members than standard business corporations do; in the latter

the worker has only her job to lose and the investors can diversify their investment, but in the former the investor and the worker are the same so she has both her job and her share in the firm. 91

The third disadvantage for the worker cooperative is that it does not have a market for control. This is, again, an extension of the inalienable nature of labor and the capital difficulties this raises; the point here, however, is that the stock market gives the standard business corporation a comparatively efficient mechanism for managing corporate governance. As we saw in chapter 3, corporate governance in standard business corporations is generally understood in terms of “exit” as opposed to “voice.” Disapproving shareholders can exercise the “Wall Street Rule” 92 and sell their shares, depressing the price of stock and signaling poor managerial performance, and opening up the disciplining threat of takeover. In this sense, the absence of such a mechanism deprives worker cooperatives of three advantages that other corporations have: the ability for members to exit, the threat of takeover, and the informational shorthand that stock prices provide to assess managerial performance. While the first of these seems more intrinsic to the cooperative form (that is, democratic participation is valued precisely as an antidote to exit), the last two may be able to be tweaked by background institutional innovation.

These then are three disadvantages that worker cooperatives have by virtue of the institutional environment in which they operate: expensive/inaccessible capital; inability to diversify/mitigate risk; and no market for control that can be used as a tool for disciplining managers. Various commenters emphasize different things: Bowles and Gintis stress the significance of access to capital 93; Elster

91 John Boatright. 2005. “Employee Governance and the Ownership of the Firm.” *Business Ethics Quarterly* 14(1): 7. See also Jon Elster. 1989. “From Here to There or If Cooperative Ownership Is So Desirable, Why are There So Few Cooperatives?” *Social Philosophy and Policy* 6(2): 94. Both make the point not only that people are generally risk averse and therefore will want to diversify investment, but that workers (who will generally have little in the way of disposable income) will generally be more risk-averse and therefore more in need of diversification.


emphasizes risk-aversion, and Hansmann emphasizes the inefficiencies of governance in the absence of an “exit” option. For my purposes I don’t think we need to determine a single cause; it is enough to note that the existence of all three doesn’t do much in the way of helping cooperatives start and thrive, and that mitigating them couldn’t hurt. I go over some plausible policy paths that might address these issues.

1. Addressing Capital Constraints

We might conceive of the lack of access to or expense of capital for cooperatives as a straightforward market failure: there are enterprises that would be profitable and efficient had they received the capital, but the debt market will systematically provide capital at too low of a quantity or too high a price.\textsuperscript{94} Seeing as market failure provides a prima facie case for government intervention it seems a plausible enough idea that there ought to be governmental subsidies for cooperatives as a way to counterbalance their difficulty in attaining credit, perhaps in loan guarantees. Miller makes a similar argument, though with a different emphasis; focusing on liberal neutrality and the bias of the market to under-invest in particular goods (like cooperative enterprise), Miller contends that a liberal society must be committed to subsidizing things like worker cooperatives out of respect for institutional neutrality to conceptions of the good.\textsuperscript{95} These are compatible arguments. Because demand for a particular type of legal and productive form of organization will be underfunded in the economy, government intervention serves as a way both to increase efficiency (by redressing a market failure) and to increase fairness in society (by combatting an institutional bias, in the service of liberal neutrality).

There are a number of potential ways in which government can intervene to correct market failure in a credit market.\textsuperscript{96} One possibility is a “directed credit” program, normally used in developing countries,\textsuperscript{97} in which banks and lending institutions are given subsidies in return for lending to particular types of projects at favorable rates. The idea here is that laws and regulatory institutions could be created that direct a certain number or percentage of loans to eligible cooperative businesses in return for government subsidies. One could also imagine a special insurance program developed that protects lenders against cooperative default. This type of approach has the advantage of tapping into the institutional competencies of people who are already in the business of lending money, and just directing them toward cooperative enterprises. The disadvantage, of course, is that not all lending institutions have experience or knowledge of cooperative businesses, and the particular types of issues they might run into.

A better alternative is to establish a cooperative bank that is dedicated solely to financing cooperatives. Many advocates suggest something like this, modelling this idea on the successful Mondragon Cooperative in Spain. One of the features of Mondragon’s complex organizational structure is the Caja Laboral Popular credit union, the central financing institution for all of the collective’s sub-enterprises. All of the cooperative firms do their business with the bank, and the bank, in turn, gives loans and oversight to the cooperative sector. Now, generally one ought to be hesitant in using Mondragon as a guide for the potential success of cooperatives elsewhere. Mondragon’s very particular history and context—that is, being born in a post-war Basque region of Spain with a strong sense of national identity, solidarity, and a common minority language—is very likely a part of its impressive record and success. Shared identity and purpose often helps solve all sorts of coordination


problems and collective action problems through shared sets of background norms, appeal to common senses of authority, and recourse to informal social sanctions. As one cooperative leader put it:

“Mondragon should be seen as an aspiration, not a blueprint.”

All of this is to say that what might come easier in the Basque region of Spain, Israel, or Quebec might be harder to get off the ground elsewhere. It will likely be imperfect and take much more on the part of the state and other formal institutions to enforce what might otherwise happen through more cooperative scripts indigenous to a particular culture (in a particular place at a particular time).

Still, there are aspects of Caja Laboral Popular that provide some insight into workable policy elsewhere. In particular, one of the ways the credit union works is that all of the Mondragon cooperatives are required to use it and only it for savings and loans. Its lending services to the sector, then, ought to be understood as the service it renders in exchange for having a regional monopoly on commercial and personal banking. Furthermore, this common institution ends up serving a crucial networking function, where businesses and individuals are put into contact with one another to share knowledge and assist each other.

This seems a more viable and sustainable policy option than directed lending (though I concede that it is politically unlikely in the current environment). Instead of directing banks to lend to cooperatives, the government establishes a cooperative bank with the condition that all cooperatives must do their business with that bank and its subsidiaries (or its sister institutions in other jurisdictions,

---

100 For a study that takes the contextual element seriously and then explores potential macro-political effects of context see Edward Greenberg. 1983. “Context and Cooperation: Systematic Variation in the Political Effects of Workplace Democracy.” Economic and Industrial Democracy 4: 206. Though the Marxist approach Greenberg uses is not totally compatible with the study here, his point overriding point that the context of the cooperative movement will determine its overall effects or success is accurate.
in the case of a federalist country). The bank could be structured as a credit union, such that all cooperatives would be able to elect and nominate members to its board of directors, and have a say in governance. In return for this granted monopoly, this bank could similarly only do its business with cooperatives and, in return for subsidies, provide affordable rates that would be less than what other financial institutions might be willing to provide. Not only would this get capital to cooperative businesses, it would also establish a sector with statutorily mandated institutions. One could imagine this bank gaining the type of authority and importance that it would be able to create policies specific to the cooperative sector, creating a body of regulation tailored to the needs and problems of cooperatives. Such a unified sector would be able to create an “ecosystem” particular to cooperatives, with intermediary shelter organizations, and complementary institutions developing in the process.

2. Addressing Risk Problems

With the development of an institutionalized cooperative sector, one could imagine ways of mitigating the risk that cooperative workers take on. Primarily, insurance products and risk-sharing schemes amongst cooperatives could be used as a way to protect members against the failures of their cooperatives. Cooperatives might be mandated (either by the cooperative bank or by statute) to portion off a percentage of their profits and contribute them to a cooperative insurance company that pays out to its members in the event that their place of work folds. \(^{102}\) By making it mandatory this would go some way to overcoming the potential adverse selection problem where only riskier businesses bought into the risk-pooling arrangement. Moral hazard might be mitigated by making use of intermediary institutions that provide oversight, guidance, and reporting on cooperative businesses.

\(^{102}\) See Ellerman, 103-105.
Another possible—though perhaps loftier—option for mitigating risk would be to establish a labor-sharing program. Under this arrangement, cooperatives would hire, or at least give preferential consideration to, workers left unemployed by contracting or bankrupt firms. In a way this is a kind of insurance that obtains in large business corporations where a worker gets protected from fluctuations in demand for her labor by the largeness of the enterprise and its ability to retool or redeploy her for some other purpose. In a cooperative labor-sharing program, this type of meso-level substitute for unemployment insurance would essentially be mimicked by an entire sector instead of an individual corporation. The cooperative sector would therefore pool risk not through pooling money, but by pooling labor; a member of a worker cooperative who has both her job and her investment tied into the same enterprise could therefore have some guarantee of employment security, if not job security.

3. The Market for Cooperative Membership

As we have seen, one of the large comparative advantages that standard business corporations have over cooperatives is the stock market, which provides a market for ownership, generating price information and discipline that makes corporate governance more efficient. This might at first seem to be something intrinsically available to the standard business corporation and unavailable to the cooperative. This, however, would ignore the various historical and institutional factors that have contributed to the creation and sustenance of stock market liquidity, which enables the existence of a market for ownership. In reality there is nothing intrinsic to the corporation that implies the existence of a market for ownership; it is the result of politics and policy. Insofar as such policies give the shareholder-centric corporation a comparative advantage over other forms of corporate governance, we might consider undoing some of these features in order to facilitate the relative viability of the

---

103 Doucouliagos, 1109-1110.
cooperative. This, however, would appear to be a form of levelling down to achieve equality; providing cooperatives a “leg up” shouldn’t come by disadvantaging other corporate forms.

This said, there does seem to be some intrinsic factors about worker ownership that makes it distinct from the shareholder model: the alienability, mobility, and fungibility of finance capital in contrast to the inalienability and relative immobility of labor. As Dow puts it: “if control rights are tied to labor supply, it is impossible to transfer [control of a labor managed firm] from person A to person B without replacing A’s labor services with B’s labor services. On the other hand, in a corporation the voting rights attached to a share of common stock can be passed from A to B without altering the physical assets owned by shareholders as a group.”104 Even if we accept that there is nothing organic about securities markets it does seem that there is something inherent to worker control that would make it difficult to construct an analogous market. Of course, in some ways the whole point of cooperative ownership is that governance of the enterprise happen as a result of collective autonomy and democratic decision-making, not through the impersonal market mechanism. The relative inexchangability of labor is precisely why cooperative advocates argue for the importance of worker “voice” over shareholder “exit."

Is there a way to keep this characteristic aspect of worker cooperatives while trying to capture some of the gains in efficiency that come from the market for ownership? To this end, I propose the creation of a mock stock market in the cooperative sector, wherein each member of a cooperative is given an equal amount of mock money to buy mock shares of all coops listed. These members could then trade these shares based on their assessments of other cooperatives. This would generate a set of prices for cooperative shares. Such prices could help members assess the performance of their cooperative and its management relative to others. As a result one could generate something like the

information created by stock markets without sacrificing the core tenets of cooperative democracy or ignoring the inalienability of labor.

Now, there is good reason to suspect that a mock market like this won’t be great at its job. As anyone who has tried to play poker “for fun” and not with money knows, simulating gambling without the personal stakes involved does not work very well. In order to generate accurate price information in a market it would seem that the risk of loss and the possibility of gain are necessary to motivate people to take their transactions seriously. Just as strategies will be distorted when people play poker without money—for instance, excessive bluffing or long-odd bet-calling—we can expect people to take excessive “risk” in a mock stock market, since there is no real punishment if one bets wrong. Alternatively, absent any possibility of gain in the event that one is right, one might expect the market to be fairly inert, with people dropping out and not really paying attention or using their mock money. To really make the mock cooperative market effective, one must introduce stakes of some kind.

It is worth noting here that this proposal with actual stakes involved would essentially look like a sectoral application of Roemer’s famous model of “coupon socialism.” Roemer was attempting to conceive of a socialist economy—that is, an institutional arrangement through which “aggregate profits are distributed more or less equally in the population”\(^\text{105}\)—that would not have to sacrifice efficiency in the manner that state socialism did. Of course, such a project goes back well into the early 20\(^{th}\) century with economists like Lerner and Lange. Roemer’s contribution was to recognize that most prior models of market socialism had not taken seriously the significance of the principal-agent problem in corporate governance, and that socialist institutions could learn much from capitalism’s mitigation of these problems.\(^\text{106}\) What was needed was a socialist economy that did not rely on public control of enterprise

---


or, more specifically, an economy that produced egalitarian results while also imposing hard budget
constraints on productive enterprises.

Roemer therefore proposed that a socialist economy could have two sets of currency: money, which could be used to purchase commodities, and stock coupons, which could only be used to buy ownership shares in corporations. The latter could not be bought, sold, or gifted; one receives coupons upon achieving the age of majority (and then, in an amount equal to others) and upon death relinquishes them to the central bank (which then redistributes them equally throughout the population). Upon purchasing shares in a corporation, citizens gain both the right to dividends (in money) and the right to vote for the company’s board of directors. If a company is earning high profits then people will be willing to pay more coupons to earn the rights to receive dividends; if a company shows promise, then people will be willing to sell their shares of high-yielding company’s in order to buy more shares of the younger company in the hopes of being able to receive more in dividends down the road. As such, stock prices will fluctuate based on the relative supply and demand for shares, itself a result of the adult population’s assessment of the corporation’s performance (or, what is more likely, a mutual fund’s assessment, which adult citizens invest their coupons in). Roemer further argues that mutual funds should be created so that workers could reap the benefits without great monitoring or discovery costs. Firms then take the coupons that they receive in exchange for selling shares, and trade them to the central bank for money at an exchange rate determined by the bank (the only legal circumstance under which coupons can be exchanged for money), which the firms can then use to purchase capital goods.

This coupon stock market thus produces three important results. First, it produces price information that can be used to assess the performance of a company. Second, it allows citizens to diversify their risk by buying shares in companies (or buying shares in mutual funds, that buy shares in

---

107 Roemer. 1994. 75.
companies). Third, it helps allocate capital effectively, since firms still get their capital from a central bank, they get it in proportion to the market value of their shares. Furthermore, because shares cannot be exchanged for money by common citizens, citizens cannot liquidate their shares and spend them; as a result, there is a form of egalitarian paternalism built into the model, making it difficult for folks to cash out long-term assets in favor of short-term consumption goods. Though these shares can produce profits for their holders, the shares themselves cannot be sold for profit, which helps prevent inequality that might originate from citizens’ poor consumption habits.

Now, this foray into Roemer’s plan is helpful for our purposes although our aims are not exactly the same. Roemer was more concerned with an egalitarian distribution of profits than he was about achieving labor ownership of enterprise (about which he was ambivalent if not suspicious). Here, I am concerned with institutional arrangements that might be favorable to worker ownership, while bracketing the question of macro-level economic equality. Still, the goals of Roemer’s economy are similar to what worker cooperatives lack and need: adequate price information, diversification, and capital allocation. Significantly what Roemer does is attach real rights to these stocks; his coupon economy is not a mock market or a cashless poker game in the way I described above. Citizens (or mutual funds) have potential gains to be made by taking the market seriously, since they might be able to earn money in the form of dividends or earn more coupons through the heightened value of such a share.

So to respond to our concerns above about the efficacy of the mock market, we might attach similar rights to them as Roemer does. Cooperative members can use their cooperative share coupons (CSCs) to purchase the rights to profits and voting rights in other cooperatives. One stipulation would be that irrespective of the number of held shares individuals could never get more than one vote, as is in keeping with the democratic principles of cooperative governance. Still, the dividend-earning rights of these shares would induce seriousness on the part of traders, and their liquidity would enable their generating of informative prices to be used by the bank for potential loans and by members to assess
strategic direction or the performance of management. It would also enable cooperative members to earn money through investments in other cooperatives, thereby allowing them to diversify and mitigate some risk.

The difficulty with this for our purposes is that giving the stock real rights has the consequence of watering down the collective autonomy of the worker cooperative. That is, if people from outside the cooperative can purchase shares and then be able to vote, have we undermined the animating idea of the worker cooperative in order to make it more efficient? To address this a number of stipulations would need to be introduced. First, similar to Roemer’s citizens, individuals would get CSCs only upon becoming members of a cooperative, and would have to give them up upon retirement (or upon leaving the sector for a stipulated period of time). Therefore, though control would be extended beyond the membership of the direct cooperative, it would be only extended to other cooperative members; it would be a stock market only accessible to worker-owners of cooperatives. In becoming a member in a cooperative you not only receive membership rights in your firm, but the possibility of membership in a certain percentage of other cooperatives as well. Thus though altering the nature of cooperative governance, the general philosophical commitment to self-determination is maintained though understood at the level of the sector as opposed to the firm. Indeed, we might see this as an institutional interpretation of the famous Rochdale principle that cooperatives must cooperate with other cooperatives. The cooperative stock market would institutionalize a way for all cooperative members to cooperate in a dispersed but effective manner.

However, one can imagine various ways of protecting the autonomy of members within their firm if this type of change is not welcome. Let us imagine that in our cooperative stock market, cooperative businesses have an incentive to sell shares and get CSCs because they can use them to buy capitalized goods (that is, though they cannot be converted into cash for the purpose of individual consumption, they can be converted for the purposes of purchasing capital goods). There might therefore be an incentive for a cooperative to sell shares to the point that its own members are in the
minority. As a response it could be stipulated that only less than 50% of shares could be sold to non-members. To do this one could assign more voting shares to members than to non-members. So, for a 100 member cooperative that wanted to sell 499 shares to the cooperative sector, they could assign 5 voting rights per member (so that among members they all have equal votes) and one voting share per non-member shareholder; the result would 999 votes in total with more than 50% held by cooperative members.

Another possibility would be simply to stipulate that voting rights only accrued to actual members in the cooperative; other non-members who buy shares with CSCs would only get rights to residuals—the equivalent of non-voting stock. This would still incentivize real engagement in the CSC economy while protecting democratic autonomy within the enterprise. Furthermore, if one is concerned not only with equal voting but equal earnings, there could be a stipulation that members cannot use their CSCs to purchase extra shares in their own cooperative; members would therefore each earn the residuals gained from the same amount of shares, and would also be forced to diversify by investing their CSCs in other enterprises.

Conclusion

Here I have argued that a relational approach to corporate governance in which we try to make internal corporate relationships cohere with overarching social values suggests a system of corporate governance where those who relationally invest in the corporation ought to have governance rights. This, however, cannot be thrust upon the corporation by mandate: people may prefer not to take on governance of the corporation, and such governance may prove to be inefficient. I have argued, therefore that a just form of corporate governance requires measures that would make the legal and institutional environment for worker ownership more viable. Another way of putting this is that if the space for norms within the corporation is constrained by efficiency concerns, then we ought to manipulate the market environment to widen that space. I have suggested some broad and inexact
means of doing this: legal means to facilitate the choice of the cooperative form and background institutions that could make the cooperative more efficient. In the latter I suggested three main institutions: a cooperative bank given a monopoly over cooperative financial services and charged with the responsibility of providing loans to cooperatives, in order to counter capital constraints; sectorally-mandated insurance and risk-pooling arrangements among cooperatives to overcome the lack of risk diversification; and, most ambitiously, a cooperative stock market that would facilitate a market for membership, and provide price information to discipline cooperative governance.

Now these proposals are all fairly inchoate and exploratory. It is possible to imagine various iterations and variations of all these schemes, which would address the many technical difficulties any one policy would likely face. For my purposes, such a scheme need not be completely specified or detailed. The point is to offer ideas about the kinds of institutions necessary to make the choice of corporate governance structure a real choice, and to enable the kinds of relationships implied by liberal democratic norms. If it seems I have gone off-topic, from discussing principles of corporate governance to a more in-the-weeds discussion about institutional mechanisms, then this is a consequence of the exercise itself. That is, one cannot really discuss a principle of corporate governance without considering background institutional environments. We can only speak abstractly of a principle of shareholder primacy or the “Wall Street Rule” because things like the NYSE, the SEC, and Delaware law courts already exist and provide the requisite background for such “principles” to have effect. As Roemer put it, “the market…does not perform its good deed unaided; it is supported by a myriad cast of institutional characters that have evolved painstakingly over time.” ¹⁰⁸ The market for corporate control that underlies modern corporate governance has had a long institutional and regulatory evolution. To give effect to a principle of corporate governance that tries to cultivate relationships and values consonant with our commitment to democracy, what is needed are not pronouncements and

¹⁰⁸ Roemer. 1994. 4
manifestos on the importance of workplace democracy. We need to develop the requisite background institutions that can enable such an economy to function well and to flourish.
In one of the most famous scenes in Antigone, King Creon bemoans the effect that money has on human morality. Incensed by the fact that Polynices’s corpse has been buried, despite his orders to the contrary, Creon immediately suspects that someone was bribed to disobey his command. Unable to account for why his authority has been disrespected, Creon falls back on the precept that the possibility of individual gain always tempts people away from fulfilling their social obligations. To understand deviant behavior, Creon’s moral psychology asserts, we look for the extrinsic monetary incentives that draw individuals toward that behavior. And yet, as we know, Antigone’s actions are not motivated by money; she is drawn to her action by a sense of obligation and duty, though an obligation and duty that she takes to supersede Creon’s rule. Creon is unable to conceive of this—that the perpetrator might be acting according to a sense of obligation, just not on obligation toward him—because of his assumptions about the strength and primacy of private gain in determining individual behavior. Of course, such an assumption is not consistent with the fact of his rule; that Creon issues edicts, and that he expects people to follow his edicts by virtue of the fact that he issued them, implies that people are motivated by all sorts of things, monetary gain being only one among them. Creon’s misunderstanding of what leads people to engage in particular actions is what produces his inability to consider the possibility that he was disobeyed because of different normative commitments, namely Antigone’s fidelity to her family and to the Gods.

---

In many ways the Chicago school has a similar problem with understanding the behavior of actors within the corporation. Faced with claims that the corporation ought to be organized along particular normative principles (or in recognition of particular social phenomena like power), Chicago denies the very possibility, by asserting the primacy of extrinsic incentives and material motivations. Finding before them a particular set of social facts, Chicago’s only explanation is that they have been effected by material incentives, that they are the result of people bargaining for the best deal they can get. Yet, like Creon, such an explanation makes the basic institution harder to explain. If Creon is right that what motivates people is individual greed, then what can account for his position as king? Similarly, if particular arrangements are the result of market bargaining, then what can account for there being a firm in the first place? Like Creon, then, Chicago must reckon with the fact that normative values necessarily underlie such a social institution; this invites the possibility that the particular normative values and set of obligations that currently structure the social institution might be challenged by virtue of their being problematic or inconsistent. The legal and normative deference to extant corporate structures, by virtue of the assumption that they represent the end-result of free and efficient bargaining, does not effectively disarm the claims that such structures are in tension with the fundamental norms of liberal democracy.

On the other hand, as we have seen, the liberal democratic norms we assert are not trumps that can be simply played against any offending institution. That we assert the necessity of constitutional democracy with regards to the state does not in itself constitute a reason to think that corporations should be organized in a similar manner. To keep up with the literary analogy, we might point out that the fact that Antigone feels obligations to a different set of commitments does not in itself justify her disobedience to Creon. Creon’s point, to give it a slightly anachronistic and modern gloss, is that political authority exists for particular political ends, which are distinct from other normative
commitments, though no less important.² In disobeying him, Antigone has undermined his authority and, in turn, undercut his ability to uphold political order. In burying her brother, Antigone has committed a category error to Creon; she does not recognize that when we are acting in the political realm as citizens or governors, our moral commitments change, requiring that we forego our personal understandings of what is good in favor of what is necessary for political society to continue. Advocates of democratizing the corporation by legal fiat are like Antigone in the sense that they both confidently assert normative values without considering that what they are asserting might be in tension with the overarching mandate of the particular institutional context in which they find themselves. Whereas for Creon’s kingship this is the imperative of social order, for the firm this is efficiency and productivity. Normative claims that are made without concern for efficiency miss something crucial about the institution they are addressing.

To use a hackneyed analogy, one can note the socializing and educational role that organized sports like football perform in society, especially among young students. Yet if one were to suggest a course of action for a league, team, coach, or player that only emphasized this socializing function, while underplaying the competitive or physical nature of the sport one would obviously be in error. Let’s say, one suggests that one ought to avoid any action that might cause pain to someone else, because it is clearly a form of anti-social behavior; therefore we will not tackle our opponents. Not only would such a suggestion have absolutely no chance of being pursued or taken seriously, it misses the most important and interesting fact about something like football: its social and socializing aspects are not in contradiction with its competitiveness or physicality, but are consequences of it.

Antigone, of course, ends tragically, in large part because of both characters’ inability to recognize the claims and motivations of the other’s actions. I want to say something similar about approaches to the corporation. Perspectives that don’t recognize the normative nature of the internal life of the corporation can vindicate illiberal and dominating institutions that undermine our most

² Ibid, 752-761.
foundational political values. On the other hand, perspectives that don’t recognize the significance of efficiency to the very existence of the corporation are both impractical and self-undermining, since they don’t acknowledge the very nature of the institution they are trying to reform. The corporation therefore requires a normative account that is distinct from, but also attentive to, an account of its economic efficiency; the normative principles that underlie corporate organization must grow from a recognition of both the potential harm of corporate power and the social function of the corporation.

In the last two chapters I have tried to show how such an approach implies changes in our current regime of corporate governance and corporate law. In terms of corporate law this requires us to recognize that the corporation is both an institution that submits participants to relational forms of power and authority that can be morally problematic and an organization that has a contractual basis. As a result the corporation is neither an outgrowth of individual initiative nor a concession of state authority but, I argued, an entity that has a real existence and therefore requires a legal status that recognizes both why it exists and how it can lead to social outcomes that are an affront to liberal and democratic values. In terms of corporate governance I contended that all of this should lead us to have a preference for corporations that assign governance rights to those parties that have a greater relational stake in the enterprise; however, the efficiency-constraints imposed on the corporation preclude us from foisting such governance structures on all corporations. Instead, I contended that we ought to arrange the background institutional environment in such a way as to reduce the costs of choosing such governance structures and maintaining them.

Even given such legal and governance structures, however, there still remains a wide range of options in terms of how those empowered within the corporation ought to act. Furthermore, because of the corporation’s efficiency constraints, there will still exist many standard shareholder-oriented corporations, even in the more just regime of corporate law and corporate governance that I am arguing for. What is needed, then, is an account of a “role ethics” for corporate managers and executives that articulates what goals and aims they ought to be striving for, and what strategies they ought to be
avoiding, in light of the latitude and discretion their roles provide for them. This is the domain of business ethics, which is the subject of this chapter. Here I attempt to articulate a theory of business ethics that fits with how I have approached corporate governance and corporate law: one that takes account of economic efficiency, but in a way distinct from the Chicago school’s hardheaded and hardhearted profit-maximization ethic; and takes account of values other than efficiency in a more realistic and viable manner than stakeholder ethics. I take as my starting point the “market failures approach” (MFA) to business ethics, a view that takes efficiency to be the primary moral principle for business, but shows how such a view leads to different ethical conclusions than profit-maximization does. As I will explain in more detail below, the MFA holds that businesses have an ethical duty not to exploit market failures. This view, I contend, is strong because it provides a robust account of business’s ethical duties within the framework of contemporary economic theory; business ethics is neither portrayed as a wet blanket draped over the C-suite nor a self-serving rationalization of business’s self-interested activities. Instead, business ethics is shown to fit within a larger scheme of social cooperation, taking seriously business’s place within that scheme.

I then argue that taken seriously, such a view requires us to take more than efficiency into account. In the MFA, efficiency is posited as the guiding ethical concern because of the institutional context in which business operates; however, if we take the institutional matrix in which business operates in its entirety, we must conclude that efficiency cannot be the only value implied by such institutions. For the same reasons that businesses must avoid profiting from market failure, I contend that they must also avoid profiting off of what I call “justice failure”—the failure of the state to establish the background conditions of liberal equality and fairness. As a result, we find that the liberal democratic values asserted in chapter 2, which I have shown in the last two chapters must affect corporate law and corporate governance, must also affect the content of business ethics and the manner in which managers make decisions. Unlike the last two sections however, which deal with the more technical aspects of prescriptive reform, this chapter articulates a far more aspirational account of what business ethics ought
to look like. This is appropriate because business ethics has a large component that is extralegal, asking that actors go above and beyond their legal or contractual duties. Although the content of business ethics ought to be informed by institutional context and be realistic in that sense, it also makes sense to offer a theory of business ethics that can account for what managers should strive for, given the most optimal possible conditions and personal dispositions. Throughout the chapter I try to signal where I am being more aspirational and where I am being more practically minded.

**Economic Theory: Key Concepts**

I begin by offering some background for three economic concepts that are necessary for understanding the MFA, but have not been discussed explicitly thus far: 1) market failure; 2) the first and second fundamental theorems of welfare economics; and 3) the theory of second best. I have alluded to or discussed these concepts throughout the dissertation. However, because the MFA was articulated with these concepts in mind, I make reference to them many times throughout the chapter, both in my exposition and my critique of the MFA. It is therefore worth taking some time to lay them out for the sake of exactness.

**Market Failure**

The modern formulation of “market failure” was given by A.C. Pigou. As we have seen, for Pigou the beauty of the market was that it could provide incentives for individuals that match the social effect of their action.³ Often this idea gets captured in the idea that markets help to internalize externalities. Pigou argued that there are two ways of understanding the way in which an action does or does not contribute to the national dividend: in terms of its marginal social net product and its marginal private net product. The former is the increase in goods and services brought about by the marginal

increase of resources that accrues to no person or group in particular; the latter is the increase in such goods brought about by a marginal increase in resources that accrues specifically to the actor who brought about the marginal increase in resources. When the market is working well, these are equal: an individual gets the same return or pays the same cost as the increase or decrease in the total stock of goods she has caused. When this happens, individual actors get rewarded personally in proportion to their contribution to social welfare; the individual pursuit of self-interest leads people to unintentionally contribute to the collective good.

However, Pigou dedicates the second part of his *Economics of Welfare* to detailing and analyzing the various reasons why markets may not equalize these two things. These all suggest good reasons for public policy to interfere with the market: “when there is a divergence between these two sorts of marginal net products, self-interest will not, therefore, tend to make the national dividend a maximum; and, consequently, certain specific acts of interference with normal economic processes may be expected, not to diminish, but to increase the dividend.”

What we call “market failure” is the result when these two concepts do not align: when the marginal social net product exceeds the private (in which case there is underproduction of goods and frustration of demand) or when the marginal private net products exceeds the social (in which case there is overproduction). The former is often defined in terms of underinvestment or missing markets, when there is no incentive for an actor to do something that would add to the social store of goods (public goods like roads and other infrastructure are a classic examples). The latter is defined in terms of negative externalities: people are incentivized to engage in actions that are to the detriment of society (pollution is the most intuitive example). These asymmetries create what Pigou referred to as a “prima facie” case for government intervention: subsidies to

---

4 Ibid, II.IX.1
incentivize actions whose absence would otherwise cause a relative decrease in private net product, and
issue taxes on actions whose prevalence would otherwise cause a relative decrease in social net product.\textsuperscript{5}

The result is that the “prima facie” case for governmental intervention is made on precisely the same grounds that the case for the market is made—efficiency. Taxing the production of goods that create externalities, and publicly subsidizing or producing goods that will be under-produced by markets is justified in terms of the efficiency such actions produce, precisely the grounds upon which the market’s “internalizing externalities” function is defended. Referring to something as a market failure is to claim that a particular market is offering incentives, distributing information, producing monopolies, or creating some other circumstance that tends to undermine (or fail to maximize) market efficiency.

\textit{The First and Second Fundamental Theorems of Welfare Economics}

That is not to say that state action is restricted solely to efficiency-promotion. Later in the 20\textsuperscript{th} century welfare economics developed to show how a commitment to efficiency need not, in principle, mean an apathetic stance toward equality. The first and second fundamental theorems of welfare economics showed, respectively, that given certain assumptions, 1) a competitive pricing system of private exchange would produce optimal allocation across all markets—or, more technically put, that all markets would reach an optimal equilibrium\textsuperscript{6}, and 2) given that there are multiple possible optimal allocations, “every Pareto-optimal allocation of resources is an equilibrium for a perfectly competitive economy, provided a redistribution of initial endowments and property rights is permitted”\textsuperscript{7}—any particularly preferred optimal allocation could be reached by altering initial endowments through tax-and-transfer schemes.

\textsuperscript{5} Ibid, II.XX.4
What all of this means is that a market can produce an efficient allocation of goods based on some particular initial holdings of market participants, and that an efficient and egalitarian outcome can be produced in the market by altering the initial holdings of market participants through governmental taxation. Combined with the account given above regarding the governmental role in efficiency promotion, the theory of welfare economics should be seen as promoting three main roles of the government: 1) subsidizing the production of goods that would otherwise be under-produced, 2) taxing the production of goods that would otherwise be overproduced (controlling externalities), and 3) achieving the egalitarian values of a society by engaging in tax-and-transfer programs.

The Theory of Second Best

It is widely recognized that while markets in theory can produce all the great results we might expect from them, in practice the assumptions upon which those theoretical models rely are absent. The general response to this by defenders of free markets is to say that while the assumptions do not hold, what we ought to do is approximate those assumptions to the greatest degree possible. What the theory of second best shows, however, is that “if one of the Paretian optimum conditions cannot be fulfilled a second best optimum situation is achieved only by departing from all other optimum conditions”\(^8\); in the real world where competitive markets don’t produce all of the hypothetical efficiency results because the assumed conditions don’t exist then the next best alternative will not be to most closely approximate the rest of those conditions. If the completely open and competitive market of economics textbooks does not exist, there is no reason to think that the best bet is to make the market as competitive as possible; we might find that we can best achieve efficiency by introducing other institutions that address the absence of those assumptions. Despite being armed with elegant models of market efficiency, the correct

response to market failure is not an unabashed “more markets,” but an experimental approach to addressing these practical problems with a variety of institutional solutions: sometimes markets, sometimes legally constituted firms, sometimes strong market regulation and taxation, sometimes the straightforward public production of goods.

The Market Failures Approach

The MFA attempts to articulate both the ethical underpinnings of markets and business, and a prescriptive program for business practice and economic institutions, with a firm grounding in the economic theories reviewed above. It is helpful, to get a grip on the type of project the MFA is, to start with the problems it was designed to address and resolve. One problem is a tendency for critics of markets and capitalism to make two interrelated conflations: to conflate profit-maximization with self-interest on the one hand, and the manager with the firm on the other. Because profit is generally something advantageous to a person, it is easy to think that “profit-maximization” as a corporate mandate is simply the extension of a principle of self-interest. As a result, it is easy to assume that business ethics must fundamentally be about identifying when profit-maximization must be constrained by ethical concerns; ethical theories that advocate profit-maximization, it is thought, must be inherently oxymoronic or disingenuous on their face, providing ethical justification for what is simply self-interest. The problem with this view, which Heath pointed out⁹, is that it conflates the firm’s profits with the manager’s self-interest. While there may be alignment between the manager’s self-interest and the successful performance of the firm, corporate scandals like those of Enron and the recent banking crises are only the most obvious and famous examples of how the pursuit of individual interest on the part of

---

managers is often in tension with the profitability of business. Thus the conflation of self-interest with profit maximization reflects confusion about the very subject of business ethics.

The second, related, problem that the MFA seeks to resolve is the confusion about why or how the idea of profit could actually be considered an ethical aim in the first place. With most ethical standards, it is more or less obvious why the ethical aim is a good one because it accords with our basic understandings of morality on a larger scale. It is good that people save lives, therefore it is intuitive why doctors ought to be ethically obligated to their patients’ good health. Pursuing profit seems different because it is not obvious how such action connects to the good of anybody other than the profit-maximizer. From this perspective, even if we don’t commit ourselves to the stronger idea that profit is somehow wrong, it is not obvious how the pursuit of profit could be considered an ethical endeavour as opposed to, at best, an ethically neutral one.

The MFA seeks to address both of these confusions by putting business ethics in its rightful context: “one cannot do business ethics without some appreciation of what justifies the system of private enterprise….we need to understand why corporations should be entitled to pursue profits, in order to understand the responsibilities of managers.”\(^{10}\) Just as one cannot understand why it isn’t a moral problem for a football player to tackle an opposing player on the field (a behavior we generally frown upon in polite society) without understanding the game of football, to understand the role of profit and self-interest in business ethics one must first understand the business manager’s context, the imperatives this context creates, and whether or not there is a justification for this environment. For business ethics this involves understanding three institutional features of the economy—the imperatives of the market, the nature of the business firm, and how such things are structured by the legal and political system—and leads to four distinct prescriptive orientations for business decision-makers: (1) a competitive orientation in the market, (2) a cooperative orientation within the firm, (3) a willingness to respect legal

\(^{10}\) Ibid, 73.
constraint, and (4) a willingness to restrain oneself from unethical behavior. One might think that (4) in the above list is the main domain of business ethics. What the MFA does, however, is insist that (4) is not exhaustive of business ethics, and that in fact the content of ethical restraint can only be made sense of once the first three classes of behavior are understood.

*Competitiveness in the Market*

What reasons might be given to endorse a competitive market? For reasons similar to those I gave in the first chapter, the MFA follows the tradition of welfare economics and favors the Paretian argument over a Lockean one or, if one prefers, the welfarist argument over the libertarian one. The Paretian defense of the market endorsed by the MFA is based on the efficiency-maximizing nature of competition. When suppliers compete with suppliers to sell to purchasers, and when purchasers compete with other purchasers to buy from suppliers, the prices at which goods trade will adjust to reflect relative supply and demand for a particular good. As prices change to reflect supply and demand, the result is a more efficient allocation of resources: “society has succeeded in minimizing the overall amount of waste in the economy....fewer resources will have been spent producing goods that no one wants, at the expense of goods that people do want.” Thus the market is preferred because of its welfare-enhancing properties. By discouraging waste and encouraging the direction of resources toward their most preferred social use, the efficiencies brought about by the market allow for a greater degree of preference-satisfaction and human welfare.

Profit maximization, it follows, is to be seen as an advantageous thing because it encourages earnest competition amongst suppliers and consumers. In order to maximize profits, firms will attempt to

---

11 I use the term advisedly. The libertarian argument is Lockean only in the sense that Nozick and others have interpreted Locke as a libertarian. There is much in Locke’s *Second Treatise* that does not square with a libertarian approach to political economy.

12 Ibid, 75.
sell to the highest number of potential customers by lowering their prices, or by altering the goods they are producing to better meet demand. Either way, the result is a more efficient use of resources, and the customers are the beneficiaries. Much like Adam Smith’s invisible hand, the price mechanism allows the drive to maximize profit to contribute to the public good. Profit-maximizing firms are therefore a good thing given the context of a competitive market economy, not because they are extensions of natural liberty, but because they encourage the use of scarce resources toward the benefit of the most people:

Thus, if we ask what the obligations of managers are, the answer can be provided quite directly. The function of the market economy is to produce the most efficient use of our productive resources possible. This can be done, roughly speaking, by achieving the price level at which all markets clear. The role of the firm in that economy is to compete with other suppliers and purchasers for profits in order to drive prices to that level. Thus managers are obliged to do what is necessary in order for the firm to maximize profits in this way. Profits show that the balance of “needs satisfied” to “resources consumed” is positive, while losses show that the resources would have been put to better use elsewhere.13

This is a straightforward application of the first fundamental theorem of welfare economics reviewed above. Competitive markets help lead to efficient allocations of goods; therefore, market competition, in and of itself, is not a bad thing and can lead to positive results for all, given the correct institutional context. This last qualification, however, is what leads to the other three features of the MFA since, as we know, perfectly functioning markets are creatures of abstract theories and not the real world. Thus, while the first fundamental theorem helps us understand why competition and profit-maximization might be ethical activities, it cannot be the whole of ethical activity; different normative orientation and activities are required because of second best considerations.

Cooperativeness in the Firm

The market in which firms operate is one of competition that encourages markets to clear, and thus generally serves to promote welfare. But as we know, markets may fail to optimally promote welfare in a number of ways. As we have seen in the discussion of Coase and Williamson in chapter 2,

13 Ibid, 77.
the dominance of firms and corporations in market economies (relative to private contractors) suggests just how many transactions require the supersession of the price mechanism in favor of more overt forms of cooperation. As we have seen, the MFA suggests that in order to understand the nature of business ethics we must first understand the normative basis of the market and its emphasis on profit. Similarly, we must also understand what the role of the firm is in the market, and how actors within the corporation fit within this story.

What the MFA takes from the Coase-Williamson story about transaction costs is that firms, and their managers, contribute to efficiency in two related but distinct ways. In the first instance they do so indirectly, by participating in organizations whose profit-maximizing behavior in a competitive market contributes to the efficient use of resources. However, the second way they contribute to efficiency is by complementing the market with cooperatively organized production, superseding the price mechanism within the corporation, in order to create goods that would not otherwise be created, or not be created as efficiently. Again, both of these functions are justified normatively by a Paretian emphasis on welfare and are therefore not in tension with each other ethically, despite the distinct institutional mechanisms used to achieve this goal.

The obligations of the manager coming from the welfarist stance of the MFA are therefore Janus-faced in orientation: generally cooperative within his firm, and generally competitive with other firms. From this we can see that the profit-maximization ethic, whether or not it is substantial enough on its own, cannot simply be understood as a dressed up defense of naked self-interest. The norm of profit-maximization requires the manager of a firm to act on behalf of the firm (or the shareholders of the firm), and therefore to act in good faith and a trustworthy fashion, while not requiring the same trustworthiness or good faith when interacting with her competitors. To see that this has teeth, one need only consider that such a norm condemns many of the various corporate scandals that have occurred over the past twenty years.
Legal Constraint

Without going too far afield, the MFA roughly follows the Pigouvian story about market failure and therefore approaches the law (in the economy) in similarly welfarist terms: the law is ideally used to constrain transactions in ways that protect society against market failure. However, whereas corporations address market failure by supplementing the market with cooperative mechanisms to create greater sets of goods, the law tries to address market failure by constraining market transactions that have a deleterious effect on social welfare. That is, because markets are imperfect, and things like externalities, missing insurance markets, and incomplete information exist, there are actually a whole range of activities that firms can engage in that would not contribute to efficiency. The law, in this view, is meant to step in to prevent or, at the least, de-incentivize, these behaviors. This comes from following the logic of the theory of second best; absent the achievement of all necessary optimizing conditions that would allow for markets to work their Paretian magic, there is good reason to think that we must deviate from other optimizing conditions as well. In this instance this calls for using non-market institutions to achieve second best levels of efficiency:

The basic rules for marketplace competition laid down by the state—including the system of property rights—are designed to limit these possibilities, in order to bring real-world competition closer to the ideal (or to bring outcomes closer to those that would be achieved under the ideal, in cases where a functional competition cannot be organized). This is the motivation that underlies not only direct state provision of public goods, such as roads, but also state regulation of negative externalities, such as pollution.14

Although a firm can maximize profits by selling faulty merchandise or by dumping its costs onto society as a whole (as is the case in pollution), doing so does not augment social welfare in a Paretian sense. Firms, therefore, contribute to efficiency not only by competing in a market for consumers or by producing goods and services that competitive markets would under-produce; they also, in following the law, refrain from doing those things that would create profit to the detriment of consumers or third parties. In this sense, it is not just market competition, but market competition in which the ground-rules

are followed, that generates the ethical results that business managers ought to pursue. As a consequence, following the law is not merely a civic duty for business leaders but an ethical one as well.\textsuperscript{15} This has all been under the assumption, of course, that the laws are as they ought to be; that is, that law is created in order to create the more efficient ground rules and is not subject to capture or poor enforcement. Taking these real world concerns into account informs the MFA’s account of ethical restraint.

\textit{The Content of Ethical Restraint}

As was stated earlier, the MFA endorses a Janus-faced orientation for business ethics, where managers face their own firms in a cooperative and trustworthy manner, and address their competitors in an adversarial manner. Stated in this way, it does not seem like the MFA differs much from approaches that claim that all businesses need worry about is profit, and following the law. To see how the MFA proposes more than this, we must recall that all the foregoing has been derived from the Pareto principle, which asserts that a state of affairs is preferable if one person is better off and no other person is made worse off. Firms are encouraged to pursue profit not for any intrinsic reason, but because doing so when others do so tends to create Pareto improvements. In this sense profit-maximization is endorsed, but endorsed instrumentally. Similarly, managers are meant to pursue firm profit, as opposed to self-interest, because the organization of economic activity within a firm is done to obtain Pareto improvements left unrealized by a price-mediated market. Again, the reason is instrumental and not principled: the shareholder does not have some moral primacy that makes their interest trump the manager’s (as a theory built on the celebration of the entrepreneur might). The manager represents others as opposed to himself because doing so allows for the cooperative organization needed to increase welfare.

\textsuperscript{15} This is, of course, assuming that the law is legitimate and does not require extraordinary behavior like civil disobedience.
In contrast, respect for the law is not to be done on the basis of some cost-benefit calculation. The law, in this view, restrains firms from engaging in activities that the price mechanism does not effectively de-incentivize. Thus, even if there is profit in breaking the laws in order to engage in these activities, doing so is to pursue profit in a non-preferred and socially harmful manner. Therefore, unlike profit and corporate cooperation, respect for the law has an ethical content in itself, because the activities prohibited therein are harmful on their face. Despite this difference, all of these are grounded in the same basic idea of efficiency. Respect for legal intervention and regulation as ethically substantive restraint, and pursuit of profit and shareholder primacy as instrumentally valuable, are both drawn out of the Pareto principle and the view of social welfare it advances.

It is with this move that we can better understand the thicker conception of business ethics that the MFA generates. The law, as Heath notes, is not a perfect instrument for protecting against socially harmful activities. Because of administrative costs, enforcement costs, and the difficulty of effectively detecting all forms of malfeasance, “the deadweight losses imposed through use of the legal mechanism can easily outweigh whatever efficiency gains might have been achieved through the intervention.”\textsuperscript{16} As a result, even after the law is used, many activities would still be legal and harmful for firms to engage in. This is where ethical leadership and decision-making kicks in: “ethical conduct in an extra-firm business context consists in refraining from using non-preferred strategies to maximize profit, even when doing so would be legally permissible. Put more simply, the ethical firm does not seek to profit from market failure.”\textsuperscript{17} This has been articulated by Norman as understanding business ethics in terms of “self-regulation,” in which ethical business leaders recognize the normative content of a regulatory scheme and act both in compliance with the regulatory scheme as it is enacted, and beyond compliance as it ought to be enacted: “you shouldn’t do X because there are clearly identifiable reasons why X

\textsuperscript{16} Ibid.
\textsuperscript{17} Ibid.
should be illegal, even though it is not in fact (or yet) illegal; profiting from X is a perversion of the market system itself.”

The result is an ethics in which competition, regulation, and normative self-regulation are all endorsed and, more to the point, endorsed on the same welfarist grounds.

One way to put this is to say that the MFA extends the theory of second best beyond institutional configuration of markets, and to the behavior of those acting within those markets. If the optimizing conditions obtain, then allocation can be done solely through a competitive market, and all market actors need to consider is maximizing their utility. When these conditions do not obtain, not only might we require different institutions, we might also require a different approach to the behavior of economic actors; we not only move away from unfettered markets, we also move away from profit-maximization and conformity to law as the only types of behavior likely to contribute to allocative efficiency. Those actors who have leadership positions within the firm have a particular duty to recognize this since the firm, and their positions within it, are created precisely to overcome market failure. To exacerbate market failure, or profit from it, is to go against the basis of their institutional role. In order to contribute to the efficiency of a second-best economy, market actors like corporate executives must avoid doing various things that would otherwise fetch them (or their firms) a profit, or do things that would otherwise harm their profits. This includes: 1) minimizing negative externalities, 2) competing only through price and quality; 3) reducing information asymmetries between firm and customers, 4) not exploiting the diffusion of ownership, 5) avoiding the creation of barriers to entry, 6) not using cross-subsidization to eliminate competitors, 7) not opposing regulation aimed at correcting market imperfection, 8) not seeking tariffs or other protectionist measures, 9) treating price levels as exogenously determined, and 10) not engaging in opportunistic behavior towards customers or other firms.

---

19 Heath 2004, 84.
This list helps to illustrate the critical nature of the MFA. The MFA is decidedly not an apologia for the status quo; it actually leads to the conclusion that the moral requirements of business managers are much more rigorous than is generally thought: were one to follow this list fully, business would not only have to refrain from legal forms of pollution and collusion, but advertising strategies that are not purely informational (as an outcome of 2, 3, and 10), or using economies of scale resulting from the size and scope of one’s enterprise to prevent other businesses from entering competition (as a result of 5 and 6). Thus even though it affirms profit as a legitimate aim of business and an adversarial ethic in competition that might at first seem ethically compromised, the MFA’s model of competition is an extremely aspirational one; were the MFA adopted by all businesses, the nature of the markets and profits would look extremely different from that which we find ourselves in now. Much of what is required by managers under the MFA still winds up being overly demanding in the context of the actual conditions of the market economy. Given this, we might say that the first order ethical requirement of managers is not to stymie efforts to bring a system into place that would allow the MFA to be followed. I will discuss this more below when discussing my critique and amendment of the MFA.

The MFA engages with the norms inherent in the practice of business so as to criticize those practices as we find them here and now, and show how such practices could be transcended in favor of a more just and better world. By beginning with the apparently amoral terrain of economic theory, the MFA takes these groundings and shows how they not only produce decidedly moral imperatives for those in business, but moral imperatives that provide the bases for a robust critique of contemporary business practice. Table 7.1 below summarizes the different types of behavior the MFA prescribes, how they connect to the institutions of political economy, and the economic theories to which they are connected.
Table 7.1. The Market Failures Approach to Business Ethics

<table>
<thead>
<tr>
<th>Ethical Orientation of Behavior</th>
<th>Guiding Idea</th>
<th>Corresponding Institution</th>
<th>Animating Economic Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competition</td>
<td>Market</td>
<td>1st Fundamental Theorem</td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td>Firm</td>
<td>Transaction Costs</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethical Limitation of Behavior</th>
<th>Constraint</th>
<th>Law</th>
<th>1st Fundamental Theorem and Theory of Second Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restraint</td>
<td>Business Ethics</td>
<td>Theory of Second Best</td>
<td></td>
</tr>
</tbody>
</table>

Values Other Than Efficiency?

Much more can be said on the merits of the MFA, but I will mention here quickly how it is distinguished from other prominent approaches to business ethics. On the one hand, as has been mentioned in passing, the MFA differs in significant ways from the shareholder value maximization camp of the Chicago School, challenging their emphasis on profit-maximization by highlighting the instrumental nature of the profit motive and, in turn, how profit-maximization can lead to market failure. Similarly, the MFA challenges stakeholder theories, which argue that corporations have moral obligations to all of their stakeholders, including customers, employees, suppliers, local communities and so on. Here, the MFA’s critique is the opposite of its challenge to shareholder maximization; it holds that stakeholder theory imports ethical standards and criteria foreign and antithetical to the corporation’s efficiency mandate. To follow the stakeholder program is to wrongheadedly give up on the welfare-promoting potential of the corporation in favor of something unworkable in the business environment.

20 Though for our purposes here the law largely plays a second best role, there still exists law in the first best world of the fundamental theorems of welfare economics in the form of property law, contract law, and the like.

21 Friedman, 1970.

And yet one wonders whether, in this position vis-à-vis stakeholder theory, the MFA hasn’t given up something important from an ethical perspective. There is an important intuition shared by many that the unethical nature of much business practice is created by precisely the adversarial and “norm-free” nature of the market with which the MFA begins. If business ethics is to mean anything, it ought to be the supplementing of this market—instrumental and useful as it may be—with a more robust sense of morality so as to blunt the sharp edge of capitalism or, if one wishes, to trim the sharp nails of the invisible hand. From this view the MFA fails in its project even before it starts. By looking to the logic of the market to derive the morality upon which market actors ought to act, one might say we have essentially capitulated to the economists and financiers; instead of attempting to challenge the hegemony of economic reasoning, it seems that the ethicists are ceding the ground to the economists themselves. In this view, the MFA’s emphasis on the market context that I highlighted above is actually its bane; it is pragmatic in the pejorative sense of being without principle, strategically aligning itself in such a manner so as to gain favor and resonance with the relevant economic discourse of the day.

As should be clear, I believe this is overstating the case against the MFA. While proponents of the MFA do build their system of business ethics up from what McMahon referred to as the “implicit morality of the market,” this is not the same as endorsing market morality as a standalone system of morals. As Norman notes, the morality of the market is but one of many competing values society must concern itself with:

A Paretian approach to business ethics cannot pretend to explicate all the relevant issues in the field—even though it is hugely significant that most legitimate ethical concerns about business activities do stem from the creation or exploitation of classic market failures...we can argue in favor of the design of any given market regulation or beyond-compliance standard by appealing to other socially desirable values it promotes. And we can surely criticize particular markets, or the activities of particular market actors (firms, employees, owners, customers, etc.), because they fail not necessarily or merely on efficiency grounds, but by the standards of other values and principles we care about. Of course, many such arguments will be flawed. But at least some justifications of this sort deserve to carry the day.24

---

The MFA, then, merely takes efficiency as one societal value, and attempts to show just how robust a system of morality can be drawn out from it. Yet, this does not imply that efficiency trumps all concerns. Other relevant considerations—like equality or justice—can be brought in from outside this system of morality, and markets can be regulated from without on their behalf.

The MFA’s claim is not that efficiency is the most important value, but that asking business ethics to concern itself with those other values is to be too demanding of the field. Concerns other than efficiency are the domain of, for example, political philosophy, which can demand the state make certain structural changes to markets so as to make them fairer or more just. Thus what proponents of the MFA demand is a less expansive notion of business ethics, and a greater understanding of how business ethics might fit within a larger scheme of morality and justice, or a “unified theory” of firms, markets, and states. Put differently, the MFA asks for a project of complementarity between business ethics and other fields of normative theory, not one of all-encompassing expansion. This is not merely a semantic distinction. Growing from Rawls’s idea that principles of justice apply only to the “basic structure” of society (the constitution and key institutions), the MFA appears to assume that the macro-level ideals of political philosophy are at best indeterminate, if not silent, on questions at less general levels, like business. Therefore, when addressing a meso-level institution like the business firm, the goal is not to bring justice to bear through an expansive scheme of business ethics, one that extends overarching principles of justice to the more local level. Instead the idea is to articulate a notion of business ethics consonant with a larger scheme of justice, implying attention not just to these overarching principles, but

also to the more local social role that business itself is meant to perform. For the MFA, this local concern is efficiency.

Though not stated explicitly, this normative schema is underwritten for the MFA by the second fundamental theorem of welfare economics described above. Markets, and market actors, can be geared entirely toward efficient allocation without downplaying other moral values because it is presumed that equality can be achieved through the tax-and-transfer redistributions of the welfare state, leaving the market to work as it will. Efficiency and equality can be achieved jointly through a division of labor between market and welfare state without the one impairing the other. Similarly, the MFA endorses a division of labor between business ethics and political philosophy, where the former works out the practical implications of the “implicit morality of the market,” and the latter concerns itself with justice in a wider sense. One need not demand that business ethics concern itself with justice in order to achieve a just society according to the MFA; a just society is better sought by letting business deal with the efficient allocation of resources, and leaving justice to those institutions better calibrated for achieving it, namely the state. This division of labor is depicted in Figure 7.1.

**Figure 7.1 The Place of Business Ethics’ within a Scheme of Social Justice**
The problem, of course, is that institutions like the state do not live up their theoretical potential any better than markets do. Indeed, as political CSR theorists have pointed out, quite often the geographical or market domain in which businesses are operating lack anything that could be called a statutory or legal regime. In these conditions no such justice-seeking institutions can be said to exist. Therefore, even if we grant that business ought to be in the business of maximizing efficiency—and therefore refraining from profiting off of market failure—we still want to know what businesses ought to do when background conditions of justice do not exist. While the MFA admirably looks at how markets fail to achieve efficiency and then suggests what this demands of the business corporation, I want to pose this critical-but-friendly question: how does the failure of justice-procuring institutions to function as we would wish them to function change the ethical demands of efficiency-procuring institutions like the corporation? Does it suggest a change in focus?

Let us return to the sports metaphor. We ask athletes to compete, and compete singularly, and to follow the moral codes derived from the ideal of competition itself, like ideas of sportsmanship and fair play. Other moral concerns enter the sphere of competition not through the contestants but through non-competitive actors: the rule-making bodies of the league, ringside physicians, and referees. As Joyce Carol Oates describes in her celebrated *On Boxing*, “the referee is our intermediary in the fight. He is our moral conscience extracted from us as spectators so that, for the duration of the fight, ‘conscience’ need not…be a factor in the boxers’ behavior.” This allows us to have values like safety and fairness established while also allowing for the fight to be as competitive, athletic, and entertaining as possible. Similarly, the second fundamental theorem allows for us to offload other moral concerns to the state and law when it comes to market competition, enabling the achievement of social equality and efficiency.

---


But what if the ringside doctor is paid off by one corner? What if the match is purposely moved to a particular state because its licensing policies are far too lax (allowing, say, Larry Holmes to fight Muhammad Ali, despite his obviously failing health)? Is it still ethical to participate in the fight, or to do so in the same manner as one would were the background conditions more optimal? Similarly, if the welfare state does not function well or does not exist at all, or if historical inequalities and injustices continue to persist despite the efforts of the welfare state, does this change the ethical content of business ethics? It seems that the answer ought to be yes.

In the remainder of this chapter I sketch out the concept of justice failure, which explains this intuition in terms of the MFA and the economic theories from which it is derived. I contend that the institutional division of labor between efficiency and equality upon which the MFA rests is not totally tenable. Because of this, I argue that we can demand that businesses concern themselves with values like equality or fairness (and that business ethics ought to try to spell out how) in the absence or failure of other institutions to do so effectively. This concept allows us to understand how concerns other than efficiency enter into what the MFA demands of business actors. It also provides us with some basic ideas of how such values ought to influence the way relationships are structured within the corporation.

**The Concept of Justice Failure**

In the previous section I’ve highlighted how the first and second fundamental theorems are key to the MFA and its presumed division of intellectual labor in normative political economy. Yet we’ve also seen that the MFA requires the theory of second best in order to explain how efficiency and regulation are not at odds. Because the assumptions of the fundamental theorems don’t actually hold, efficiency requires moving away from pure market mechanisms in our second best world. In other words, the second best theory shows how markets and efficiency are closely related but, because markets and the world don’t operate the way in which economists assume, other institutions are required. For this reason
the MFA can assert both its adversarial side and its cooperative rule-following side on the basis of efficiency.

However, while the MFA takes the theory of second best seriously when it comes to the first fundamental theorem, it does not seem to apply it to its own reliance on the second fundamental theorem. Just as we depart from free markets so as to achieve the second best possible efficient allocation, there is good reason to think that we should depart from purely statist approaches to distribution in our pursuit of justice or fairness. We demand institutions like the state act in particular ways so as to achieve second best efficiency, and we ask corporate executives to constrain and restrain themselves according to the spirit of those actions; perhaps there is good reason to think that market actors ought to be asked to shoulder some of the burden of justice in order to achieve second best social justice.

Another way of putting this is that the MFA’s scope of critique is too narrow. The project is built on the norm of efficiency because efficiency is the implicit morality of the market. It seems, however, that there is a more substantive implied morality when it is not simply the market being considered, but the market and its place within a larger scheme of social cooperation. Unless the MFA wishes to assert efficiency as the most important aspect of justice (as opposed to one component part of it), the morality of pursuing efficiency must rest on the background of a larger scheme of social equality; the first fundamental theorem presents a compelling moral case for the market because the second fundamental theorem shows how it can be pursued without sacrificing equality. When this latter assumption falters—when we can no longer assume that the welfare state is actually handling equality—the rationale for basing a moral code solely on efficiency seems less obvious. Again, here one can respond by asserting that efficiency is in fact the most important value either in principle or in our second or third best
If we are unwilling to assert this—or if wish to understand what our aspirational moral duties in business ought to be even if, here and now, efficiency as the main value must suffice—then we must introduce other values into business ethics.

To this end I suggest justice failure as an analogue to market failure. I use the term “justice failure” to distinguish the concept from “government failure.” Government failure, like market failure, refers to sub-optimal efficiency resulting from government action; justice failure, although also concerned with state action, is not about the failure to achieve efficiency, but the failure to achieve some morally desirable form of equality (whatever that notion of equality is). Whereas market failure suggests failures of the market to achieve all possible Pareto improvements, justice failure suggests failures of the welfare state to achieve all possible movements toward equality consistent with efficiency. Similarly, just as the MFA argues that business must refrain from intentionally profiting from market failure in order to avoid contradicting the implied morality of the market, businesses must also refrain from intentionally profiting off of justice failure in order to avoid contradicting the implied morality of the larger social scheme that justifies the implied morality of the market.

This is intentionally vague. Whereas the concept of market failure comes replete with a particular understanding of efficiency (the Pareto standard, in contrast to utilitarianism or cost-benefit analysis) and is therefore more conclusive in its prescriptions, I present the concept of justice failure without a corresponding metric of equality. Without some specific content or definition of justice or equality, this will be a fairly broad and generic suggestion. If one adopts some particular notion of equality one would get a more determinate understanding of what would constitute justice failure. The main claim is that some idea of justice or equality should enter into business ethics in a critical approach to political

30 While I strongly doubt any proponent of the MFA would assert the primacy of efficiency in principle, Heath seems to suggest that practically speaking, efficiency might be the best value to pursue in our non-ideal world.
32 This does not include failures to “level down” toward equality, which is why I add the qualification “equality consistent with efficiency.”
economy; this claim can be made while being more or less agnostic on which particular idea of justice or equality that should be. One can take some transcendent vision of justice to be universally correct and use it to calibrate business ethics, or one can proceed contextually—taking a particular society’s concept of justice or equality as the standard to determine justice failures. However, given that our current non-ideal world furnishes us with numerous situations that would qualify as justice failures under almost any account of equality or theory of justice, it seems that we can perhaps start with those and not worry too much about about articulating a full theory of justice or equality. The larger point is to understand how macro-level principles of social justice can be seen as shaping the meso-level program of business ethics not just from without, but through its content.

This conceptual move may not seem obvious for at least two reasons. First, since the theory of second best is only explicitly about efficiency and not equality it is not clear how relevant or applicable this theory is for the second fundamental theorem. Yet, the second fundamental theorem is parasitic on the first: the claim that a fair and efficient allocation can be reached through initial redistributions is based on the first theorem’s conclusion that markets will always reach Pareto optima in the first place. If we have disabused ourselves of this latter belief because of second best considerations then it seems we must also disabuse ourselves of the former: that somehow the best possible egalitarian distribution is attainable purely through the tax-and-transfer systems implied by the second fundamental theorem.

The second trouble for this conceptual move is motivating the assignment of duties to the corporation in particular. Even if we grant that justice failures exist and that such failures ought to affect the way in which market actors act, it is not obvious why I have singled out corporations as needing to take on such onerous responsibility. For market failures such an assignment makes sense: because firms and corporations exist to help deal with the problem of market failure, it makes sense to contend that they ought not to exacerbate that which they are meant to mitigate. Because corporations exist primarily

for reasons of efficiency specifically, and not social justice generally, such an argument cannot be used to justify a corporate duty with regards to justice failure. There are two additional reasons, however, why the duty to consider justice failures ought to attach to corporations. The first is pragmatic. In this view we might say that though actors generally ought to consider justice failure, corporations wield far more power, influence, and volume with regards to things like hiring, investment, governmental lobbying, and so forth. For pragmatic reasons then, corporations have duties that individuals do not—though it would be no bad thing were individuals to take justice failures into account as they go about their business.34

The other related reason why we should think that corporations ought to have such a duty is that in many aspects of social life corporations have taken on roles formerly occupied by the state. This insight informs the basis of political CSR scholars who argue that given our post-Westphalian world the division of labor that used to exist between corporation and state is no longer wholly tenable. As a result corporations have taken on roles in fashioning regulation, governance, and enforcement in policy realms abdicated by the nation-state.35 I generally think this paradigm overstates its case a bit and exaggerates the extent to which the state has been displaced. Still, there is more than a little merit in the empirical claim that corporations have a regulatory and governance function that extends beyond the more traditional purview of efficiency and profit. Because they are performing functions in lieu of states, it is sensible to ask that they take on the ethical duties that result from the state’s incapacity or inability to do the work of securing the egalitarian basis for the just pursuit of profit. The problem of justice failures therefore falls onto the corporation to consider.

34 It is an open and interesting question whether other institutional market actors—like labor unions—might have such duties as well. My intuition is to think that they do, though I do not explore that here.
Justice Failure and Liberal Neutrality

It is worth noting that there is a large extent to which we already apply the logic of second-best to questions of justice or equality, as can be seen in affirmative action policy. Critics of this policy will sometimes refer to the policy as racist or sexist in that it makes legal the use of race or sex (or sexual orientation, religion, etc.) in deciding access to education or employment. While this type of argument may very well betray a complete misunderstanding of what constitutes racial or sexual oppression, it also betrays a misunderstanding of the argument for affirmative action. There are few advocates of affirmative action on the basis of race who don’t also hold that Harlan’s vision of a “color-blind constitution,” as expressed in his dissent in *Plessy v. Ferguson*, is the ideal to which we ought to strive. This, as Frank Michelman notes, seems to be the case of Justice William Brennan, who thought Plessy’s prophetic dissent was one of the best in judicial history, while also joining on the historic *Bakke* decision (amongst others) that laid the jurisprudential basis for affirmative action policy.36

What accounts for this? The answer, of course, is that both goals—a color-blind constitution and affirmative action policy—are justified by the same value, something like equality or social justice. Color-blindness in the law is the ideal, but its achievement depends upon a number of conditions obtaining. While I don’t have the time or expertise to elucidate all of these, surely the American case is one where these conditions do not obtain. Regardless of its success in achieving its aims, the intent of affirmative action is to achieve second-best racial equality or justice. Given the failure of governments, markets, universities, militaries and so on to achieve the background conditions necessary for the colorblind constitution—that is, given the existence of justice failure—affirmative action advocates argue that we must move away from the color-blind approach and toward an approach that actively combats the problems of racial exclusion. I say that this is the same logic as the second-best approach to political economy applied to legal equality because the overall goal remains constant; just as we move

from a laissez-faire market to a complex of markets, firms and statutory intervention in the service of efficiency, what changes in the move from color-blindness to affirmative action is the recognition of empirical contingencies that would make the former unhelpful or counterproductive in the pursuit of racial legal equality.

Unger captures this logic well in his discussion of how a supposedly neutral political process might very well lead to further entrenchment of social inequality:

[The goal] is to create a political process that can serve as an impartial device for summing up the wills of individuals about the proper role of the state...it is carefully designed to prevent manipulation by transitory and inflamed majorities who, misguided by the demagogues or fools, might wreck the underlying pure structure of power and coordination. But precisely because government cannot easily disrupt the social order, it becomes the victim and protector of this order. It turns into a pervasively biased method of collective choice. The search for the neutral method for summing up the opinions of the citizenry diverts us from the more realistic attempt to create a polity that would in fact be more open to self-revision and more capable of dismantling any established or emergent structure of social division and hierarchy.  

The liberal idea of law, from this perspective, attempts to establish a neutral and objective system of politics and law that applies to all equally and thus ensures that this system establishes equality amongst the citizenry. Yet, when the assumptions of the model don’t obtain—when legal and material resources are distributed unequally throughout society on class and racial lines, for example—the application of an objective and formally equal legal system will further exacerbate extant inequality. This was best and most famously articulated by Martin Luther King Jr. in his chastisement of white moderates for their resistance to his method of civil disobedience: “I had hoped that the white moderate would understand that law and order exist for the purpose of establishing justice and that when they fail in this purpose they become the dangerously structured dams that block the flow of social progress.” Sticking to the liberal ideal of formal justice in the face of structural injustice can often have the effect of amplifying the latter. To achieve equality here, it seems the best bet is not to approximate the ideal, neutral procedure as

closely as possible. Instead, we should recognize that the non-realization of certain assumptions might require letting go of other commitments in order to most closely attain the substantive ideal of equality (at the expense of the procedural ideal of objective formalism).

So, when we are confronted with a market system that is meant to allocate goods and services most efficiently, but which is set against a background in which the welfare state (if one exists) has failed to redistribute resource endowment effectively, or offer opportunities of training and access equally, what are we to do? To say that the market must stay the course of efficiency is to resign oneself to the possibility that one might further perpetuate the initial inequality in the process. It also does not seem plausible to try to fold these concerns into the more substantive duties that the MFA derives from efficiency: while a duty to refrain from profiting off of externalities or information asymmetries is quite a demanding and admirable proposition for business ethics, it is of a different character from, say, dedicating resources to overcome systematically unequal educational opportunities. The former is proposed because it gets us as close to the Pareto frontier as possible, absent the complete markets and complete information necessary for the market mechanism to do so; the latter, on the other hand, is fundamentally about win-loss transfers, and thus involves comparisons of distributive states that are Pareto non-comparable. The justice failure approach explains how the failure of the welfare state to achieve a particular level of equality affects the moral duties of business actors, the same way the failure of political and social conditions affects the legal approach to questions of race or sex.

This links up the idea of justice failure with the MFA in helpful ways. The MFA is importantly both reconstructive and prescriptive: it gives both a normative account of why certain institutions exist (the market, the firm, and the regulatory state) while also offering direction to how businesspeople ought to act. Similarly, I think, the idea of justice failure might offer this dual purpose. To continue with the example of anti-discrimination law and affirmative action, justice failure better explains why such policies ought to exist in the economic realm. It is not obvious that they can be explained on the efficiency terms the MFA uses to explain the ethical basis of following the spirit of regulatory law—even
though there are surely efficiency-gains to be had from ensuring equal access, justifying such policy in terms of efficiency misses the point. Yet affirmative action and anti-discrimination law deviates from the MFA in how it brings justice to bear. Whereas the MFA asks justice to be brought into the equation through background redistribution and welfare programs, affirmative action policy explicitly affects, and intervenes in, the way in which businesses act. They are not simply legally constrained from profiting off of market failure, but morally compelled to bring values (equality, diversity, etc.) into their business and hiring calculations.

The idea of justice failure helps us make sense of this. Due to the historical failure of the United States to effectively integrate descendants of slaves into its most important political, economic, and cultural institutions, we have a case of justice failure that demands that business practice take on this historical burden as well. Absent the redistributive (and, perhaps, reparative) institutions that could establish something like an equality in initial positions upon which an efficient market could work its magic, we demand that businesses follow particular rules of action that reflect this historical fact and institutional deficit. Of course, legal and statutory interventions here will be blunt instruments; as I have been arguing, the necessity to intervene legally into the decision-making processes of businesses comes about because the state has failed to achieve background conditions that would promote equality. In a sense then we are asking for statutory action to make up for the inefficacy of statutory action. This isn’t necessarily as absurd as it sounds: requiring businesses to engage in affirmative action is an order of magnitude more manageable than using large-scale transfers of wealth and government programs to solve the structural factors that perpetuate racial inequality and injustice. However, it should make us aware of just how limited state action might be.

40 Despite the variety in ethical descriptions of affirmative action, all seem to agree that fundamentally it is about the failure of society to integrate a particular “protected class,” either affirmatively working toward their equal treatment, or demanding special treatment in order to achieve equal results. (Guy Adams. 1997. “Racism, Community, and Democracy: the Ethics of Affirmative Action.” Public Productivity & Management Review 20(3): 245).
For this reason “justice failure” would also imply something further with regards to business ethics: not merely following the laws intended to promote equality, but also following the spirit of the law, and actively combatting the inequality that business leaders find in front of them. This is analogous to what the MFA demands. Because regulatory interventions will not always be effective (or enacted), business actors must self-regulate and avoid profiting off of market failure in order to contribute to efficiency; similarly, the normative logic of justice failure does not merely suggest enacting and following affirmative action or anti-discrimination law, but avoiding business strategies that take advantages of their non-enactment or non-enforcement. Cheryl Wade, for instance, suggests something akin to self-regulation in her call for stricter self-monitoring of employees’ compliance with anti-discrimination law.\(^{41}\) Other more substantive proposals would suggest procedures and practices for ensuring greater representation of people of color in positions of power, amongst the client base, interpretations of fiduciary duty that emphasize a duty of care and empathy, and dedicating resources to training those who might otherwise not have the types of educational opportunities to succeed in the market.\(^{42}\) It might even require that members of such groups be paid higher wages, though there seem to be all sorts of prudential concerns (e.g. creating envy and hostility in the workplace toward precisely those groups) that would complicate pursuing such a measure. Of course, one might think of even more drastic measures to demand of businesses when it comes to issues of racial injustice, in addition to issues of gender, class, or a host of other social inequalities that bring the very non-ideal functioning of the welfare state into stark relief. The point, however, is that such proposals make sense to include in a program of business ethics, and that they are best made sense of through the concept of justice failure,


through businesses’ moral duty to refrain from exacerbating those inequalities that contradict the social basis of the just pursuit of efficient allocation.43

**Practical Elements**

It must be said again that, like the MFA, a program of business ethics built around the concept of justice failure is deeply aspirational. If any one firm were to refrain unilaterally from non-informational advertising, or attempted to alleviate huge historical injustices, that firm would not likely be long for the market. Realistically, these problems are all bigger than a single market actor, which operates in a competitive and unforgiving market; as a result, even if most businesses were interested in living up to their ethical obligations, there would likely be a collective action problem involved in remaining so, since each would have reason to fear that the other would do the competitive-but-inefficient-and/or-unjust action to capture a profit. Furthermore, because the state’s inability to deal effectively with these problems is what gives rise to the extralegal ethical obligation in the first place, there is little likelihood of state-imposed regulation to solve this collective action problem.

So, while it is hard to imagine such a vision being viewed as an apologia for the status quo, it might very well be criticized as overly idealistic. I am comfortable with this. The goal is to articulate a view of business ethics that is appropriate to the particular institutional setting of the market, not to facilitate its easy implementation by dulling its critical bite. Even if such a vision is hard to reach here

43 Throughout this section I have had in mind more social and progressive understandings of equality, and what such concerns would demand. However, I think the concept of justice failure would apply to a range of worldviews. I follow Sen in thinking that equality is a generic feature of virtually every normative social theory. The question isn’t whether or not a theory is egalitarian or not, but what social goods the theory demands to be equalized and for what goods the theory will tolerate distributive inequality (Amartya Sen. 1996. “On the Status of Equality.” *Political Theory* 24(3): 395). All this is to say: the concept of justice failure could apply to even a libertarian or socially conservative theory of justice, since those theories will have concepts of equality that might be frustrated by malfunctioning institutions. If I might speculate about perspectives that I share little in common with: a libertarian might view the failure to fully establish full equality of formal property rights (which result from, for example, redistributive taxation) as a sort of justice failure; she might then argue that corporations ought to be given special tax privileges as a way to counter the inequality of liberties that the welfare state is establishing. As should be clear by now, such a view does not appear interesting or persuasive to me (and it may very well not be endorsed by actual libertarians). All I mean to say is that although my example of justice failure tend toward the left, the concept should still apply to other worldviews.
and now, it is still important to articulate clearly what we ought to be doing, lest we come to confuse it with what we happen to be doing.

Still, there are two ways we might imagine such an ethical vision being put into practice. The first is that businesses ought to pursue the types of regulation and agreements that would allow them to do the ethical thing. We might think of it this way. Absent the context in which a business could follow such ethical duties without being put out of business, the ethical duties implied by market failures and justice failures are supererogatory. However, there will be no great loss or harm imposed on market actors by pursuing the context in which they could act out their ethical obligations. We might say, then, that the first ethical obligation of businesses is not to lobby against state and non-state regulation that might make the market environment more hospitable to ethical action; the second would be to actively lobby in favor of such action.

Another practical way to implement such a vision is to effect such ethical action within the relationships internal to the corporation. As we have seen, the corporation introduces norms as a means to counteract the opportunism cultivated by the market. However, the kinds of norms that should be used remains an open question. As we have seen, managerial authority is not a given, and mandatory democratic control is not a panacea. Though the extent and reach of these norms will be constrained by the minimal and maximal normative horizons discussed in chapter 4, we might demand that non-efficiency values apply to these relationships, as I argued in the last chapter in my discussion of the various institutions that might give cooperative governance a better chance of success. However, even in that scheme, I contended that we must allow for a plurality of governance structures. The question arises here, though, whether one has an ethical obligation to try and foment particular arrangements of relationships within the corporation as a means of countering justice failures.

On this note advocates of the MFA have been largely silent. Though the idea has been raised that such ethical obligations might require particular arrangements within the corporation, there has been no
stance indicated one way or another. Here I would like to make a modest suggestion of a more stringent approach to the norms which govern intra-firm organization given a backdrop of social cooperation: the norms adopted to facilitate production within the corporation should not directly contradict the spirit or integrity of the values being implemented by the state or other social entities which structure the marketplace; that is, they should not directly contradict the core liberal values of individual liberty, equality, or democracy, or the core values that individual liberal societies have collectively come to endorse.

For example, most welfare state democracies are committed to some notion of equal opportunity, that a social system is fair only insofar as no members of society are excluded from opportunities due to concerns other than merit or work effort. Welfare states attempt to do this by, for instance, providing welfare support for people in extreme poverty, and through programs that subsidize educational opportunities for those who might not otherwise afford it. Following Norman’s idea of self-regulation, in which businesses ought to commit themselves morally to the spirit of the laws they are obliged to follow legally, I would argue that norms ought not to contradict this idea of equal opportunity and access.

As an example, take housing cooperatives. These are businesses in which ownership rights of the building are allocated to those who live in the building as opposed to a landlord. On its face, housing cooperatives are extremely egalitarian and democratic since they apply the one-member one-vote rule to all resident-owners, thus securing equal relations and distributions of power amongst them. Yet, such norms can actually run in the face of social egalitarianism when they are applied to extremely rich apartment buildings, as is the case in many New York City cooperatives. In this case egalitarian norms within the firm are not being used for the purposes of social egalitarianism, but rather are used to circle the wagons, so to speak, and keep people of lower classes out of the building. Here I would want to say

Norman notes that one can distinguish “Paretian justifications of the competitive marketplace, on the one hand, from very different norms and justifications relevant to relations within a firm, on the other” (Norman 2014, 26-27).
the norm being used to organize the housing firm is out of sync with the norms we endorse as a society. This doesn’t mean housing cooperatives as such are necessarily out of sync, however. For instance, also in New York City, there are many Housing Development Fund Corporation cooperatives which restrict membership to those earning less than a certain income threshold and which also put caps on resale values. Here, the same democratic egalitarian norms are being implemented but in a manner which is in harmony with overall goals of equal access, equal opportunity, fair housing, and the like.

I use the example of cooperatives because it follows from my discussion in the last chapter, and tempers an overly sanguine view of cooperatives and their egalitarian potential. The more generic point is that what appear to be egalitarian or fair norms when only the individual firm is being taken into account, might look less fair or less egalitarian when that particular firm is looked at from the societal point of view; more equality might not mean more egalitarian, and more democracy might not mean more democratic While this is very vague, it would seem that those norms other than efficiency that further the wider social values endorsed by society are the best norms to be implemented within firms. Programs for increasing diversity, coupled with pay equity schemes would seem to be in good keeping with such liberal values in light of historical circumstance, and are things that could be enacted upon by corporations without losing much in the way of competition. Corporate governance structures like cooperatives that are used to empower historically disadvantaged groups or workers who are often structurally disadvantaged in bargaining, are to be lauded; those same structures when used to protect privileged positions are to be criticized.

The purpose here is not to spell out in detail all of what is required by corporate managers, or whether any particular corporate organization is ethically mandated by my approach. It is rather to give a framework for how managerial decisions might be assessed from an ethical perspective, and to give

---

some ideas on how third or fourth best decisions might be made in light of this framework. In the context of the institutional and legal reforms discussed in the last two chapters, a managerial and corporate culture in which this ethic was followed would bring the corporation more in line with core liberal values without having to sacrifice the corporate contribution to efficiency.
Conclusion

My analysis started with some basic presuppositions that represent the underlying normative commitments of the liberal democratic market societies we find ourselves in. I took these to be: 1) market-based economies are beneficial to society because they maximize efficiency and, in turn, increase human welfare, and 2) liberal democratic norms—individual liberty, formal equality, and democratic decision-making—are preferable to other norms around which social cooperation could be organized. While providing justification for seeing these norms as at least plausible, if not powerful, I did not endorse them as necessarily correct or good. Rather, the point was to start with the normative presuppositions intrinsic to our social structure and institutions and show that the corporation—taken as a significant and important part of this institutional structure—does not fit comfortably therein; its intentional and bureaucratic mode of organization runs counter to the way we conceive of the market’s efficiency, and its hierarchical and collective nature runs counter to our liberal democratic commitments.

As a result, a political theory of the corporation is needed to give an account of whether and how the corporation could be made compatible with such values. I contended that such an account would entail answers to six questions. The first three questions address the corporation’s efficiency and constitute what I call the “productivity component” of a theory of the corporation. They are: 1) The Question of Efficiency (“how and why is the corporation an efficient way of coordinating economic activity?”); 2) The Question of Constraint (“what are the limits and constraints of corporate organization?”); and 3) The Question of Constitution (“which people and what relationships fall within the domain of the corporate and which do not?”). The final three questions address the corporation’s
relationship with liberal democratic values and constitute what I call the “normative component” of a theory of the corporation. They are: 4) The Question of Corporate Law ("what ought to be the legal limits of corporate authority with respect to its internal affairs or, put differently, what are the grounds for contesting corporate authority and calling it to account?"); 5) The Question of Corporate Governance ("Who ought to make decisions in the corporation, on whose behalf, and why?"); and 6) The Question of Business Ethics ("Upon what normative considerations, and according to which criteria, ought corporate decision-makers base their decisions or, how ought we appraise the executive decision-making of corporations?"). Taken together, answers to these questions provide an understanding of the corporation that can explain how it fits both within the efficiency-mandate of the market and liberal values we associate with political society.

I began with a historical reconstruction of the Chicago School’s account of the corporation, which I take to be the most influential and important political theory of the corporation. I argued that this theory begins with a transaction cost account of the firm and then flips the theory on its head by conceiving the corporation as a privately-owned market. As a result, its efficiency can be accounted for by the same explanation that accounts for market efficiency, namely the more-or-less spontaneous cooperation that makes the best use of available information. In table 8.1 below I summarize how the Chicago School answers the productivity component questions:
Table 8.1 The Chicago School’s Productivity Component

<table>
<thead>
<tr>
<th>Productivity Component Questions</th>
<th>Chicago School Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Question of Efficiency</td>
<td>The corporation contributes to efficiency by facilitating party-relevant transactions through the effective use of specialized information</td>
</tr>
<tr>
<td>The Question of Constraint</td>
<td>The corporation’s efficiency is constrained by the costs of bureaucracy inherent to such organization.</td>
</tr>
<tr>
<td>The Question of Constitution</td>
<td>The corporation is a legal fiction standing in for the nexus of contracts with which all parties of the corporation implicitly transact.</td>
</tr>
</tbody>
</table>

By seeing the corporation as a nexus of contracts that are voluntarily entered into for the benefit of each party, the Chicago School sees no problem on the basis of liberal and democratic norms, and therefore offers a straightforward account of the normative component. The corporation is consonant with such values in this view because it is simply an outgrowth of individual choice and contract, a spontaneous, albeit legally enabled, expression of individual freedom. I summarize how the Chicago School answers the questions of the normative component in Table 8.2 below.

Table 8.2 The Normative Component of the Chicago School

<table>
<thead>
<tr>
<th>Normative Component Questions</th>
<th>Chicago School Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Question of Corporate Law</td>
<td>The degree to which managerial decision can be legally challenged should be as minimal as possible. Fiduciary duty is to be interpreted with a maximum of deference for managerial decisions. The corporation has no legal purpose other than what is explicitly state in contract.</td>
</tr>
<tr>
<td>The Question of Corporate Governance</td>
<td>There is no authority to speak of within the corporation; the corporation is merely contracted relationships held in check by markets. Despite the corporation’s orientation toward shareholder primacy, “shareholder democracy” is a misnomer; shareholder vote mechanisms exist solely to specific contracts, not to legitimate power within the corporation.</td>
</tr>
<tr>
<td>The Question of Business Ethics</td>
<td>Unless explicitly told to pursue some other end, managers have an ethical duty to maximize shareholder value.</td>
</tr>
</tbody>
</table>

I then argued that this account of the corporation is unsatisfactory. To view the corporation as simply a nexus-of-contracts is to focus on a particular explanation of why
individuals cooperate within a firm by ignoring the role that norms and social scripts play in determining how they cooperate. Scripts of cooperation, norms of managerial hierarchy, habituation to taking orders—all of these play crucial roles in functioning organizations, and explain how the corporation is able to reduce transaction costs that arise in the market; that is, they are used to overcome the opportunism that causes the under-supply of many goods and services on the market. As a result, to take preferences as strictly exogenous—as Chicago’s contractarian account does—misses how corporations achieve efficiency-gains precisely by intrinsic and/or extra-voluntarist means. In contrast to the Chicago account I offered my own account of the productivity component based on what I called a “norm-governed productivity” account of corporate efficiency, summarized in Table 8.3 below. This account was based on a more socio-theoretic reading of Ronald Coase’s work and the theory of transaction costs that has been influenced by him.

Table 8.3 The “Norm-Governed” Productivity Component

<table>
<thead>
<tr>
<th>Productivity Component Questions</th>
<th>“Norm-Governed Productivity” Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Question of Efficiency</td>
<td>The corporation contributes to efficiency by overcoming transaction costs through cooperative scripts, norms, and extrinsic motivation.</td>
</tr>
<tr>
<td>The Question of Constraint</td>
<td>The corporation’s efficiency is constrained by the non-scalability of such explicitly cooperative schemes.</td>
</tr>
<tr>
<td>The Question of Constitution</td>
<td>The corporation is defined by the realm of generally norm-oriented, as opposed to generally consequence-oriented, cooperative production.</td>
</tr>
</tbody>
</table>

This account of corporate efficiency helps us understand why the Chicago School’s normative component is problematic; the fact that authority and power within the corporation are used to facilitate more efficient kinds of cooperation now does not mean that such authority is justified or that individuals would not prefer some other less hierarchical form of organization. As a result I have argued that the corporation must take liberal democratic norms into account in their organization and decision-making. However, this analysis also
helps us understand the kinds of constraints corporations face. Corporate bodies must promote efficiency, not merely out of a moral concern for social welfare, but also as a matter of fact; corporations must organize themselves such that they can capture efficiency-gains better than the market can. Corporations don’t merely compete with other corporations in the market; they compete with the market itself. This has important ramifications for a normative theory of the corporation. It is tempting to think that we ought to legally require corporations to be organized in all sorts of ways that better fit with our notions of fairness or equality.

However, such demands are only as viable as the corporation remains comparatively efficient; if corporations cease to provide efficiency gains over the price mechanism then people can always resort to private contracting. I therefore contended that the corporation creates a space for normative considerations that is bounded. On the one hand these norms cannot be so thin that they are unable to overcome the distrustfulness and the opportunism that is cultivated by market arrangements; I refer to this as the minimal normative horizon. On the other hand, the norms cannot be so thick that they force the corporation to underperform in the market; I refer to this as the maximal normative horizon. Within this space, norms can influence the way the corporation is organized and oriented without sacrificing its requisite efficiency. The trick, then, for bringing liberal democratic values to bear on the internal relationship of the corporation is to not only use the compulsory force of law to effect a more palatable regime of corporate organization; we must also, and perhaps primarily, arrange background institutions in order to widen the space for intra-corporate norms, and to cultivate an ethical disposition toward respecting the institutional context in which market activity takes place.
I argued that legally we should conceive of the corporation as having a status distinct from both the individuals that constitute it and the state that grants it privileges. Instead, we should think of the corporation as a relational entity whose particular relational constitution makes it a “real entity.” These crucial relationships, I contend, ought to be structured by law and corporate governance so as to comport with and support liberal democratic values. However, because these liberal democratic values include a healthy amount of respect for individual choices, the types of compulsory mechanisms available must be carefully articulated. I suggested that corporations ought to be compelled to organize themselves in manners that befit the values they espouse, and that a variety of regulatory and institutional background conditions ought to be introduced to encourage corporations to organize themselves as cooperatives. Finally, given that such a regime will still involve managerial prerogative and shareholder-oriented corporation, cultivating certain ethical dispositions for corporate decision-makers is important. I argued that from the perspective of ethics and social justice, corporations should be sensitive to the larger institutions that serve to promote efficiency and justice, and should avoid taking advantage of their underperformance in these areas; they should avoid exacerbating or profiting off of market failures and justice failures. Taken together, these comprise my normative component of a theory of the corporation, summarized in table 8.4 below.

**Table 8.4 My Normative Component**

<table>
<thead>
<tr>
<th>Normative Component Questions</th>
<th>My Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Question of Corporate Law</td>
<td>The corporation is a “re(lation)al entity, with legal status and obligations distinct from its members. Managers are not owed the deference or license implied by the business judgment rule and the at-will doctrine. The externalization of risk that limited liability creates ought to be mitigated through mandatory bonding or insurance</td>
</tr>
<tr>
<td>The Question of</td>
<td>A plurality of governance structures is to be expected. However, legal and</td>
</tr>
<tr>
<td></td>
<td>institutional conditions are introduced to encourage corporations to organize themselves as cooperatives. Finally, given that such a regime will still involve managerial prerogative and shareholder-oriented corporation, cultivating certain ethical dispositions for corporate decision-makers is important. I argued that from the perspective of ethics and social justice, corporations should be sensitive to the larger institutions that serve to promote efficiency and justice, and should avoid taking advantage of their underperformance in these areas; they should avoid exacerbating or profiting off of market failures and justice failures. Taken together, these comprise my normative component of a theory of the corporation, summarized in table 8.4 below.</td>
</tr>
</tbody>
</table>
Corporate Governance

institutional measures ought to be introduced so as to encourage governance rights to be assigned to the patron group that relates most closely to the corporate enterprise. Voting and elections in shareholder oriented corporations must be taken more seriously.

The Question of Business Ethics

Managers can pursue corporate profit but in a manner that respects the letter and the spirit of regulatory law. They should avoid profiting off of, or exacerbating market failure and justice failure.

Methodological Reflection

By way of conclusion it is worth reiterating the scope and bounds of this study in order to better understand the nature of my claims. As suggested in the introduction and the first chapter of this dissertation, this critical study can be seen as constrained in two ways. In the first instance, the main subject of the study was the Chicago School and its theory of the corporation. Instead of attempting to draw up a full-fledged theory of the corporation from scratch, I began with what I took to be the most dominant and influential account of the corporation in scholarly and policy discourses, and then subjected it to critique. Politically, there is good reason for this. If one wants a normative theory to have any practical uptake, it makes sense to begin with the discourse that governs actual practice. Intellectually, such a strategy is motivated by the belief that there is often some reason for why social institutions and structures are ordered as they are—which isn’t to say that there is always some good reason for it. If we wish to achieve or articulate a better social order or scheme of social cooperation, one must start with understanding why the current system is the way it is, in order to understand how it can be transcended.

The second way in which this study was constrained, was by accepting the normative values that I took to be implicit to liberal democratic market societies. One could take the Chicago School as the subject of criticism while still using some other set of values – something perhaps more substantively egalitarian or communitarian– as the grounds of such
a critique. Instead, I constrained my normative presuppositions to two fairly conventional assumptions: 1) liberal democratic norms and institutions—i.e. individual primacy, formal equality, and democratic decision-making—are preferable to other alternatives, and 2) there is a good form of capitalism that satisfies our intuitions regarding justice. Thus, instead of trying to derive non-controversial normative criteria from accepted first principles, I take the norms implicit to our political and economic institutions and ask whether the reigning justificatory account of the corporation can be made sense of on those grounds.

As I have suggested, all of this is part and parcel of a methodological commitment to immanent critique. Instead of trying to find a vantage point outside ourselves and our practices from which to conduct social critique, immanent critique starts from the norms underlying those practices themselves. We might say this study can be understood as an immanent critique in two distinct but related ways. On the one hand, I have offered an internal critique of the Chicago School by trying to show that its account of the corporation fails in terms of its own emphasis on efficiency. This can be understood as an immanent critique of social institutions (and not just an internal critique of legal and economic theories): insofar as the contemporary regime of corporate governance and corporate law presents efficiency as an implied value, I have attempted to show that we must reject such institutionalized practices as failing on their own terms. In this sense, I offer an immanent critique of our current corporate structures. Secondly, insofar as the corporation is a crucial institution in modern liberal democratic societies, showing such an institution to be out of step with the core liberal democratic values can be understood as an immanent critique of liberal democratic and market societies more generally. We are forced to reckon with the fact

---

that much of our economic productive relationships are arranged in ways that are not in accord what our values should dictate.

I therefore do not contend to be some kind of critical philosopher-king looking at the sun to see the true forms of the corporation, as if such a thing were possible. Instead, following Walzer, I strive to be a theorist standing in the cave, deriving the critique from within.\textsuperscript{2} Aside from being more theoretically coherent, this strategy has the further, and perhaps more important, benefit of articulating critique in a way that can be comprehended by those engaged in the practices themselves. As Sabia puts it: “normative truths, however discovered, must be both communicable and persuasive to have such effect, and so the carriers of extracultural or transcendental insights must express their criticism and prescriptions in way that can be understood and justified…only in this way can such normative truths gain substantive meaning and bite and, therefore, possible authority and efficacy.”\textsuperscript{3} When Socrates’ philosopher returns to the cave, what he sees cannot be effectively expressed to its inhabitants; immanent critique comes equipped with an idiom through which the critique can be communicated. Even if one does come to criticism possessing a transcendent value to be applied to particular practices or norms, immanent critique presents itself as a pragmatic, and perhaps necessary, tool for carrying this out.

The consequence of such a methodological commitment, it must be conceded, is that one’s prescriptive conclusions will be constrained both in content and in kind. They will be constrained in content because, ultimately, they must appeal to values implicit to the society

\textsuperscript{2} Such a methodological misstep is not limited to the idealistic left. If we start with the equally heroic assumption that the market can efficiently allocate everything then we’ve also obviated the need for the corporation or a theory of it. In fact, this is more-or-less what the Chicago theory of the firm does.

to which they are addressed. This inherently imposes certain conceptual horizons. For example, I started with norms of liberal democratic capitalism because they represent the normative basis of the society in which we find ourselves. This has direct consequences for the kinds of conclusions I can draw. The liberal privileging of individual choice, and capitalism’s emphasis on efficiency, are key constraints in the kinds of values that can be foisted onto the corporation. Were one to approach the question of corporate governance solely from a commitment to eudemonism, communitarian solidarity, or some other such value, the conclusions would be different. One would be less concerned, say, with creating institutions like a cooperative stock market or insurance products, and would perhaps simply try to find ways to outlaw more atomizing and isolating forms of corporate organization.

As someone who personally does share values like these listed above, I admit that such conclusions are attractive. A commitment to immanent critique, however, forces me to exercise more self-restraint in the degree to which I appeal to my own moral beliefs. This is because both liberal values and economic efficiency—which I take be immanent to our social institutions—impose all sorts of constraints on how we can use law and public policy. Thus, though a eudemonist might believe strongly in the importance of worker cooperatives in society, a eudemonist who is trying to be sensitive to the values implicit to his society may not be able to be so full-throated in his beliefs. Mill is exemplary of this. Despite his

---

4 My own personal views are much closer to eudemonism and communitarianism. I personally believe that a way of life that is overly individualist is missing something crucial about what it means to be human, and I find utterly negative understandings of freedom to be vacuous. I think Hegel is right that though “the bond of duty can appear as a restriction…the truth is, however, that in duty the individual finds his liberation.” To be free will always entail certain kinds of restraints derived from our obligations and commitments to others. I therefore favor organizations that take such bonds seriously, that cultivate such dispositions in their members. My interest in the corporation as a subject for normative theorizing and a site where salutary social norms may be implemented, is perhaps explained by these normative commitments. Because I don’t think these beliefs are intuitive, obvious, or widely shared, I have avoided appealing to them in making my arguments.
libertarian and utilitarian streaks, Mill was also a deeply eudemonistic thinker, concerned with how to cultivate flourishing and individuality. As we saw in chapters 2 and 4, Mill believed that an economy dominated by worker cooperatives was both possible economically, and preferable morally and politically. However, he did not think such a mode of organization could be forced onto its participants. One reason was certainly out of a respect for individual liberty. As Mill famously contended, even if an individual is unenlightened with regards to his own self-interest, society should not force that enlightenment upon him: “considerations to aid his judgment, exhortations to strengthen his will may be offered to him, even obtruded on him, by others; but he himself is the final judge. All errors which he is likely to commit against advice and warning, are far outweighed by the evil of allowing others to constrain him to what they deem his good.”

Thus even if worker cooperatives would clearly be good for workers, if workers do not choose to enter them, we have no basis to force them to be chosen.

The other reason why Mill did not think cooperatives should be compulsory, as we saw, was efficiency. Mill believed market competition was the ultimate judge of what was efficient with regards to productive organization, and therefore was extremely important for understanding how production ought to be organized. This was grounded in his utilitarian view that what fundamentally mattered was human welfare; therefore, to legally impose certain organizations onto the market, without concern for efficiency, was to reduce overall social utility. In our post-Rawlsian and post-Pigouvian world, such assumptions are less obvious: utilitarian moral philosophy is now more suspect, and we have a better understanding of when and how competitive markets reduce human welfare. Still, the overall

point is not so easily dismissed: if we care about justice, we must care about human welfare, and if we care about human welfare, we must care about efficiency. What Mill was really getting at was what Janos Kornai would articulate over a century later, that in order for an economy to benefit society, businesses must have relatively hard budget constraints.⁶

Businesses will tend to favor socially optimal allocations if they are forced to respond to price signals, alter their production accordingly, and not offload their costs onto others. Forcing a particular kind of corporate organization—even one that is more just than some other alternative—might have very large negative consequences for human welfare, by messing with enterprises’ ability to match their production to price signals.

Of course neither the respect for individual liberty nor concerns for economic efficiency are insurmountable; indeed, both efficiency and liberty are often at odds with one another and sometimes require one to trump the other. Still, liberal values and economic efficiency end up constraining the content of our prescriptions because they force us to be temperate with regards to the state’s ability to effect social change in a positive manner.

Now, as I have tried to show, this hardly leads us to a laissez-faire or utilitarian view. There is still ample room for fairly radical prescriptive conclusions starting from such presuppositions. But it does require some creativity, and less reliance on straightforward legal fiat, to achieve.

A commitment to immanent critique will also mean that normative prescription will be constrained in kind. By this I mean that prescriptions will necessarily be of a slightly more realist bent than some other methods of critique might allow. G.A. Cohen, for example, famously contended that one could articulate normative principles independent of empirical

---

consideration about feasibility. From the perspective of immanent critique such a position seems impossible. Because we start with an empirical account of the phenomena we are studying, and then subject those findings to critical scrutiny based on the values they imply, one cannot then turn around and make normative recommendations without similar attention to real world concerns. Though we want to critique practices as we find them, we are not entitled to assume the messy and nitpicky factors away; the real world is both the subject of critique and the context in which our prescriptions must be made.

As we have already seen, being sober about the constraints that the market imposes on the corporation has the effect of constraining the kinds of reforms we can make. Had this study been more idealist, I might have started with the assumption that we don’t need a market to allocate things efficiently. With this assumption, I would not have needed to think about the constraints that corporations face through market competition. In fact, I would not have to worry about transaction costs at all since I could just assume those away as well! In that case then, yes, we could imagine all sorts of ways in which the corporation could be organized given soft budget constraints. But the blank canvas on which we paint our bright and idealistic vision would have been purchased by essentially assuming the problem away. Absent a developed competitive market we would likely not have the modern business corporation in the first place, and so we would not need to worry about its moral legitimacy.\(^7\) While we don’t want theory to be a mere rationalization of the status quo, we don’t want it to be idle speculation either. In order for a theory to be critical, it must be critical of something, and therefore be relatively close to the real world of practice to know what that something is.

\(^7\) Such a methodological misstep is not limited to the idealistic left. If we start with the equally heroic assumption that the market can efficiently allocate everything then we’ve also obviated the need for the corporation or a theory of it. In fact, this is more-or-less what the Chicago theory of the firm does.
If we want to criticize contemporary corporate practice we must first start with the context in which corporations operate, which is the competitive market. The consequence of this, however, is that we cannot then ignore the way the competitive market operates when articulating a different vision of how the corporation should be structured.

Of course, I could have subjected different things to criticism and started with different empirical assumptions. If instead of criticizing the Chicago School’s theory of the corporation I had subjected the whole complex of institutions we associate with capitalism—the market, private property, the corporation, etc.—to critique, this would have implied many different changes to this study. As an example, I would not have started with the existence of a competitive market, since whether or not the competitive market is justifiable would have been the question to be answered in the first place. Instead, I might have started with certain assumptions about mass society, the division of labor, and so on, and asked whether capitalist institutions live up to the normative bases implied by those factors. This would have allowed me a greater freedom in giving a prescriptive account of the corporation, since I would not have started with the context of capitalism; indeed, we might have found that in the transcendence of capitalism, the corporate institution becomes less relevant and necessary. Of course, such conclusions would not have been unconstrained either; they still would have to address themselves to the problems that capitalism addresses itself, like the division of labor or the problem of scarcity.

All of this is to say that the theory of the corporation that I have offered as a counter-proposal to the Chicago School’s theory is very much the result of how I conducted my critique. This is, of course, always true for any theorist; the types of constraints and presuppositions we impose on ourselves will have the effect of pushing our conclusions in
certain directions. Because I started with the Chicago account first, my own counter-proposal is more concerned with efficiency than it might have been had I, for example, began with a theoretical account of egalitarian socialism that I then applied to the corporation. Because I have not challenged the capitalist or liberal context in which the corporation operates, my conclusions are more concerned about efficiency and individual liberty than they might have been had I subject those values to criticism as well. However, taking such things as given also has the benefit of making my conclusions more germane to the social situation we find ourselves in here and now. Had my theory of the corporation started with the assumptions of egalitarian socialism or the non-existence of the capitalist marketplace, one could not try to change the corporation without first changing the market, our set of social values, and so forth. As a result, though my conclusions might be more biased toward conventionalism than they might have been, they are also more ready to be acted upon than they would have been had I adopted less conventional assumptions.
Bibliography


**Thought.** London: Chapman and Hall.


Williamson, Oliver. 1988. “Corporate Finance and Corporate Governance.” *Journal of*


