A NORMATIVE MODEL OF PROFESSIONALIZATION:
IMPLICATIONS FOR BUSINESS ETHICS

by

Sareh Pouryousefi

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

© Copyright by Sareh Pouryousefi 2013
A Normative Model of Professionalization:
Implications for Business Ethics

Sareh Pouryousefi
Doctor of Philosophy
Department of Philosophy
University of Toronto
2013

Abstract

To say that there is a lack of consensus on the essence of professionalization is an understatement. For almost a hundred years, scholars have disagreed about the traits, attributes, processes, temporal sequences, and socio-historical structures that fundamentally define professions. My dissertation draws on an information theoretic framework to provide a novel analysis of the role of professions in society. According to my normative model, professions are trust-creating and trust-preserving institutional structures, which respond to market failures that arise due to information asymmetries in the market for professional services. Given the limitations on market-based and governmental solutions to information asymmetries, I argue that norms should be viewed as a fundamental transaction-cost minimizing professional governance mechanism.

What are the moral obligations of the professional? The implicit morality of the market for professional services involves achieving the end of economic efficiency. Accordingly, professional moral obligations involve a set of deontic constraints that promote Pareto-efficiency in response to information asymmetries. After providing a conceptual framework that outlines
the relation between professional and ordinary morality, I argue that social institutional roles sometimes permit, to a limited extent, what would otherwise be forbidden. Given certain institutional structures and safeguards, professional roles can, thus, be sui-generis sources of moral obligation.

The culmination of my arguments about professions and professional morality is a contribution to the business ethics literature. Since the early twentieth century, scholars have argued over whether the managerial role may be theorized using the normative and theoretical trappings of the professions. A survey of the management literature reveals that an adequately nuanced analysis of the professions and the normative nature of professionalization has yet to be put forward. To fill this gap, I provide a normative model of management professionalization that is sensitive to the socially beneficial nature of professional work while avoiding glorified altruistic characterizations. I argue that managers in private economic entities are professionals, properly understood, since in addition to external market-oriented incentives, they typically appeal to internal, trust-creating norms to promote Pareto-efficiency within the firm. Professionalism thus involves following efficiency imperatives despite the prevalent moral hazard problems surrounding the managerial role.
Acknowledgments

First and foremost, I am grateful to my supervisor, Joseph Heath, for his excellent mentorship and advocacy. He has been a tireless supporter of my research, reading countless chapter drafts, and providing comprehensive and insightful comments. I could always count on Joe to send me lightning-quick email replies, to be generous with his time and frank about his opinions, and to prioritize my best interests. I am indebted to Joe for his relentless commitment to my success, and for challenging me to think through the philosophical depth of the problems I sought to address in my dissertation. He has been a model teaching mentor, and attentive to the institutional dynamics surrounding me as a female, non-native English-speaking philosopher. I learned from Joe to muster “the will to one thing,” to take the social sciences seriously, and when in doubt, to “say something interesting”.

I also owe thanks to others at the Department of Philosophy at the University of Toronto – to Margaret Opoku-Pare, Mary Frances Ellison, and Suzanne Puckering for their amazing administrative support; to Donald Ainslie, Amy Mullin, and Arthur Ripstein (during each of their terms as chair) for ensuring that I had a first-rate graduate experience, and for providing me with support when I asked for it; to Frank Cunningham for his guidance through my coursework year; to Wayne Sumner for setting an example for me as a professional philosopher, and for serving as my internal examiner; to David Dyzenhaus for holding my hand as I gained my research footing, and for being a steady source of inspiration; and to Arthur Ripstein, who alongside David, served as my reader. Arthur and David provided excellent advice and unfailingly prompt feedback on my drafts, for which I am very grateful. Their penetrating criticisms have made this a better thesis. Along with Joseph Heath as my supervisor, I could not have had a stronger dissertation committee.

I am indebted to my external examiner, Wayne Norman, for his kindness and his advice, and for the thoughtful comments in his detailed appraisal of this dissertation.

Thanks also to Rose Jones and Joseph Heath, and to the post-docs and graduate students at the University of Toronto’s Centre for Ethics, where I was a Graduate Associate. The bright and beautiful Trinity College surroundings, the interesting talks and colloquia, and the supportive reading groups and workshops all played an important role in my education.
In addition to my formal committee members, I have been fortunate to meet faculty members at other institutions who have taken an active interest in my work. I am indebted to Jeff Frooman (University of New Brunswick) for his mentorship and support, and for introducing me to the intellectual community at the Society for Business Ethics; to Dirk Matten and Andy Crane (York University) for their teaching mentorship and their warm support during my visit at the Schulich Centre of Excellence in Responsible Business; to Wes Cragg and the staff at the Canadian Business Ethics Research Network for generous research and professional development support throughout my PhD; to Thomas Beschorner (University of St. Gallen) and the Transatlantic Doctoral Academy for fruitful research advice and stimulating international workshops; and to Jeffery Smith (University of Redlands) for his insightful comments on my research.

I owe special thanks to Parviz Shahriari – the shahriar (king) of mathematics in Iran – for being my first and most influential mentor; to Susan Gardner (Capilano University) for converting me to philosophy; to my undergraduate advisor, Ray Jennings (Simon Fraser University) who got me excited about logic, and who had the foresight to prepare me for an academic career before I knew I wanted to pursue one; to Mark Winston (Simon Fraser University) and the Semester in Dialogue for ingraining in me the importance of thinking critically and independently; and to my instructors at the University of Toronto Judo Club, especially Sensei Jorge Comrie, for teaching me the art of maximum-efficiency-minimum-effort – being expertly thrown, arm-barred, and choked allowed me to forget about the dissertation in between bowing in and out of the dojo.

The support of my friends was instrumental throughout the dissertation writing process. Thanks to Florence Ballard, Lana Kuhle, Ian Drummond, Susanne Mueller, and Kristine Smitka. I’m also grateful for inspiring conversations with my talented Philosophy Department colleagues Stephen Thierman, Eran Tal, Doug MacKay, Farshid Baghai, Lauren Bialystok, and Patrick Turmel. Thanks to my parents Sholeh Nouri and Hamid Pouryousefi – for staying strong in the old country, revolution and war and all, during the first 17 years of my life, and for their love and support as I strive to be everything I had not known I could be: independent, unburdened, and unafraid.

Above all, my greatest debt is to my partner and fellow harvester, Stephen Fraser Thierman. Steve has been the source of unflagging emotional support, inspiration, happiness,
perfect espressos, excellent cocktails, and gourmet meals. Alongside the unique and purr-plenty affections of our sweet cat, Bubba Walnut, Steve’s love has sustained me through the course of this dissertation.
For Thierman & Bubba
Table of Contents

Acknowledgments.......................................................................................................................... ii

Table of Contents.......................................................................................................................... viii

List of Figures................................................................................................................................ x

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

1  The Essence of Professionalization: A Historical & Theoretical Survey ...................................... 25
   1.1  Etymology and History, Prior to the 20th Century ..................................................................... 26
      1.1.1 Connection with the Guilds ................................................................................................. 31
      1.1.2 The History of the Anglo-American Professions ................................................................. 34
      1.1.3 Professional Regulation & the State .................................................................................... 40
   1.2  Definitional Uncertainty .......................................................................................................... 44
      1.2.1 Functionalism .................................................................................................................... 48
      1.2.2 Structuralism .................................................................................................................... 61
   1.3  Morality and the Professions .................................................................................................... 68
      1.3.1 Professions as Inherently Moral Entities ............................................................................. 68
      1.3.2 Professions as a Means of Promoting Self-interest ........................................................... 71
      1.3.3 Conclusion ....................................................................................................................... 74

2  Functionalist Professional Morality ............................................................................................... 77
   2.1  Durkheim ............................................................................................................................... 81
   2.2  Tawney .................................................................................................................................... 85
   2.3  Parsons .................................................................................................................................... 90
   2.4  Some Criticisms ..................................................................................................................... 94

3  Information Asymmetries, Regulation, and the Professions ......................................................... 97
   3.1  Professional Regulation .......................................................................................................... 99
      3.1.1 The Public Interest Theory of Professional Regulation ................................................... 103
      3.1.2 The Private Interest Theory of Professional Regulation .................................................. 109
   3.2  Information Asymmetries: A Backgrounder .......................................................................... 125
      3.2.1 Efficiency ......................................................................................................................... 129
      3.2.2 Principal-Agent Relationships & Games of Asymmetric Information ............................ 132
   3.3  Modeling Information Asymmetries: The General Case ....................................................... 152
4 The Mystery of Professions

4.1 Professions as a Private Economic Entity

4.2 Professions as Trust-preserving Institutional Mechanisms

4.2.1 Professionalization as a Response to Moral hazard & Adverse Selection

4.2.2 The Definition of Professions

5 Professional Morality via Roles

5.1 Professional Morality, Efficiency, & Ordinary Morality

5.2 Roles & Role Morality

5.2.1 Partiality

5.2.2 Division of Moral Labor

5.2.3 Adversarialism

5.3 A Conceptual Solution

6 Management Professionalism

6.1 Is Management a Profession? A Literature Review

6.1.1 Professional Traits

6.1.2 Lack of Critical Attention

6.2 Trust and Management Professionalism

6.2.1 Managers as Professionals

6.2.2 The Separation of Ownership & Control

6.3 Managerial Role Morality

6.4 Implications for Business Ethics

6.4.1 Conclusion

Bibliography
List of Figures

| Figure 3.1 | The Principal-Agent Game | 133 |
| Figure 3.2 | Five Games of Asymmetric Information | 155 |
| Figure 3.3 | The Nash Equilibrium Solution for Strategic Games – An Example | 158 |
| Figure 3.4 | A Game of Incomplete Information – Follow-the-Leader III | 159 |
| Figure 3.5 | After the Harsanyi Transformation – Follow-the-Leader III | 161 |
| Figure 4.1 | The Professional-Client Game – Moral Hazard Hidden Information | 181 |
| Figure 4.2 | The Professional-Client Game – Signaling | 182 |
| Figure 5.1 | The Spectrum of Morality: Self-regarding vs. Altruistic Ends | 200 |
| Figure 5.2 | A Three-part Categorization for Professional Morality: Permissible/Forbidden/Obligatory | 203 |
| Figure 5.3 | The Division of Moral Labor: Personal vs. Impartial Realms | 227 |
| Figure 5.4 | The Division of Moral Labor: Roles & Institutional Morality | 233 |
| Figure 6.1 | Professionals & Managers in the Private Economic Market | 278 |
Introduction

To say that there is a lack of consensus on the essence of professionalization is an understatement. Debate about the definition of professions goes back almost a hundred years, and scholars continue to disagree about the traits, attributes, processes, temporal sequences, and socio-historical structures that fundamentally define professions. For the most part, it appears that scholars have accepted definitional uncertainty and moved on. My dissertation considers the definitional question anew, and draws on a game theoretic framework to provide a novel analysis of the role of professions in society. Based on my normative model of professionalization, I argue that managers in private economic entities are professionals, properly understood, and I provide a dynamic analysis of the managerial role and its implications for business ethics.

I present professions as an institutional mechanism for responding to market failures that arise due to information asymmetries. Because of information asymmetries in, for example, a patient-physician relationship, the patient may struggle to evaluate the physician’s training and experience, as well as the physician’s incentives to recommend and provide accurate diagnosis and treatment. The complex nature of medical service thus tends to prevent the patient from adequately evaluating due diligence and the quality of the care she receives. Patients’ distrust of professional medical services, as well as prospective doctors’ misgivings about entering the medical profession, can both lead to inefficiencies arising from lost transactions. For their part, professional associations typically respond to information asymmetries through institutional mechanisms, such as educational certification, experiential pre-requisites, codes of ethics, peer assessment, and complaint investigation. Given the limits to market solutions that counter
information problems in the market for professional services, trust-creating institutional norms are also core features of professional governance mechanisms.

Amidst the various attempts at definition, two schools of thought on the normative nature of professionalization have emerged: one focused on altruism and one focused on monopolization. On the one hand, a large body of literature has praised professionals for their dedication to the public good, their service ethics, and their lack of interest in personal gain. Critics of professions, on the other hand, tend to view the ethos of service orientation as a rhetorical strategy that legitimizes the power, and the social and economic prestige, that accompanies professionalization. My approach, which presents professions as mechanisms that enable cooperation by responding to market failures arising from information asymmetries, dissolves this dichotomy, since the mutual benefit created through cooperation is the half way point between monopoly and altruism. Professionals can thus claim to be good for society when they are guided by appropriate professional norms, but not because they are altruistic saints. At the same time, the behavior of so-called professionals who draw on their institutional autonomy and authority to enable self-serving cartelization can be viewed as a distortion, since that behavior deviates from the socially beneficial purpose of the professions.

Another central question of my dissertation is: what are the moral obligations of the professional? Our everyday moral sensibilities are fundamentally intertwined with our professional identities, and our moral obligations acquire meaningful texture when studied through the lens of professional roles. Whether we are teachers, lawyers, or physicians, our professional roles impose obligations that, at least to some extent, have normative authority. I argue that professional morality involves a set of deontic constraints that promote efficiency in response to information asymmetries in the market for professional services. I then propose a
conceptual framework that demonstrates that social institutional roles can sometimes permit, to a limited extent, what would otherwise be forbidden. Partiality, the moral division of labor and adversarialism are three increasingly controversial variations of this kind of permission. Providing a moral justification for such permissions is one of the most challenging problems in the study of professional ethics. The specific institutional circumstances within which roles take place play a crucial role in determining the normative nature of professional obligations.

The culmination of the arguments in the first five chapters of the dissertation is a contribution to the literature in business ethics. Even though professions are private economic entities, they have historically been portrayed by some as ethically motivated, and there is a long-standing literature that documents professional moral norms and codes. Like professions, corporations are private economic entities; but there is a fair bit of skepticism about the moral climate of business, and some scholarly and popular accounts continue to view the notion of “business ethics” as a contradiction in terms. As Emile Durkheim asked in Professional Ethics and Civic Morals, “[t]here are professional ethics for the priest, the lawyer, the magistrate.... Why should there not be one for trade and industry?” But business was traditionally not considered a profession, and the scholarship remains inconclusive over whether the managerial role should properly be understood as a professional role.

I argue that managers are professionals, since they appeal to trust-creating, trust-preserving institutional mechanisms to make Pareto-improvements by responding to information asymmetries in the corporation. These institutional mechanisms draw on external market-oriented incentives as well as cooperative norms. Within the corporation, management professionalism involves following a set of deontic constraints, in the form of efficiency imperatives that guide managers to fulfill obligations towards the corporate chain of command,
despite the prevalent information asymmetries. A failure to fulfill efficiency-promoting
imperatives can thus be viewed as a distortion of managerial professionalism. One noteworthy
feature of my argument – which focuses on professional responsibility within the firm – is that
managers owe moral obligations towards employees (a group that is typically overlooked in the
traditional shareholder-primacy theory of the firm). There are also crucial theoretical and
practical implications that follow from establishing a standard of managerial professionalism
towards the rest of the market and the broader social structure, for which my work serves as a
necessary grounding theoretical step.

Having provided a broad overview of the dissertation, I now discuss some of the important
theoretical assumptions and underlying economic considerations relevant to my project. This
discussion is then followed by a précis of the six chapters in the dissertation.

*A Normative Model of Professionalization*

My task during the first five chapters of the dissertation is to outline a normative model of
professionalization. I then apply this model to the field of management in chapter six. A
“reconstructive” normative model articulates the implicit norms and ideals that structure our
practices.¹ But in order for such a model to be credible, it cannot be merely prescriptive. We thus
need an empirically-informed understanding of what professions do, demonstrated by the kind of
work they have historically done, in order to lay the groundwork for a more sophisticated
approach. In effect, such a normative model of professions provides an account of the normative

¹ See Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*,
purposes already *implicit* in the practices of professionals. Accordingly, I dedicate chapters one and three to discussing the history of professions and the economic models that explain their essence.

A core assumption of my approach is that certain normative professional purposes have efficacy. Specifically, I hold that the nature of professions is, at least in part, a consequence of the mutually beneficial moral purpose of professional market interactions. The relation between the nature of professions and their purpose might be direct. For example, clients might solicit professional legal services, and lawyers might join the profession, because the regional bar associations purport to serve moral purposes. But it is also possible for the relation in question to be attenuated. Even if we consult legal services for reasons independent from the moral purpose of the legal profession, the profession arguably has been developed with certain underlying normative purposes at its core, and those design features resonate to some extent through current legal institutions.

A central problem is that two extreme positions dominate the theoretical reconstruction of the normative commitments underlying the professions. The two purposes that are often cited as justifying professional obligations are: self-interest, on the one hand, and altruism, on the other hand. Each account corresponds to a normative model. In the first account, the central function of professions is to promote professional self-interest, with the goal of gaining cartelizing market control. In the second account, the central function of professions is to provide altruistic professional service to the public, with the unselfish goal of improving the wellbeing of professional clients. In this dissertation, I introduce a third normative purpose called “mutual

---

benefit.” Here, the central function of professions is conceptualized as the creation of mutual benefit from cooperation, through making Pareto-improvements in the market for professional services. This model views professional institutions as playing a complementary role to that of the market, since professional mechanisms correct market failures by appropriately responding to information asymmetries.

One virtue of this proposed model is that it avoids the long-standing dichotomy between the first two approaches. As I show in my first chapter, for over a hundred years, definitional uncertainty has surrounded the professions. A pernicious consequence of this lack of consensus is normative confusion, which can in turn lead to a weakened defense of the value of professions in the social sphere. From a public policy perspective, such normative confusion stalls reforms and improvements to the standards of professional services. The definitional uncertainty and normative confusion surrounding the professions makes it difficult to justify the status of one set of professions, while legitimately objecting to other occupations deemed unworthy of the title of “profession.” Another problem with the uncertainty surrounding the professions is that the prevailing models have difficulty explaining the tremendous growth of professionalization. Scholars disagree over whether the historical predecessors of professions are the guilds. What they do agree on is that professions began to grow in number in the nineteenth century, and, as influential works such as Wilensky’s “The Professionalization of Everyone” have documented, became a widespread phenomenon in the twentieth century.

The cartelization model would explain the history of professions by suggesting a pessimistic trend, where new professions increasingly grew by taking advantage of the state’s authority. But beyond the challenge from motivational skepticism, this model struggles to explain the ethos of service orientation in self-regulating professions. The sustained emphasis on
codes of ethics and norms of peer-assessment thus raises doubts about the characterization of self-serving gains as the core professional interest. Meanwhile, the altruistic model would explain the growth of professions by making the somewhat optimistic claim that an increasing number of occupations have sought to serve their clients better. This model struggles to explain the inflated remuneration of professionals and traditional barriers to entry (e.g. relocation, experience, and educational requirements).

In comparison, my proposed model of professions offers a more plausible explanation of the exponential growth of professions and their unique institutional design. As I argue, information asymmetries underlie all institutional features of professions. Educational pre-requisites, licensing and standard-setting, codes of ethics, peer-assessment, and other professional institutional mechanisms, can be explained in terms of standard information theoretic responses to market failures that arise due to information asymmetries. Information asymmetries can also provide a more fundamental and comprehensive explanation for the growth of professions, given the professional capacity to provide non-market governance solutions when markets fail. As Kenneth Arrow puts it, “… when the market fails to achieve an optimal state, society will, to some extent at least, recognize the gap and nonmarket institutions will arise attempting to bridge it.”

As markets grew in the course of the twentieth century, professional market mechanisms can be understood as having grown alongside them, supplementing market mechanisms (e.g. competition and reputation-building) with non-market institutional mechanisms (e.g. norms and codes of ethics). Professional codes of ethics and professional

---

associations are some of the oldest and most sophisticated models of non-market governance in the history of theorizing about private economic entities.

Rather than trying to provide a purely prescriptive definition for the professions and their normative features, I seek to identify the normative considerations that are embedded in existing professional institutions. This model provides an improved self-understanding of the nature of professional work for professionals themselves, so that they may more effectively achieve the goals implicit in their practice, and provide a legitimate case against distortions of professionalization.

In the next section, I discuss some of the underlying economic assumptions relevant to my normative model.

**Economic Background**

Economists often theorize about private economic entities (such as corporations and professions) by comparing their characteristics with markets. The reason for this tendency, as many economic historians have pointed out, is that markets have played a pivotal role in western economic systems since the Middle Ages. Until recently, and despite the setbacks of the current financial crisis, markets have continued to impact growth and industrialization around the world. The growing phenomenon of large modern corporations may be evidence for the fact that certain transactions are most effectively carried out within the firm, as opposed to the market. In this context, the question has been whether hierarchical systems, such as large firms, can operate as

---

efficiently as the unorganized or decentralized market.\textsuperscript{6} The literature that attempts to answer this question is broadly called the “new institutional economics.”\textsuperscript{7} My study of professions is conducted within this realm, since I define professions as institutional structures that arise when professional service associations can carry out certain transactions more efficiently in comparison to the market.

Ronald Coase was the first economist to inquire about the activities that are most efficiently carried out within economic entities like firms, as opposed to markets.\textsuperscript{8} Coase laid the foundations of the transaction cost approach to governance. A brief outline of this approach will prove insightful.

All transactions purport to serve the mutual benefit of the economic agents involved. Governance mechanisms and institutional constraints are meant to ensure that agents are protected from free-riders who might attempt to cheat the reciprocal nature of economic interaction. The cost of imposing such a governance system is called a “transaction cost”. Coase pointed out that because of transaction costs, some activities are most efficiently governed within the hierarchy of firms, while others pose the lowest transaction costs when carried out in the market. What was needed was an assessment of the net benefits and net transaction costs under different organizational and governance alternatives. Coase identified negotiating agreements and determining appropriate prices as key costs of transactions that are mediated by the market.\textsuperscript{9}

\textsuperscript{9} Ibid.
Oliver Williamson later built on Coase’s transaction cost theory and explained variations in governance structures through underlying differences in the attributes of different transactions.\textsuperscript{10} Crucially, Williamson highlighted the impact of the fear of opportunism on the choices of market actors. He pointed out that individuals might decide to organize activities within corporations or within the market, depending on their expectations about the opportunistic behavior of economic actors. For example, market power (i.e. the ability of an individual economic actor to influence the trading price of a given good) is a source of transaction costs, because of the risk of opportunistic behavior that arises from having insufficiently competitive suppliers in the market. Thus, in some ways, the transaction cost theory of the firm can be thought of as a “market failure” theory of the firm,\textsuperscript{11} since in cases of externalities, information asymmetries, and market power, the market in effect fails to operate according to the traditional invisible hand theorem.

Under such circumstances, transaction cost theory would suggest that it may be more efficient to conduct transactions within firms, using non-market models, such as hierarchies.

But as Herbert Simon\textsuperscript{12} and generations of critics of new institutional theory\textsuperscript{13} have pointed out, we need to attend to the unique non-market organizational mechanisms that go beyond


\textsuperscript{12} Simon, “Organizations and Markets,” 47.

\textsuperscript{13} Critics have argued that transaction cost economics ignores differential capabilities in structuring economic organizations, neglects power relations, identity and trust within corporations, and overlooks evolutionary considerations regarding market processes. Recent criticisms of transaction cost economics have been raised by sociologists, non-mainstream economists, and management scholars. It is important to note, however, that most of the basic criticisms of transaction cost economics apply generally to game-theoretic microeconomic analysis of the firm. See George B. Richardson, “The Organisation of Industry,” Economic Journal 82 (1972): 883-96; Charles Perrow, Complex Organizations: A Critical Essay, 3rd ed. (New York: McGraw-Hill, 1986); Bruce Kogut and Udo Zander, “What Firms Do? Coordination, Identity and Learning,” Organization Science 7 (1996): 502-518; Mark S. Granovetter, “Economic Action and Social Structure: The Problem of Embeddedness,” American Journal of Sociology 91 (1985): 481-510; Richard N. Langlois and Paul L. Robertson, Firms, Markets, and Economic Change:
agency theory, asymmetric information, transaction costs, opportunism, and other concepts in neoclassical economics. One such key non-market solution involves the formation of alternative organizational forms that draw on ethics.

Certain efficiency losses, e.g. those caused by opportunism, need not be solved by hierarchies, but can be partially alleviated through “norms” and “codes of conduct.”\textsuperscript{14} For example, even if suppliers do not have the right incentives to honor an agreement, and buyers are not privy to this fact due to information asymmetries, suppliers may decide to honour agreements in order to establish and promote a reputation for honesty and fairness. Employers whose treatment of employees is shielded from outsiders by asymmetries in information, and who lack the right kinds of incentives to provide satisfactory employment conditions, may similarly be motivated to uphold a reputation as a desirable employer that can in turn attract high-quality future employees.

Even further, and more importantly, institutional norms might motivate efficiency-maximizing strategies for non-instrumental reasons, since individuals draw reason-giving force from the institutional norms guiding their work. For example, professionals such as doctors, lawyers, and civil servants work within institutional settings infused with service-oriented

purpose. Organizational culture (e.g. loyalty and promise-keeping) and broader cultural norms (e.g. in family-owned vs. multi-national corporations) may also play an important role, as does individual identity-building and the sense of belonging and comradeship gained through shared work experiences and interactions.

Within the transaction cost economic framework, my study of professions makes the observation that, when market solutions have been exhausted or when markets fail, certain transactions are more efficient to conduct within professional associations and through using non-market solutions. Professional institutions provide normative guidance and a culture of professional purpose and service orientation. To be sure, market solutions, including improved information communication and incentive setting, play a crucial role in the professional organizational form, and I would also underscore the importance of the reputational effects of professional behavior. However, I want to draw attention to the fact that when market solutions and reputational mechanisms have been exhausted, or in cases where the background assumptions for such solutions do not hold, professional organizational forms can still provide services at lower relative transaction costs. Thus, my approach is appropriately subversive of neo-classical economics by attending to the impact of and institutional implications of market failures. At the same time, by giving ethics a key governance role, my work pushes the boundaries of the traditional perspective offered by new institutional theory.

---


In his discussion of the nature of the firm, Simon suggests that we visualize our economy as seen by an alien visitor who watches the earth through a special telescope that detects social structures. Here, firms show up on the telescope as “solid green areas with faint interior contours marking out divisions and departments.” Market transactions, on the other hand, show up as “red lines connecting firms, forming a network in the spaces between them.” There are also a set of blue lines within firms, representing hierarchy, and connecting employees with their employers. As Simon points out, in the eyes of the alien visitor:

… The greater part of the space below would be within the green areas, for almost all inhabitants would be employees, and hence inside the firm boundaries. Organizations would be the dominant feature of the landscape. A message sent back home, describing the scene, would speak of “large green areas interconnected by red lines.” It would not likely speak of “a network of red lines connecting green spots”.

Continuing with Simon’s thought experiment, I hold that beyond the red, green, and blue lines, we need to start paying special attention to the set of yellow lines that represent ethical norms. Norms play a crucial role in our daily organizational lives. On the telescope, yellow lines would show up parallel to all visible blue lines and red lines. Take the norms guiding promise-making, for example. These norms fundamentally affect our market and authority relationships. The more people can be trusted to do what they promise to do, the lower the transaction costs of enforcing contracts. The more employees can be trusted to do what is asked of them, the less the agency costs of employee shirking. Consider the norm of loyalty as another example. Green areas on the map that represent armies (of nation-states, guerrilla warfare, or “terrorists”) have prominent blue lines of hierarchy running through them. But at least as prominent as the blue lines are the yellow lines of loyalty that have the power of convincing soldiers to give up their lives to fulfill their duties. Or, consider also the norm of service-orientation (in government

---

18 Ibid.
bureaucracies like post-offices and consular services). The more service providers can be trusted to prioritize client service, the less costly it will be to set up intricate incentive schemes and policing mechanisms in order to improve service delivery and standards.

Yellow lines enforce blue and red line governance mechanisms and explain the force of internal organizational incentives, such as identity and culture. The yellow lines also show up in places where the red and blue lines do not, e.g. in relations between employees within the green area of a firm, representing the sense of identity and solidarity that employees derive through serving in unions and on employee committees. As well, yellow lines connect green areas of various sizes and shapes, more prominently than the sparse blue lines of hierarchy among select organizations. Respect for the rule of law, through industry-wide adherence to the spirit of regulations, for example, draws on honesty, trust, and solidarity in intricate ways that hierarchical and market solutions cannot explain.

While Simon’s telescope showed the same concentration of green areas with red lines connecting them in China, Russia, and the United States, the yellow lines would differ in thickness and color density depending on what problem and what part of the world we are analyzing, since norms vary depending on cultural and social settings. The yellow lines also differ among various green areas. Crucially for this dissertation project, the green areas that draw on the work of professionals typically exhibit especially prominent yellow lines. This is why I propose that we may define professions with special attention to the prominent yellow lines that distinguish them from other private economic entities.

My proposal that managers are professionals, properly understood, in effect amounts to a reconstructive call to observe that a basic set of yellow lines already runs through corporations, just as it does in the professions. Ethics, to some extent, is already an implicit moral feature of
corporations. Professionals have historically analyzed and developed yellow lines, and ranked them among key institutional mechanisms, so that we might imagine a yellow glow hovering about them from afar. Building on the existing blueprint of yellow lines in corporations, I argue that developing ethics and the use of non-market solutions in corporations can improve efficiency within corporations.

Now that I have sketched out the background economic debate and introduced the normative model I seek to establish, I provide a brief outline of the six chapters in my dissertation.

**Précis**

The **first chapter** provides a survey of the historical and theoretical accounts of the nature of professions over the past century. I begin by discussing the etymology and history of professions. After surveying some of the historical confusions surrounding the use of the word “profession,” I provide a broad-brush history of the Anglo-American professions from the Middle Ages up to the present time. Much of the twentieth century literature on the professions centers on the proper definition of profession. I provide a survey of the attempts to provide a definition and reflect on some of the controversies and difficulties encountered in theorizing about the professions. These definitional attempts may be characterized under broadly “functionalist” and “structuralist” banners, which I discuss in turn.

The first theoretical movement in support of professions in the twentieth century was a functionalist one. Functionalist scholars explain social phenomena in terms of their role in society. According to Durkheim, “the function of a social fact … consists of the production of
socially useful effects.”\(^{19}\) The military and the family, for example, have social functions, since their objectives are the defense, nurture, and continuity of society.\(^{20}\) Applying this approach to the discussion of professions, if the heart’s function in the body is to pump the blood, then the profession’s function might be, say, to provide social solidarity. Professions thus exist, according to functionalists like Durkheim, because they serve a particular social need.\(^{21}\)

In contrast to functionalism, the structuralist movement (which gained momentum in the mid-twentieth century) was critical of professions and the political and economic repercussions of professionalization. Scholars most prominently associated with this analysis, such as Larson, Starr, and Abbott, brought attention to the “underlying structural processes,”\(^{22}\) “structural effects,”\(^{23}\) and “structural necessities” of the professions.\(^{24}\) While the functionalists studied the utility of professional expertise in order to explain the professions,\(^{25}\) the structuralists challenged the professionals’ claim to legitimate expertise, and criticized professions for being monopolies concerned with prestige and their members’ artificially inflated remuneration.\(^{26}\) The structuralist literature also questioned the objectivity of professional scientific expertise, which was a cornerstone of the functionalist approach.\(^{27}\)

Disagreements over the normative role of professionals in society lie in the background of these definitional debates. On the one hand, some believe that professionals have a substantial

---

26 Ibid.
27 Ibid.
moral commitment to the welfare of their clients and share “an ethic of service.”  

On the other hand, however, many critics of the professions argue that the ethos of service orientation may have been a rhetorical strategy for professionals to deflect criticism of the power and prestige they acquired through professionalization, and that they are, in fact, “wolves in sheep clothing, monopolists who live by the rule of caveat emptor, but lack the integrity to admit it.” 

In sum, my first chapter shows (i) that scholars have yet to reach a consensus on the traits, historical processes, or essential features that define professions, and (ii) that critics have highlighted a number of problems with the process and normative features of professionalization.

In the second chapter, I argue that many scholars of professions (both proponents and critics) have failed to make an important distinction between the demands of functional professional morality, on the one hand, and the demands of all-things-considered general morality, on the other. Seeing this distinction clearly is important because of the morally relevant spectrum that exists between these two models. Earlier in the introduction, I noted the dichotomy between two dominant normative understandings of the purpose of professions: altruism vs. cartelization. While some of the early proponents of professions take functionalism to imply that professionals are by definition altruistic, others, such as the structuralist critics of professions, take functionalism to imply that professionals are by definition egoistic and self-serving. In chapter two, I narrow down my focus on functional professional morality, and present a dual

---

understanding of the professional normative purpose: functional institutional rule-following, on the one hand, vs. general, all-things-considered morality, on the other hand.

As I show, according to functionalists, professional morality requires individuals to promote the social purpose of the profession through a set of institutional guidelines. Determining what motivates professionals to follow these institutional rules is beyond the scope of the functionalist approach, as are broader considerations about general morality. This social purpose is much more limited than what many commentators have understood functionalists to imply, viz. an all-things-considered morality that implies professional rule-following, in addition to a variety of other stringent moral requirements, applicable at personal as well as societal levels. I draw on the works of three early functionalist theorists of professions (Durkheim, Tawney, and Parsons) in order to emphasize the difference between traditional functional professional morality, on the one hand, and all-things-considered morality, on the other.

Having analyzed the history of professions and some of the controversies surrounding the different normative models of professionalization, in the third chapter, I explore the essence of professionalization from an information economic perspective. The debate over the implicit morality of professions continues in this chapter, with a focus on professional regulations (e.g. registration, certification, and licensure). While the public interest theory of professions holds that regulations promote and facilitate professional contributions to the public interest in the face of market failures, the private interest theory of professions holds that regulations are in fact a decoy for enhancing the benefits accrued by professionals through cartelization frameworks. Both sides of the debate present empirical evidence to support their positions, and I review this literature in an effort to gain a better understanding of how professional regulations guide and govern professionals to do what they do. This chapter completes my analysis of the nature of
professions, and lays the descriptive groundwork for the normative model of professions I develop in chapters four and five.

In Part I of chapter three, I discuss the public and private theories of professional regulations. Professional regulations are a set of formal and informal rules and norms. These norms govern professional social and economic interactions and outline professional institutional structures. According to the public interest theory of regulations, professional regulations are necessary because of the prevalence of market structure imperfections, externalities, and imperfect information. Private interest theorists, on the other hand, argue against professional regulation. According to this stance, professional regulations function as market allocation mechanisms, impose anticompetitive restrictions, facilitate professional cartelization, promote special interests, restrict entry to the professions, legitimize professional authority, and impose inefficiency in the market for professional services. Most importantly, some private interest theorists outright deny the existence of market failures that might justify professional regulation.

The most significant type of market failures in the professional services market are due to information asymmetries. In an effort to explain the prevalence and problematic nature of information asymmetries, I discuss information economics in Part II of the third chapter. I focus on the problem of selection (adverse selection) and the problem of incentives (moral hazard), and survey some of the game theoretic strategies that may be employed to respond to the lack of information. For example, employers may prefer to know about the talent and productivity of a potential employee, and insurance companies may like to know the likelihood that their clients might become sick or have accidents, but this information is not readily available (adverse selection). Alternatively, employers may like to know how much effort their employees exert at their work, and insurers may prefer to know what precautions their clients take to avoid getting
sick or getting into accidents, but this information is often unobservable (moral hazard).

Resolving moral hazard and adverse selection problems is a complex task. I model a number of standard game theoretic problems of information, and provide some contract design solutions (e.g. the Harsanyi doctrine, the pooling vs. separating equilibria, and the revelation principle) for resolving these problems.

In **chapter four**, I describe trust as the “mystery of professions” that distinguishes professional associations from other private economic entities. I start by making a core descriptive background assumption: that the nature of professions and professional regulations is, at least to some extent, socially beneficial. Within my reconstructive approach, making this descriptive assumption about the socially beneficial nature of professions corresponds to a certain implicit morality of the professions. My analysis in the first three chapters provides the necessary empirical evidence about what professionals do, and what they have historically done, that can go towards supporting this stance.

I present a normative model of professions that describes them as trust-creating, trust-preserving institutional mechanisms that create mutual benefit by responding to information asymmetries and by raising efficiency in the market for professional services. I argue that information asymmetries in the market for professional services are the key for explaining the professional reliance on trust. Widespread disagreement has persisted in the literature on professions over the professional feature that qualifies as *the* fundamental defining trait. Power, knowledge, codes of ethics, autonomy, etc. can each explain some set of scenarios well, but fall short of explaining all others at the same time. I argue that that these characteristics are all symptoms of the lack of trust and can be best explained in terms of information asymmetries.
The earlier chapters focused heavily on sociology, history, and economics. This focus was necessary to set up the descriptive support for the reconstructive normative model I advance in chapter four. By drawing on this model, I can explain why professionalism involves being “ethical,” as in socially beneficial, mutually beneficial, and efficiency promoting, as opposed to unselfish and altruistic, as many scholars of professions contend. At the same time, I can use this framework to explain that cartelizing professional institutions are a distortion of professionalism (because they take a toll on efficiency), thereby refuting cartelization as the core feature of professionalization. Conceiving of the normative nature of professions as involving the facilitation of trust is somewhat of a middle road between the dichotomy of cartelization and altruism. The information theoretic characterization of professions also allows me to stay clear of any ambitions to define professions using broad, generalizing platitudes.

Next, in chapter five, I discuss professional morality. I begin by arguing that the implicit morality of the professional market consists of imperatives that are generated when the goal of economic efficiency is taken as an end. Avoiding market failures arising from information asymmetries constitutes one such imperative that promotes efficiency. The normative force of professional moral obligations is in turn derived from the mutually beneficial efficiency gains achieved via the trust-creating professional institutions that respond to information asymmetries. The efficiency promoting notion of trust has normative significance, according to this approach, because the implicit morality of markets for professional services involves, at minimum, achieving economic efficiency as an end. Professional obligations can thus be moral because of the mutually beneficial efficiency that is achieved by professional institutional mechanisms. Specifically, the professional role sometimes requires incumbents to promote efficiency in the market for professional services, whereas promoting efficiency is ordinarily simply permissible in those circumstances outside the role.
I then consider instances where the professional role seems to either require or permit us to do what would otherwise be forbidden. As I show, partiality, the moral division of labor and adversarialism are three increasingly controversial variations of this kind of permission. Providing a moral justification for such claims is one of the most challenging problems in the study of professional ethics. This analysis is important since much of professional practice occurs precisely in this moral space. I suggest that roles are a useful framework for deliberating about the moral relevance of such permissions. The way social institutions conspire to achieve outcomes involves differentiating responsibilities into distinct roles. After discussing the literature on social roles, I propose a conceptual framework that demonstrates that institutional roles can sometimes permit, to a limited extent, what would otherwise be forbidden. In other words, I argue that roles can be constrained instances of deontic weakening, and sui generis sources of moral obligations. This understanding of roles in turn informs my stance on the relevance of professional moral obligations, as less than all-things-considered moral, but nevertheless moral.

Even though roles at best constitute circumscribed instances of deontic weakening, they merit attention because of their importance in shaping our everyday moral experience. Our lived social interactions, after all, occur at the level of institutional experiences and social roles, not at the level of abstract morality. A serious descriptive engagement with our organizational and institutional moral experiences can thus inject an element of “facticity”\(^{31}\) into the debate over professional morality. Towards this goal, in my final chapter, I apply my argument for a normative model of professionalization to the institutional environment surrounding managers in private economic entities.

\(^{31}\) Habermas, *Between Facts and Norms*, 288.
Is professionalization an appropriate concept for analyzing business and the role of managers in society? I begin chapter six by illustrating that a survey of the management literature on this question reveals a lack of consensus, and a deficit of nuanced definitional and normative analysis. The underlying definitional uncertainty surrounding the professions has not received adequate attention among management scholars, and, more problematically, professions are presented as uniformly altruistic promoters of the public interest, even by those who argue against viewing managers as professionals.

I argue that managers may be viewed as professionals under my normative model of professionalization, since, in addition to drawing on market-oriented incentives, they typically appeal to norms such as trust and loyalty to make Pareto-improvements in the corporation. I outline a set of managerial efficiency imperatives generated towards members of the corporate chain of command, including employees, boards of directors, and shareholders. Professionalism thus involves following efficiency imperatives that guide managers to fulfill principal-agent obligations, despite prevalent moral hazard problems. These moral hazard problems make it challenging to properly observe or evaluate the quality of work and the effort level of managers. Based on this approach, the failure to fulfill efficiency-promoting imperatives is a distortion of managerial professionalism.

As I show, an analysis of information asymmetries surrounding the relationship between managers and owners can provide a complementary explanation to the mainstream understanding of the phenomenon of the separation of ownership and control, and the growth of the salaried managerial position in the twentieth century. Instead of merely explaining the phenomenon via optimal agency costs and external pecuniary incentives, I draw attention to the presence of norms that may have signaled managerial trustworthiness as well as efficiency-
promoting capacities to owners. Therefore, I show that the separation of ownership and control in corporations is precisely the kind of transaction that might have been lost in the absence of appropriate norms surrounding owners and managers in corporations. In this manner, the increased reliance on trust that is traditionally typical of professionals (as opposed to, say, used car salespersons) would fit with a reconstructive account of the historical development of socially beneficial management practices.

Finally, drawing on my definition of professions and my approach to professional morality, I outline the normative features that characterize managers as concerned with more than merely self-serving – yet less than altruistic – ends. I show that the managerial role generates a set of deontic constraints towards the end of efficiency-promotion within the internal hierarchical chain of command of the corporation. I then argue that given certain background conditions, the managerial role can be a sui generis source of moral guidance. Partiality, division of moral labor, and adversarialism are three instances of controversial permissions, granted by the managerial role, to do what would otherwise be forbidden. Thus, my normative model of management professionalism provides an appreciation of the limited, but nevertheless genuine force of managerial moral obligations.
The Essence of Professionalization: A Historical & Theoretical Survey

What is the essence of professionalization? What is the definition of a “profession”? Scholars of professions have not reached a consensus on an answer to these questions yet. This chapter provides a survey of the historical and theoretical accounts of the nature of professions over the past century. In section 1.1, I provide an etymological account of the term “profession,” and considers the historical development of professions. Beginning with the guilds, (arguably) the historical predecessor of today’s professions, section 1.1 explores some of the historical confusions surrounding the use of the word “profession,” from the Middle Ages through to the twentieth century. In section 1.2, I focus on the extensive literature produced in the twentieth century, and survey key approaches (broadly categorized under functionalism and structuralism) to thinking about the essence of professionalization. Despite the large volume of contributions, confusions surrounding the definition of professions persist to the present day, and a number of scholars have even called attempts to define professions futile. A recent literature review concluded that scholars who continue to write on the professions have “accepted definitional uncertainty and moved on.”

In section 1.3, I argue that a dispute over the nature of professional morality lingers in the background of many definitional contributions. While some argue that professions are a locus of

---

morality, others insist that professions constitute a more cynical attempt to rig the market and create monopolies\(^2\) (or in more precise terms, cartels,) for members of professional associations. Thus, a satisfactory account of the definition of professions must take a position on the dispute in the “morality vs. monopoly” theoretical debate.

1.1 Etymology and History, Prior to the 20\(^{th}\) Century

**Etymology**

The word profession in the English language is derived from the classical Latin *professio*, denoting “to declare publicly, own freely, acknowledge, a vow,” which is derived from the past participle *professus* of the verb *profiteri*.\(^3\) In the time of Cicero and Tacitus, *Professio* referred primarily to an oath, vow or declaration, and this meaning continued in the monastic culture of the early and high Middle Ages.\(^4\) Before 1500, to “profess” had only a religious meaning in English, and the passive form “to be professed” meant “to take the vows of some religious order, esp. to become a monk or nun.”\(^5\)

---

\(^2\) A monopoly is a market that is made up of a single entity, in contrast with the perfectly competitive market, which is made up of infinitely many entities. Economists have historically expended time and energy studying these two extreme forms of markets, i.e. monopolies and perfectly competitive markets, because of the lack of strategic interaction in these environments. In a monopoly, there is no strategic interaction by definition, and in a perfectly competitive market it would be unreasonable for any given entity to attempt to track the infinite strategic actions of its competitors. See Prajit K. Dutta, *Strategies and Games: Theory and Practice* (Cambridge: MIT Press, 1999), 75. In comparison to a monopoly, a cartel is made up of a group of firms offering the same services, and exercising monopolistic control in the market. Professions are cartels (although they are often called monopolies in the non-theoretical literature) since although each member of a professional organization enters into transactions with various clients, no single member is large enough to qualify for the title of monopoly. See Ira Horowitz “The Economic Foundations of Self-Regulation in the Professions” in *Regulating the Professions*, eds. Roger D. Blair and Stephen Rubin (Lexington, Massachusetts: Lexington Books, 1980), 9.

\(^3\) *Oxford Latin Dictionary*, s.v. “professio”.

\(^4\) Ibid.

\(^5\) *Oxford English Dictionary (OED)*, s.v. “profess”. 
The noun “profession” is defined by the *OED* as a “declaration, promise or vow made by one entering a religious order;” “the action of entering such an order;” or “the fact of being professed in a religious order.”⁶ This use of the noun can be traced to the 13th century, and by the 14th century, the meaning was extended from a vow to an act of belonging to a group of individuals who made similar vows.⁷ “Profession” was thus subsequently taken to be “a particular order of monks, nuns, or other professed persons.”⁸ In Catholicism, “the profession of faith” continues to signify a public avowal of faith that describes the shared beliefs of a religious community.⁹ In *The Profession of Faith of the Savoyard Vicar*, for example, Rousseau provided instructions for the proper religious education of pupils.¹⁰

Starting in the sixteenth century, the verb “profess” was taken to mean “to declare oneself expert or proficient (in some art or science).”¹¹ This declaration entailed that one made something “one’s profession or business.”¹² In 1577, for example, one can find passages such as: “Ozias as we reade professed husbandry.”¹³ Similarly, in 1651, Thomas Hobbes referred to those “that professeth the study of the Law.”¹⁴ During this period, “profession” came to be known as “the occupation which one professes to be skilled in and to follow.”¹⁵

*Confusions regarding etymology*

---

⁶ *OED*, s. v. “profession”.
⁷ Ibid.
⁸ Ibid.
¹¹ *OED* s.v. “profess”.
¹² Ibid.
¹⁵ *OED*, s. v. “profession”.
Historically, there has been confusion surrounding the etymology of the term “profession.” According to T. M. Stinnett, “few words are so loosely used as profession.”\(^{16}\) In order to describe “typical professions,” classic contributors to the literature such as Alexander Carr-Saunders and Paul Wilson\(^ {17}\) explore historical claims about vocations that were considered professions “by common consent.”\(^ {18}\) But although the authors imply that a common historical consensus did in fact exist on the meaning of professions,\(^ {19}\) they arguably do not provide adequate scholarly support for such a claim.\(^ {20}\) The etymology of the term “profession” was similarly overlooked by influential scholars of the history of professions, such as Burton Bledstein, who focused on the history of words “middle,” “middle-class,” “amateur,” and “private”, but not “profession” or “professional.”\(^ {21}\) This tendency towards weak historical scholarship is common in writings on the professions.\(^ {22}\) Often, scholars use the literature on the subject of professions in the twentieth century to comment on the historical development of professions—as opposed to conducting research about the nature and meaning of professions in the intended historical era.\(^ {23}\) Conversely,


\(^{18}\) Ibid., 3-4.

\(^{19}\) Ibid., 289.

\(^{20}\) Bruce Kimball, in his meticulous research of the literature on professions, argues that the Carr-Saunders and Wilson are merely “draping themselves with the appearance of historicity” and “relied upon anachronistic interpretations of historical testimony.” Bruce Kimball “The True Professional Ideal” In America: A History (Cambridge: Blackwell, 1992), 3.


some scholars define the term profession in the contemporary setting, without considering the historical usage of the word or the implications of this history.\textsuperscript{24}

Beyond the terms “profession” and “professional,” there are etymological confusions surrounding key terms such as “learned professions.” Some scholars assume that the term referred to theology, law, and medicine in the Middle Ages, or in the fifteenth or sixteenth century, although they do not cite instances of this usage during those specific periods.\textsuperscript{25} The \textit{Oxford English Dictionary [OED]} defines the learned professions as follows: “a vocation in which a professed knowledge of some department of learning or science is used in its application to the affairs of others or in the practice of an art founded upon it. Applied spec. to the three learned professions of divinity, law, and medicine; also to the military profession.”\textsuperscript{26} As Bruce Kimball has pointed out, the problem is that although the examples provided by \textit{OED} identify each of these individual vocations as a profession in the era after the middle ages, the term “three learned professions” in fact does not appear in the literature prior to 1888.\textsuperscript{27} Instead, the original use of what became the conventional term “the three learned professions” was established by Joseph Addison in \textit{The Spectator} in 1712, when he noted “I come now to that Point of Precedency which is settled among the three Learned Professions.”\textsuperscript{28} Still, many authors assume

\begin{footnotes}
\footnotetext{26}{\textit{OED} s.v. “profession”.}
\footnotetext{27}{Kimball, “The True Professional Ideal” \textit{In America: A History}, 100.}
\end{footnotes}
that the term “learned professions” has had this tripartite meaning prior to the eighteenth century, as a result of consulting the *OED*.\(^{29}\)

Another similar historical confusion exists around the different religious connotations of the term “professional.” Although in the second half of the eighteenth century the religious sense of the word appeared prominently in texts,\(^{30}\) the *OED* refers to this usage of the term professional as “archaic” and provides only one example of it dating from 1420.\(^{31}\) More recent scholars and eminent sociologists, such as Hughes, similarly often leave out religious vocations from their list of professions.\(^{32}\) Problematically, as the term “professional,” in the sense of “occupational,” became gradually introduced to common usage in the late eighteenth and early nineteenth centuries, this oversight regarding the profession of religion may have misled some scholars to think that the “occupational” reference was the exclusive or even the original meaning of the term.\(^{33}\)

A final historical controversy worth noting is over the use of the term “professor.” Some argue that the term professor was a more “grandiose” and “prestigious” title prior to the nineteenth century, but that it later lost this glow and applied to practitioners of all vocations.\(^{34}\) Others argue for the reverse process,\(^{35}\) noting that the usage of the term “professor” after the

---


\(^{31}\) *OED*, s.v. “Professional”.

\(^{32}\) See, for example, Everett C. Hughes, “Professions” 1963 reprinted in *The Professions in America*, ed. Kenneth S. Lynn (Boston, 1967), 2.

\(^{33}\) Kimball “The True Professional Ideal” *In America: A History*, 141. This misinterpretation was epitomized in Samuel Johnson *A Dictionary of the English Language* (London, 1755), s.v. “professional”.

\(^{34}\) See, for example, the *Oxford English Dictionary*, and also Bledstein, *The Culture of Professionalism*.

nineteenth century slowly came to denote high-level specialized individuals in the education
field, as opposed to an inclusive term for practitioners of all vocations, leading to the current use
of the term university “professor.”

1.1.1 Connection with the Guilds

Guilds appear to be the historical predecessors of the professions, in so far as professions
are groups of individuals who provide services vital for the survival of their community. A
guild is a social or occupational institution, created by a group of individuals, which pertains to
their work, skill, or craft. Some scholars argue that the economic motive of establishing
monopolies was a primary motivation for forming guilds, while others believe the guilds have
in common a distinctive, unifying morality. As I will show in section 1.3, the antagonism
between the two sides of the argument, one viewing monopoly and the other morality as the
essential feature of the guilds, is replicated in the debate over the nature of professions.

“Expediency and Means of Elevating the Profession of the Educator in the Public Estimation,” 276, 304, in The
Educator. Prize Essays on the Means and Expediency of Elevating the Profession of the Educator in Society
(London, 1839).

36 Kimball elaborates that John Walker had defined “professor” as “one who publicly practices or teaches an art,”
but when Lyman Cobb abridged Walker’s dictionary in 1828, he eliminated the word “practices,” thus making the
professor less of a practitioner and more of an abstract position. See John Walker A Critical Pronouncing Dictionary
(1791; Menston, UK, 1968), s.v. “professor”; Lyman Cobb, Cobb’s Abridgment of J. Walker’s Critical Pronouncing
Dictionary (New York, 1828), s.v. “professor.”

37 Elliott A. Krause, Death of The Guilds: Professions, States, and the Advance of Capitalism, 1930 to the Present

38 Kraus, Death of The Guilds, 2.

39 Anthony Black, Guilds and Civil Society in European Political Thought from the Twelfth Century to the Present

40 See the extended discussion in section 1.3 of this chapter.
Discussing the professional group and its role in society, Durkheim notes that the “the name in history of this professional group is the guild...”\(^{41}\) He provides a description of the long history of the guilds, dating back to the pre-historic era in Rome and to a tradition related by Plutarch and Pliny attributed to the King Numa.\(^{42}\) With the fall of the Roman Empire, and the civil wars that destroyed trade and industry, guilds disappeared for some time.\(^{43}\) But Durkheim notes that around the eleventh and twelfth centuries, “in all European societies the guilds, after being in eclipse for a time, began a new existence.”\(^{44}\) In the thirteenth century, guilds “flourished” and thus formed “a necessity for the foundation of the City to the Empire at its zenith, from the dawn of Christian societies to the French Revolution.”\(^{45}\) Durkheim believes that guilds have persisted through time because they respond to “some need at once profound and lasting.”\(^{46}\) This need is to live a moral life. As he puts it: “if [a guild] is to serve its purpose it must be above all through its moral consequences, for each trade or craft association would have to become the focus of a moral life *sui generis.*”\(^{47}\) Durkheim then argues that this theme of morality can be found throughout the history of the guilds, and that even though they were eradicated after the French Revolution, their moral essence would reappear in the modern industrial professions guided by professional ethics.

Durkheim has been influential in suggesting that the rise of professional associations was connected to ancient and medieval guilds, but not everyone agrees with him.\(^{48}\) As I show below,

\(^{41}\) Durkheim, *Division of Labor in Society* (W. D. Halls trans.), 17.
\(^{42}\) Ibid.
\(^{43}\) Ibid., 18.
\(^{44}\) Ibid. (citing Levasseur)
\(^{46}\) Ibid.
\(^{47}\) Ibid.
\(^{48}\) Kimball calls this connection “tenuous” (Kimball “The True Professional Ideal” *In America: A History*, 192), although he does not provide a substantive argument to this effect.
attention to the structure of organization in the guilds reveals a different stance on the essence of guilds, one that presents them as monopolistic as opposed to moral in nature.

From about 1150-1400, guilds became the main way of organizing work in many European cities and towns.\(^4^9\) Associations of individuals could decide their own guild rules, decide who might enter the association, how training and apprenticeship would be done, what dues would be charged and what fines would be imposed for breaking guild rules. Guilds were hierarchically organized, and the guild masters bore responsibility for the obligations of the group, with a guild court overseeing the resolution of disputes.\(^5^0\) Guilds had control over the workplace since the guild masters owned the means of production, and guilds controlled the market through their monopoly over the products made or services provided.\(^5^1\) Guilds also had a relation with the state, through which they would lobby, beg, and bribe the local powers to protect their monopoly, which in turn allowed them to set their own price, and to control the training of skill in the field. Early town governments reportedly complained of “restraint of trade” and of price-rigging by the guilds.\(^5^2\)

Resolving the dispute over the moral or monopolistic nature of the guilds is beyond the scope of this chapter, but noting the dispute itself highlights the dichotomy, which continues to this day in the literature pertaining to professions, between morality, on the one hand, and monopolies, on the other. I discuss this dichotomy in more detail in section 1.3 of this chapter.

In the next section, I provide a history of the Anglo-American professions by drawing on two sources: Carr-Saunders and Wilson’s canonical contribution to the literature on the


\(^{50}\) Black, *Guilds and Civil Society in European Political Thought from the Twelfth Century to the Present*, 12-29.

\(^{51}\) Ibid., 123-28.

professions (from 1933), and Kimball’s systematic evaluation of the history of professions (from 1992). I use Carr-Saunders and Wilson to trace the historical development of the Anglo-American professions from the middle ages up to the 18th century, and then discuss Kimball’s historical analysis of professions over the 18th, 19th, and 20th centuries. Together, Carr-Saunders and Wilson’s and Kimball’s contributions provide a broad overall survey of the history of professions.

1.1.2 The History of the Anglo-American Professions

Carr-Saunders and Wilson famously note that during the eleventh century “a great movement towards association began to sweep like a wave over the cities of Europe.”53 This movement led to the “formation of associations round many aspects of social life, and among them the performance of specialized functions and the carrying on of specialized crafts.”54 By the end of the fifteenth century, this movement was universal among all classes of individuals in cities.55 Associations included the “clergy, regular and secular, of all grades; the legal, medical, and teaching professions; the merchant, the shopkeeper, and the craftsman; the persecuted alien and the despised water-bearer.”56

Shortly before the year 1200, the universities arose and teachers and students formed “gilds of learning” following the model of medieval traders and craftsmen.57 The role of guilds of all

53 Carr-Saunders and Wilson, The Professions, 289. This sentence appears in many contemporary contributions to the definition of professions.
54 Ibid.
55 George Unwin, The Gilds and Companies of London (London: Methuen and Company, 1908), 172. See also E. H. Phelps-Brown, The Economics of Labor (New Haven: Yale University Press, 1962) for a discussion of guilds in India which were hereditary and may have been the precursor to the caste system.
kinds was to require a formal license for practicing a craft. Thus, in the case of universities, “the possession of a university degree was originally nothing else than the possession of a diploma to exercise the function of teaching.” ⁵⁸ Although universities were not originally ecclesiastical, they soon came under the domination of the Church as the administration of England came to be placed predominantly in the hands of the clergy. ⁵⁹ Educated individuals in the early Middle Ages were clerks in the ecclesiastical order, and writing and accounting were rare gifts for laymen, especially since all letters were written and accounts were kept in Latin. ⁶⁰ Thus, the earliest phases of a number of vocations that later turned into professions were passed within the Church. ⁶¹ Teachers, lawyers, physicians, and civil servants were all members of the ecclesiastical order, and students in universities viewed education as a “door to the church,” which at the time implied “the door to professional life.” Ecclesiastical preferment also determined maintenance and promotions within the professions. ⁶²

But although at one time most specialized functions that we would call professions today were in the boundary of the ecclesiastics, this changed in the fully developed medieval system. The law, for example, was divorced from the church by the beginning of the fifteenth century. ⁶³ During this same period, the surgeons, apothecaries, scriveners, and also the common lawyers, emerged as secular trading guilds. Thus, these professions did not emerge out of the clerical order in England, unlike physicians, civil servants, and teachers. ⁶⁴ The vocations which had grown out of the Church became secularized after some time, with the exception of teaching, and

---

⁵⁸ Ibid., 6
⁵⁹ Carr-Saunders and Wilson, The Professions, 289-290.
⁶¹ Carr-Saunders and Wilson, The Professions, 290.
⁶² Ibid.
⁶³ Ibid.
⁶⁴ Ibid., 291.
this process was complete by the end of the sixteenth century.\textsuperscript{65} The transition occurred partly due to the growth and consolidation of these vocations, but also due to the receding role of the Church. For example, the secularization of the Royal College of Physicians occurred slowly and was finalized in 1663.\textsuperscript{66} Teaching, unlike other professions which had grown within the Church, remained in clerical hands long after the Reformation.\textsuperscript{67} Only in the eighteenth century did laymen begin to acquire more important posts at universities, and the same was true of headmasterships of public schools and fellows of colleges.\textsuperscript{68}

The medieval guilds began to decay starting in the sixteenth century. As Carr-Saunders and Wilson put it “the new social and economic forces … were antagonistic to the ancient forms of association, and created conditions which, for more than two centuries, were unfavourable to the rise of new forms.”\textsuperscript{69} By the late eighteenth century, the authors report that the French Revolution had a strong impact on limiting the guilds since “associations of members of the same trade or profession were forbidden” altogether.\textsuperscript{70} Elaborating on hostility towards what the authors call “professional association,” they note that almost until the end of the eighteenth century, the “exclusiveness, selfishness, and slothfulness” of professional associations were criticized.\textsuperscript{71} Although they do not elaborate on the relation between “guilds” and “associations” in their widely influential contribution, Carr-Saunders and Wilson seem to use the term “guilds”

\textsuperscript{65} Ibid.
\textsuperscript{66} See William Munk, \textit{Roll of the Royal College of Physicians} (1861).
\textsuperscript{67} Carr-Saunders and Wilson, \textit{The Professions}, 293.
\textsuperscript{68} Ibid., 294.
\textsuperscript{69} Ibid., 299.
\textsuperscript{70} Law of 14\textsuperscript{th} June, 1791. Marten Saint-Léon \textit{Histoire des Corporations et Métiers} (1897), 624, quoted in Carr-Saunders and Wilson, 299. Unfortunately, Carr-Saunders and Wilson do not provide an explanation for this monumental anti-guild development and its rationale. Further, the authors explicitly limit their study to the professions in England and Wales, and yet include a French quote about the state of the ancien regime, without noting the relevance of the Continental European experience on English professions – this is problematic since the word “profession” has had different meanings in the Anglo-American and Continental traditions.
\textsuperscript{71} Carr-Saunders and Wilson, \textit{The Professions}, 300.
to refer to groups involved in activities which are the “origin of those techniques around which professions are built.”

Addison influentially identified the “learned professions” as divinity, law, and physic. Carr-Saunders and Wilson explain that divinity is on Addison’s list because it was once the only profession, or the basis on which all other professions were built, and that “physic and law had a place on the list by ancient right.” They explain that the other professions are excluded from this list because by the eighteenth century, “professions were regarded first and foremost as gentlemen’s occupations.” While the professions did not offer large material rewards, they provided a “safe niche in the social hierarchy.” Thus surgeons and apothecaries are not mentioned since they were “not vocations fit for gentlemen,” teachers are omitted because they were already included in the Church, and architects and civil servants had not formed groups or formal associations yet.

It is widely recognized in more recent literature from the twentieth century, that the term “profession” commonly referred to the fields of theology, law, medicine, and education during the eighteenth and nineteenth centuries. But opinions vary regarding the various esteems of the

---

72 Ibid., 298
74 Carr-Saunders and Wilson, The Professions, 295.
75 Ibid.
76 Ibid., 294-5.
77 Ibid., 295.
78 Ibid., 294.
79 Ibid., 295.
80 See, for example, William J. Good, “The Theoretical Limits of Professionalization,” in The Semi-Professions and Their Organization, ed. Amitai Etzioni (New York, 1969), 267-8; Donald O. Schneider, “Education in Colonial American Colleges, 1750-1770, and the Occupations and Political Offices of their Alumni” (PhD diss., George Peabody College for Teachers, 1965), 166; Earl F. Cheit, The Useful Arts and the Liberal Tradition (New York, 1975), 1-2; Barbara J. Harris, Beyond her Sphere: Women and the Professions in American History (Greenwood, CT, 1978), 116; Cindy Sondik Aron, Ladies and Gentlemen of the Civil Service: Middle-class Workers in Victorian
four professions. Although theology was the most esteemed among these vocations until the fifteenth century, it slowly started losing ground to the other three. Thus, in the first edition of *An American Dictionary of the English Language* (1828), the “university” was defined to include “the four faculties of theology, medicine, law, and the sciences and arts,” which had by that point become conventionally known as the “learned professions,” as confirmed by entries in subsequent editions of Webster’s dictionary. “The professions” thus conventionally referred to the clergy, attorneys, physicians, and teachers of these three fields.

Although Kimball notes that the “profession of arms,” was frequently cited by British observers in the nineteenth century, Carr-Saunders and Wilson did not include armies under the title of professions, since “the services which soldiers are trained to render is one which it is hoped they will never be called upon to perform.” In the early days in the United States, Americans only occasionally thought of the military as one of the learned professions, and

---


85 Carr-Saunders and Wilson, *The Professions*, 3.

military obligation was inherent in citizenship rather than reserved for a specialized elite.\(^{87}\)

Another potential profession, at least in the United States, was commerce. But commerce did not qualify as a profession due to the “mediocre social credentials” of merchants during the early generations of merchants in colonial United States.\(^{88}\) Although merchants began to gain social recognition,\(^{89}\) still, commerce was not thought fit to acquire the title profession through to the end of the nineteenth century, since “the professional man was expected to avoid the pursuit of wealth.”\(^{90}\) This, as we shall see later in my final chapter, is an important idea.

It was finally in the twentieth century that observers suggested that “business should be, and to some extent already is, one of the professions.”\(^{91}\) Talcott Parsons, the influential sociologist who produced some of the most important early writings on the professions, originally wrote (in 1937) that “professions and business were fundamentally distinct due to their respective motivations of altruistic service and acquisitiveness.”\(^{92}\) But by 1939, Parsons adopted a fully functionalist approach and argued that “business was indistinguishable from a profession insofar as it was functional and rationally organized.”\(^{93}\) Still, scholars generally regarded business as an “emerging or marginal profession” through to the 1960’s.\(^{94}\)

\(^{87}\) Ibid. See also Theodore J. Cracelk, Mr. Jefferson’s Army: Political and Social Reform of the Military Establishment, 1801-1809 (New York, 1987).


\(^{90}\) Kimball “The True Professional Ideal” In America: A History, 8.


A few words are now due about professional regulations, given the key role they play in disputes over the nature and history of professions. I turn to this discussion now, before moving on to an evaluation of the definition of professions.

1.1.3 Professional Regulation & the State

In the Anglo-American tradition, Carr-Saunders and Wilson explain that “[a]ll those professions, whose ancestry can be traced to the Middle Ages, were subject to regulation by the state,” whereby the state intervened, granted privileges to, or imposed duties upon those vocations.95 At that time, no principles or specific methods were employed by the state to direct interventions in the work of professionals—at least not any that distinguished professions from other vocations.96 The objective of regulation was twofold: to ensure “competence, honest dealing with clients and fair treatment of those of inferior status,” and at the same time, to “maintain discipline and supervisory duties.”97 By the end of the nineteenth century, the English Parliament began to regulate associations systematically, in order to provide some guarantee that practitioners of certain professions possessed a minimum level of competence.98 Starting at this time, the Parliament required that a list of all professional practitioners be created, and that only competent individuals be allowed to get on to the list.99

---

95 Carr-Saunders and Wilson, *The Professions*, 304.
96 Ibid., 305.
97 Ibid.
98 Ibid.
99 Ibid.

The professions that made it unto the list were divided into five groups: i) those that provided a “vital” service (e.g. medicine, dentistry, veterinary surgery, pharmacy, nursing, and midwifery); ii) those whose service was “fiduciary in a marked degree” (e.g. legal groups including barristers, solicitors, and patent agents). The clientele of these first two types of professions was “every-man”, that is, anyone who may have required the services of these practitioners;\(^\text{100}\) iii) those professions which had an intimate relation to public safety (e.g. the merchant navy and mine managers); iv) those professionals who were employed by the state (e.g. civil servants and teachers in grant-aided schools);\(^\text{101}\) v) finally, since registering with the state implied self-government and conferred certain privileges, architects and the registered teachers requested to be regulated by the state.\(^\text{102}\) Once registered, bodies with statutory authority created rules to govern admission and expulsion from the professions.\(^\text{103}\) Each occupation had to fight for its own government regulations, and the members’ identities were drawn from their individual professional institutions.\(^\text{104}\) Occupations that successfully acquired the status of “profession” through state support began to grow in numbers in the late nineteenth and early twentieth century.\(^\text{105}\)

Historically, the temporal stages of professionalization in cooperation with the state in the Anglo-American tradition were as follows. The occupation became full-time, a professional association was formed, qualifying examinations were created, political pressure was produced

\(^\text{100}\) Ibid., 306.
\(^\text{101}\) When the state became the employer of many professionals as well as other workers, it set up tests of competency in order to “ensure efficiency of its professional servants and to do away with any suspicion of favoritism.” (Ibid.)
\(^\text{102}\) Ibid.
\(^\text{103}\) Ibid., 307.
by members of the profession in an effort to receive Royal Charter (or legal protection) for their specific work area, academic qualifications were established with the help of higher education authorities, and rules were established to ensure long-term professional development. In the UK, professions have remained, for the most part, responsible for professional education, and many were granted the authority to self-regulate.106

In continental Europe, the state was more active in organization, training, and employment in the professions,107 and status and security were historically gained by attendance at state-controlled, elite institutions of higher education. These institutions afforded individuals “elite positions in the civil service or other technical-managerial positions.”108 For example, in 19th century Russia or Poland, it was prestigious to be a graduate of a gymnasium, regardless of one’s specific occupation.109 Being a university graduate or an Akademiker was of utmost importance in Germany,110 and attending one of the grandes écoles was highly valued in France.111 Thus, “primary identity was not given by occupation,” but rather by the status gained through “elite education no matter what the particular specialty.”112 And in all these cases, the state was intimately involved in the educational system.

Consider, for example, the historical development of professionalization in Germany. After the occupations became full-time, local licensing was granted, followed by the introduction of academic degrees by the state, the creation of voluntary national professional associations, as

107 Ibid.
108 Ibid.
109 Ibid.
110 Ibid.
111 Ibid.
112 Ibid.
well as a national chamber system. Professions were (and largely remain) state-regulated in Germany, and the state was responsible for professional education and admittance to the profession. Even today, the professional regulation policies of Germany are a lot more dependent on the state than in the UK or the United States. The European Union’s overall regulatory policy has arguably moved away from the British model and towards the German model, particularly through recent widespread professional transparency directives.

Professionalization grew in the early 20th century, and eventually spread to a myriad of occupations as demonstrated by classic contributions to the literature such as Harold Wilensky’s “The Professionalization of Everyone?” A survey of the literature over the past hundred years shows the process of professionalization has applied to and affected a plethora of occupations and practitioners, including social workers, geologists, web designers, biologists, independent hoteliers, chiropractors, osteopaths, midwives, alternative medicine practitioners, physicians, interior designers, librarians, psychiatrists, dentists, dental hygienists, life scientists, financial planners, human resources managers, consultants, occupational and physical therapists, statisticians, musicians, investment planners, sound recording engineers, acupuncturists, archeologists, correction workers, herbalists, cardiac sonographers, human geneticists, adult and

---

113 In Germany, national chambers are membership organizations made up of individuals and enterprises that engage in trade, industry and services. Chambers provide assistance to their members while at the same time advising and influencing government to promote favorable policies and regulations. They are self-administered and not-for-profit, and are distinct from trade associations and employers’ organizations due to their multi-sectoral structure and broad policy missions. German chambers can be traced back to the medieval European merchant guilds. Further, chamber systems under public law are influenced by the so-called "chambres de commerce" introduced in France during the Napoleonic period. See Markus Pilgrim and Ralf Meier, “National Chambers of Commerce: A Primer on the Organization and Role of Chamber Systems” (1995) http://www.cipe.org/sites/default/files/publication-docs/Chamber_Primer.pdf (Accessed Feb 6, 2013).
continuing educators, pharmacists, obstetricians, funeral directors, anthropologists, volunteer administrators, realtors, and many more. Professionalization has also occurred in institutions and industries such as shipping, film, IT, women's movements, indigenous populations, the military, the church, the police, NGOs, not-for-profits, start-up firms, commercial theatre, executive search markets, sports, agriculture, and the state legislatures, among others.\footnote{Kimball points to “the inauguration of an almost unnoticed new trend in the structural organization of professions” as a contributing factor to the growth of professionalization: starting in 1961, fifteen states in the United States enacted legislation that allowed professionals to form corporations and associations taxable as corporations. (Kimball, “The True Professional Ideal” In America: A History, 313).}

Much of the twentieth century literature on the professions seems to center on the “definition” of professions. The main research question in this definitional project has been: what is the distinction between professions and other occupations? In section 1.2 of this chapter, I provide a survey of the attempts to answer this question.

1.2 Definitional Uncertainty

The notion of “definitional uncertainty” is reflective of the state of the literature on the professions. I classify the efforts towards providing definitions under two broad approaches – “functionalist” and “structuralist”– and discuss these approaches in turn. But first, I reflect on some of the controversies and difficulties encountered in the process of formulating a definition for professions.

On the difficulty, alleged uselessness, and failure to construct a definition
To say that there is a lack of consensus on the definition of “profession” is an understatement. The debate over definitions goes back almost a hundred years to Abraham Flexner, and scholars continue to disagree regarding the traits, attributes, processes, temporal sequences, and socio-historical structures that should be emphasized in theorizing about the essence of professions. Distinguished scholars such as Andrew Abbott use the term profession “very loosely.” Others such as Laurence Veysey call professionals “a very miscellaneous sector of the population” and suggest that “it is best simply to give up the effort to abstractly define the term professional.” Veysey thus proposes to define professions “as nothing more than a series of rather random occupations that have historically been called that in our culture.”

Sociologists today are said to no longer take for granted what a profession is or to know what kinds of workers ought to be the subject of the sociology of professions. Efforts to produce a unified definition have thus been viewed as misguided and useless by many authors. Wilbert Moore and Gerald Rosenblum’s 1970 bibliography of what they call “a truly tremendous literature” on professions is a “selected” list and yet it spans 56 pages. The list of contributions to the literature has only continued to grow since then, and attempts have been made towards

118 Abraham Flexner, "Is Social Work a Profession?" School and Society 1, 26 (1915): 901-11.
providing a definition of professions by synthesizing the various existing approaches.\textsuperscript{125} Still, the use of the term tends to vary conceptually and substantively across contributions.

There are also certain characterizations of professions that could pose a challenge to the project of defining professions. For example, Morris Cogan argues that the definition of profession tends to draw controversy and is either enthusiastically accepted or defensively rejected due to its emotive quality.\textsuperscript{126} According to this stance, rational discourse regarding a definition is challenging because “to choose a definition is to plead a cause, so long as the word defined is strongly emotive.”\textsuperscript{127} Definitions also tend to draw criticism from some group or other of aspirants to the title of professional, who seek to enjoy the myriad advantages accrued through the title.\textsuperscript{128} Similarly, Eliot Freidson observes that occupational groups may take up the term “either to flatter themselves or to try to persuade others of their importance” and these groups do not have much in common except for “a hunger for prestige.”\textsuperscript{129} Howard Becker argues that the word profession singularly refers to “a social symbol” attached to a select number of occupations.\textsuperscript{130} Others call the definition of professions “a matter of personal temperament.”\textsuperscript{131} Finally, as Abbott puts it “because the term ‘profession’ is more an honorific than a technical one, any apparently technical definition will be rejected by those who reject its implied

\textsuperscript{126} Cogan, “The Problem of Defining a Profession,” 105.
\textsuperscript{127} Ibid.
\textsuperscript{128} Ibid.
judgments about their favorite professions and ... to start with a definition is thus not to start at all."\textsuperscript{132}

In sum, reviewing the mass contributions to the definitional project in the twentieth century, the only coherent description of the state of the literature seems to be that "scholarship concerned with the professions is in an intellectual shambles."\textsuperscript{133} Commenting on the lack of consensus on a unified definition, Abbott laments that "English barristers do not necessarily train in university but rather by apprenticeship and eating dinners 'in hall.' American clergy do not generally have codes of ethics .... Yet both groups are unmistakably professions. ... People don’t want to call automobile repair a profession because they don’t want to accord it that dignity."\textsuperscript{134} And according to the most recent literature review on the definitions of professions, most researchers no longer care to resolve the definitional uncertainty.\textsuperscript{135}

Similar to their historical predecessor, i.e. the guilds, professions have experienced periods of popularity and periods of decline. Recall that the guilds could be traced to Roman times, but that they disappeared due to historical circumstances for some time, until they became widespread again in Medieval European cities. Another marked period of decline occurred around the time of the French Revolution, but guilds continued later in the form of professional associations and grew exponentially during the last quarter of the nineteenth century and the first quarter of the twentieth century.\textsuperscript{136} As I show in section 2 below, this growth was accompanied

\textsuperscript{133} Eliot Freidson, “Are Professions Necessary?” In The Authority of Experts: Studies in History and Theory, ed. Thomas L. Haskell (Bloomington, IN, 1984), 5.
\textsuperscript{134} Abbott, The System of Professions, 8.
\textsuperscript{135} Evetts, “Short Note: Sociology of Professional Groups: New Directions,” 133.
\textsuperscript{136} Carr-Saunders and Wilson view guilds and professional associations as connected along a temporal sequence. (See The Professions.) Kimball has objected to the historical connection between guilds and professions. (See “The True Professional Ideal” In America: A History.)
by a scholarly literature which fell under a mostly functionalist banner, and by and large promoted the value of professions and professionalism through to the middle of the twentieth century. Around this time, a critical turn in the literature occurred, and scholars began to focus on the problematic structural process of professionalization. This latter movement to a large extent dominated the scholarship until the final decade of the twentieth century. In the past two decades, a select number of contributions have sought to reach beyond the criticisms of professions, and underscored a more balanced position by reminding us of the more positive role of professions in society while remaining cautious of their shortcomings.

1.2.1 Functionalism

The first wave of enthusiasm about professions in the twentieth century was supported by a functionalist theoretical movement. Simply put, functionalism involves explaining social phenomena in terms of their role in maintaining or reproducing society. As Durkheim explains, “the function of a social fact … consists of the production of socially useful effects.”

For example, the military, industry, and the family each have social functions, since their respective objectives are the defense, nurture, renewal, and continuity of society. In his Division of Labor in Society, Durkheim noted that he sought to determine “the function of the division of labor, that is to say, what social need it satisfies.” Here, Durkheim’s objective was to determine “function” in terms of the impact of the division of labor on social cohesion and social solidarity. Now consider applying this functional approach to the discussion of professions. If

---

139 Durkheim, The Division of Labor in Society (W. D. Halls trans.) 45,49, and 54.
140 Ibid., xxx.
the heart’s function in the body is to pump the blood, then the profession’s function might be, say, to provide social solidarity. Professions thus exist, according to functionalists, because they serve a particular social need.\textsuperscript{141}

Broadly speaking, functionalist scholars tend to share the following characteristics: a) they view society as a social system, or as a “whole,” made up of various interconnected “parts”;\textsuperscript{142} b) they assume that dynamic interactions among these “parts” tend to push societies towards an “equilibrium”;\textsuperscript{143} c) they take an interest in studying social order; d) they are interested in studying social structures in terms of contributions to evolutionary developments in society; e) finally, they view social consensus as the basis for social order.\textsuperscript{144}

A brief note is also in order about the important role of science in the functional framework. Scientific knowledge serves a social function according to functionalism,\textsuperscript{145} and Merton’s taxonomy of scientific knowledge arguably lies in the background of sociological work that celebrates professionalization as the functional expression of intellectual progress.\textsuperscript{146} According to this taxonomy, four principles characterize the scientific knowledge upon which professional experts rely: (i) “universalism,” i.e. evaluating claims independent of personal or political considerations; (ii) “communism,” i.e. taking scientific knowledge as common property; (iii) “disinterestedness,” i.e. requiring neutrality in the examination of evidence; and (iv)

\textsuperscript{141} Ibid., 42-3, 167-70
\textsuperscript{143} Durkheim, \textit{Division of Labor in Society} (George Simpson trans.) 365, 397.
\textsuperscript{145} Kimball, “The True Professional Ideal” \textit{In America: A History}, 311.
“organized skepticism,” i.e. requiring testing and scrutinizing of all analytic claims. Parsons argued that two additional elements play crucial roles in modern professions: v) rationality, i.e. the search for objective truth, and vi) functional specificity, i.e. technical competence which gives rise to functional authority. Functionalist scholars of professions have drawn on science as a source of professional authority. Commenting on the rise of professions after the scientific revolution, Carr-Saunders and Wilson note: “many professions are based upon sciences; and nothing, short of the onset of a glacial age in the history of human mental activity, could now check the onward march of these sciences.”

From this perspective, the historical rise of professional associations is a downstream consequence of the scientific revolution. For several centuries, less than a dozen professions provided necessary skilled intellectual services, but things changed after the scientific revolution and “the flood-gates opened” for the creation of new professions. One result of the scientific movement was the stimulation of inventions, which used to be conducted by untrained persons without interests in pure sciences. This changed when military engineers brought traditional skills to the mechanical arts, and the art of civil engineering was derived. Thus, science “descended into the arena of practical problems” leading to revolutionary changes in society. As more new crafts were built upon the sciences, social and industrial organizations were created

148 Ibid. These elements are also present in other core occupational structures such as business and government administration in modern society. See Parsons, “The Professions and Social Structure,” 37.
150 Carr-Saunders and Wilson, The Professions, 296. See also 284, 295-7, 307, 365, 491, 499.
151 Carr-Saunders and Wilson, The Professions, 296. This claim is controversial. For criticisms of this stance see Sciuli, Professions in Civil Society and the State, and Kimball “The True Professional Ideal” In America: A History.
152 Carr-Saunders and Wilson, The Professions, 296.
153 Ibid., 296-7.
as well. Industrial organizations needed accountants, secretaries, bankers, insurance services, surveyors, auctioneers, and real estate agents. Further, large-scale industrial organization implied large-scale social organization, and hence local government services were needed to provide the necessary surrounding conditions in addition to the technical circumstances necessary for the professions.

Formative functionalist contributors in the early twentieth century include Spencer, Veblen, Whitehead, and Tawney. Influential classic functionalist scholars of professions include Carr-Saunders and Wilson, Parsons, Marshall, and Hughes. I briefly discuss Carr-Saunders and Wilson and Parsons below in an effort to explain the functionalist stance on professions.

**Carr-Saunders and Wilson**

Carr-Saunders and Wilson applied a functional approach to the professions in their 1933 book, *The Professions*, which was celebrated at the time as “undoubtedly the best general study in the field.” Toward the end of the twentieth century, the book was still regarded as “the

---

154 Ibid., 297.
155 Ibid.
160 Carr-Saunders and Wilson, *The Professions*.
161 Parsons, “The Professions and Social Structure.”
163 Everett Hughes, “Professions,” Daedalus 92 (1963), 655-68. Hughes’ ethnographic approach to professions arguably left him off the radar of Parsons’ followers, but he nevertheless had a formative influence in the field.
164 Frances P. DeLancy, *The Licensing of Professions in West Virginia* (Chicago, 1938), v.
standard general history of the professions in England.” Carr-Saunders and Wilson set out to study historical claims about “vocations of ancient lineage which by common consent are called professions.” The authors reported that research on professional associations in England and Wales was non-existent at the time of their writing, and that the historical development of professional associations and their respective social, economic, and ethical problems were by and large neglected. Still, Carr-Saunders and Wilson provide no systematic explanation of the connection between professional associations, professions, and the guild system, and claim that any attempt to provide a definition or distinguish between professions and other vocations would be “arbitrary.” They have thus been cited as one of the first examples of authors who believe that the definitional project is flawed.

Carr-Saunders and Wilson examine and evaluate characteristics of professionalism, and provide a number of clues to necessary and sufficient conditions surrounding typical professions. Francis Bacon is quoted in the book’s very first line, in effect legitimizing the authors’ overall historical message about professions: “I hold every man a debtor to his profession, from the which as men do of course seek to receive countenance and profit, so ought they of duty to endeavour themselves, by way of amends, to be a help and an ornament thereunto.” The authors take this quote to imply that, for Bacon, the term “profession” indicated “certain vocations with peculiar characteristics, and in this sense it has been in use for

166 In a departure from common consent, however, the authors insist that only formal associations count as real professions. See Carr-Saunders and Wilson, *The Professions*, 3, 284, 294-5, 298, 309.
168 Ibid., 4.
170 Carr-Saunders and Wilson, *The Professions*, 3.
171 Ibid., 4.
centuries.” Following Carr-Saunders and Wilson, various scholars of professions have drawn on Bacon to imply that these distinctive features, whatever they are, have existed for a few centuries.

Carr-Saunders and Wilson come close to providing a definition for professions when they note that “a profession can only be said to exist when there are bonds between the practitioners, and these bonds can take but one shape—that of formal association.” Although among the first objectives of professional societies was the desire to “promote ‘study’ activities,” with the goal of discussion, research, and publication in the chosen field, pure study societies did not count as professional associations. The authors explain the additional functions of professional associations as follows:

…members of this club had more in common than the study interest… they called themselves civil engineers, architects, and so on. To them these titles indicated men who had attained to a certain degree of competence in their own sphere. But the public accorded these titles to any one who laid claim to them, whether competent or not, and in consequence the skilled practitioners came to desire that the competent should somehow be distinguished and protected.

At the heart of professional associations, then, was an effort to distinguish the skilled from the unskilled, and to signal or communicate this distinction to the general public who sought the services of professionals, but did not have the necessary knowledge to make accurate determinations about skill level. The strategy for achieving this effect was to limit professional

---

173 Carr-Saunders and Wilson, *The Professions*, 1. Unfortunately, a unified description of these specific characteristics, and a methodical, historical evidence of their continuous existence is missing from this work.
174 See, for example, Stinnett, *The Profession of Teaching*, 14.
175 Carr-Saunders and Wilson, *The Professions*, 298. This requirement of formality seems to arbitrarily exclude instances of association that are informal yet operational, and for all practical purposes function as a formal group. The authors do not justify the necessity of this formality.
176 Carr-Saunders and Wilson, *The Professions*, 301.
177 Carr-Saunders and Wilson, *The Professions*, 301. The authors also mention collegial relationships and dinner clubs as reasons for forming associations, but alongside “study,” these purposes do not seem sufficient for formalizing a group.
178 Carr-Saunders and Wilson, *The Professions*, 301.
admission to those who could show evidence of high standards of competence. For example, the president of the Institution of Surveyors in 1877 recommended requiring students to pass examinations prior to joining the institute, so that “their respectability and character is secured and that some degree of guarantee should be given to the public that they are not unfitted for the work entrusted to them.”\(^{179}\)

In addition to helping the public distinguish between competent and incompetent service providers, professional associations sought to help the public distinguish the honorable practitioners from dishonorable ones. The formulation of ethical codes aimed at this goal, according to Carr-Saunders and Wilson.\(^{180}\) The most prominent historical example of a code of ethics, the Hippocratic Oath, is said to date back from the fourth century B.C.\(^{181}\) and versions of this Greek oath were adopted through to the eighteenth century.\(^{182}\) In 1803, Thomas Percival’s *Medical Ethics* presented the baseline for a more complex system of professional ethics,\(^{183}\) and Percival has been widely recognized as the founder of modern codes of medical ethics.\(^{184}\) This work was accompanied by a number of other publications by other English authors in the early nineteenth century, including W. O. Porter’s *Medical Science and Ethicks*\(^{185}\) and Abraham Banks’ *Medical Etiquette.*\(^{186}\) Around the same time, the Manchester Medico-Ethical

---

182 Chauncey D. Leake ed., *Percival's Medical Ethics* (Baltimore, Williams & Wilkins, 1927), 36.
Association\textsuperscript{187} and the British Medical Association’s medico-ethical committee\textsuperscript{188} were set up in response to the widespread concerns about ethics in the medical profession.

Beyond addressing ethics problems, codes of ethics also served the function of creating trust for professionals. T. H. Marshall argues that “ethics codes are based on the belief that between professional and client there is a relationship of trust, and between buyer and seller there is not.”\textsuperscript{189} As Ivan Waddington has explained, according to this traditional stance on professional ethics, the professional group in effect takes it upon itself to “guarantee the integrity of its members by the development and enforcement of codes of professional ethics.”\textsuperscript{190} It was thus hoped that through professional standards and codes of ethics “the public will come to realize that in giving patronage to members of the association they were assured of honest as well as of competent service.”\textsuperscript{191} And with this trust, came certain perceptions about social class. High technical standards and honorable service brought “prestige” to the association, since members of occupations that formed new professions sought the social status “fit for gentlemen.”\textsuperscript{192}

Carr-Saunders and Wilson finally observe that after a while, the emphasis changed from “respectability and status” to “protection of interests,” which in turn became analogous with a

\textsuperscript{187} As Waddington reports (footnote 9), the “Rules and Bye-laws of the Manchester Medico-Ethical Association” were published in 1848 and reviewed anonymously in an article entitled “Medical Ethics,” {	extit{Brit. For. med.-chir. Rev.}} (1848) 2: 1-30.
\textsuperscript{188} Waddington explains (footnote 10) that while the Association was still called the Provincial Medical and Surgical Association in 1853, it became the British Medical Association in 1855.
\textsuperscript{190} Waddington, “The Development of Medical Ethics – A Sociological Analysis,” 38. In contrast with this traditional stance that defines professional ethics in terms of the relations between practitioners and clients, an alternative approach views professional ethics as a more practical response to problems among professionals. Evidence for this reading exists as far back as Percival’s work on medical ethics. Waddington provides a historical analysis of nineteenth century writings on medical ethics and reveals a persistent concern with the “relationship between practitioners.”(39)
\textsuperscript{191} See, for example, The Records of the Society of Gentlemen Practisers (1897), 1, and the preamble to Charter of the Institute of Chartered Accountants (1880) (Quoted in Carr-Saunders and Wilson, The Professions, 302.)
\textsuperscript{192} Carr-Saunders and Wilson, The Professions, 301-3.
“relatively high level of remuneration.” This is how “protective activities” were added to the list of characteristics possessed by professions. Professional associations also became repositories of special knowledge and were approached by public and private bodies for advice. The later-formed associations often had semi-public mandates to present their views on matters of public policy.

To summarize, the professional associations that formed in the early nineteenth century in England and Wales had (i) interests in study and research; they acquired functions related to (ii) technical competence (which was achieved by raising standards for joining the group) and (iii) honor of their members (which was accomplished through ethical codes); (iv) protection of material interests as well as (v) roles in public activities were later added to the functional attributes of these professional associations.

Following Carr-Saunders and Wilson, other scholars provided historical and empirical studies of the professions. Caplow and Hughes provided historical analyses that tracked the temporal sequence of the development of professional associations. Most influentially, Harold Wilensky’s "The Professionalization of Everyone?” considered the history of eighteen American occupations, and identified typical paths through which established professions had developed. Wilensky defined professions in terms of a process –a regular, temporal sequence of

193 Ibid., 302.
194 Ibid. No rationale is provided for this shift from social closure (achieved through the pursuit of prestige, derived from competent and honorable service) to economic closure (achieved through protective activities in pursuit of higher remuneration).
195 Ibid., 303.
196 Ibid.
199 Wilensky, “The Professionalization of Everyone?”.
stages that an occupation passes through to qualify for being a profession.200 The typical sequence, according to Wilensky, was: i) the occupation became full-time; ii) training began at universities and a core group of teachers (as opposed to simply practitioners) was established; iii) graduates of training programs formed professional associations, which were organized hierarchically and competed with neighboring occupations; iv) political attempts were made to receive protection by law in order to demarcate areas of competence; and, finally, v) a set of rules in the form of ethics codes were adopted, which aimed at keeping technical standards high, promoting the service ideal, and protecting clients.201

In the next section, I turn to discussing the work of Talcott Parsons. Parsons’ work on professions remains, arguably, the most theoretically informed literature on the nature of professions.

*Parsons*

Parsons’ focus on the professions was part of his project as an influential functionalist.202 Parsons’ 1939 essay “The Professions and Social Structure”203 was reprinted later in *Essays in Sociological Theory, Pure and Applied*, leading to his canonical article in the *International*
Two core assumptions in the background of Parsons’ work are empiricism and positivism. Empiricism, in this context, refers to the claim that scientific knowledge originates from “systematic observation [or] sensory experience.”

Positivism, here, designates the dual contention that “the methodological procedures of natural science” constitute the basis for social-scientific investigation, and that “the outcome [of such] investigations can be formulated in terms parallel to those of natural science” – namely, encompassing generalizations that are uniform or universal.

Parsons argued that professionalization is a process, “which in one aspect is almost synonymous with that of rationalization,” for the creation of disinterested professionals. He showed that three elements play functionally significant roles in the modern profession, and that these elements are also present in other core occupational structures such as business and government administration in modern society. These elements are i) “rationality” (which includes the search for objective truth), ii) “functional specificity” (which relies on the essential characteristic of technical competence and gives rise to functional authority) and iii) “universalism” (which yields disinterestedness as a social institution).

Parsons identified the antecedents to his approach to include the works of Durkheim and Weber. The

---

206 As Camic explains, “scholars have generally agreed that Parsons’ own definitions of the terms empiricism and positivism are ‘somewhat idiosyncratic’ when judged by modern or historical usage.” (433, footnote 2)
207 Parsons, “Professions,” 536-7, 544.
philosophical foundations of Parsons’ works may be traced to the epistemological works of Karl Popper.\textsuperscript{211}

Parsons defines professions as follows:

In sociological terminology, a profession is a cluster of “occupational” roles, that is roles in which the incumbents perform certain functions valued in the society in general, and by these activities, typically ‘earn a living’ at a ‘full time job’. Among occupational role-types, the professional is distinguished largely by the independent trusteeship exercised by the incumbents of a class of such roles of an important part of the major cultural tradition of the society. This means that its typical member is trained in that tradition, usually by a formally organized educational process, so that only those with the proper training are considered qualified to practice the profession. Furthermore only members of the profession are treated as qualified to interpret the tradition authoritatively and, if it admits of this, to develop and improve it. Finally, though there usually is considerable division of labor within such a group, a substantial proportion of the members of the profession will be concerned largely with the ‘practical application’ of the tradition to a variety of situations where it can be useful to others than the members of the profession itself. The professional man is thus a ‘technical expert’ of some order by virtue of his mastery of the tradition and the skills of its use.\textsuperscript{212}

Thus, Parsons identified carrying out socially valuable functions as a core defining feature of professions. Technical expertise, remuneration, social/cultural merit, training, educational prerequisites, professional authority, and client (in addition to) peer service orientation are also highlighted in this definition.

Parsons’ view is sometimes called “structural functionalist,” which is an accurate title in so far as Parsons relies on Durkheim’s conjecture that the structural division of labor exists by

\textsuperscript{211} The same may be said of Merton’s work. For the “Mertonian” conception of the scientific enterprise in Parsons’ work, see Parsons “The Present Position and Prospects of Systematic Theory in Sociology,” 212–37. As Camic, clarifies in “The Making of a Method: A Historical Reinterpretation of the Early Parsons,” Parsons was determined to make sociology a “science.” See Karl Popper’s \textit{Objective Knowledge: An Evolutionary Approach} (Oxford: Clarendon Press, 1972).

virtue of social functions. Variations of this definition appear in countless attempts to provide a definition or a socio-historical description of the development of professions, for example by Caplow, Hughes, Wilensky, Moore and Rosenblum, Elliott, Vollmer and Mills, Etzioni, Cogan, Becker and Millerson.

The functionalist definition of professions dominated the literature in the first half of the twentieth century and to a large extent celebrated the role of professions in society. The historical/temporal representation, epitomized by Carr-Saunders and Wilson, alongside the trait-based definitional project of Parsons, produced streams of followers that attempted to demarcate professions as a distinct occupational group. Most functionalist scholars agreed that professions were occupations distinguished from others by “their orientation to serving the needs of the public through the schooled application of their unusually esoteric knowledge and complex skill.”

I now move to a discussion of a second approach to defining the professions, which can be broadly referred to as Structuralism.

---

213 See Durkheim, *The Division of Labor in Society* (W. D. Halls trans.), xxxv. Note, however, that continental scholars (e.g. Durkheim) typically use the term profession to refer to all occupational groups, as opposed to specialized, distinct groups, as was done in the Anglo-American context. On this distinction see Kimball “The True Professional Ideal” in America: A History. (313, footnote 47)

214 Caplow, *The Sociology of Work.*

215 Hughes, *Men and Their Work*; Hughes, “Professions.”

216 Wilensky, “The Professionalization of Everyone?”


221 Cogan, “Towards a Definition of Profession”; Cogan, “The Problem of Defining a Profession.”

222 Becker, “The Nature of a Profession.”


1.2.2 Structuralism

One of Parson’s most basic assumptions was that the “function” served by professions was a positive one—their contributions were a contribution to “society.” Gaining momentum in the second half of the twentieth century, structuralist scholars began to take a critical stance against professions and professional associations. In defining the distinctive properties of professions, they focused on various key concepts such as monopoly, power, and autonomy. Although some authors continued to be influenced by Parsons, a widespread shift in focus among scholars of professions gained momentum after the 1950s. While the sociological literature prior to this period focused “primarily on the analysis of professional norms and role relations and on interaction in work settings,” it also tended to neglect the political and economic repercussions of professionalization. Literature after this time, however, focuses on “the political influence of professions, on the relation of professions to political and economic elites and the state, and on the relation of professions to the market and the class system.” While sociologists had so far focused on medicine, law, engineering and architecture, the list slowly started to grow, as professionalization did through to the late 1980’s and early 1990s, and began to include myriad other disciplines such as education, social work, women’s study, management, politics, and many others.

Before discussing this literature, it is important to note that economists had been critical of the “closed, monopolistic character of the professionalized labour market” throughout the first half of the twentieth century. A number of scholars in political science had also been

---

225 Ibid., 19-20.
concerned about “professions as privileged private governments,” and some policy makers saw professional experts as “overnarrow and insular in their vision of what is good for the public.”

With the growth of professions in the twentieth century, the list of professional achievements was extensive. As Bledstein puts it, these achievements came about in part because “[s]cience as a source for professional authority transcended the favoritism of politics, the corruption of personality, and the exclusiveness of partisanship.” Still, as social and historical critics were quick to point out, professionals drew on science en route to an explicit pursuit of class privilege and undemocratic authority. The overall functional approach to science first came under attack through Thomas Kuhn’s influence. Kuhn argued that the community of scientists constituted a plurality of groups committed to various paradigms, and that scientists did not, and could not, invoke a set of general scientific rules. Different paradigms had various capabilities to resolve problems, and the foundations of expert knowledge were thus ripe with conflicts and disagreements among communities of scholars (as opposed to a uniform functional consensus). As Bernstein explains, “[s]ituating scientific controversy within its social and historical contexts has become the preoccupation of scholars ever since Kuhn’s intervention in

Reuben Kessel, “Price Discrimination in Medicine” The Journal of Law and Economics I (1958): 20-53; Walter Gelhorn, Individual Freedom and Governmental Restraints (Baton Rouge: Louisiana State University Press, 1956), 106. These authors are critical of professional licensing to varying degrees. For example, while Friedman’s solution is to abandon licensing, Gelhorn believes licensing can perform a real function through appropriate reforms.


229 Bledstein, The Culture of Professionalism, 90.


the literature.” Expertise has historically been deployed by credentialed elites. Thus, naturally, the empirical and conceptual findings of functionalists about the process of developing and demarcating technical expertise inspired arguments from critics regarding the development of monopoly power and the exclusionary nature of professionalization.

First, sociologists began to attend to the legitimating capacity of professional expertise, and underscored the structural factors that endorsed the authority of professionals in their social environments. Cogan has documented the growth of critical scholarship after the 1950s, and Friedman and Kuznets, Becker, and Hughes were among those who first questioned the widespread enthusiasm for the professions. As Gross and Osterman note, this critical movement was so successful that by 1972 “virtually every one of the major professions [was] deeply troubled, and most [were] in the process of renewal, reconstruction, reform and,
occasionally, revolution." Criticism of the professions continued to grow through to the late 1980s and the 1990s.

Scholars most prominently associated with this analysis, such as Larson, Starr, and Abbott, tend to discuss the “underlying structural processes,” “structural effects,” and “structural necessities” of the professions. The anthropological school of “structuralism” and the Marxist sense of “structures” are both in the background of this critical movement, and both claim that social structures help explain the human experience. Whereas the functionalist literature studies the utility of professional expertise in order to explain the professions, structuralist literature underscores the organizational structure of professional associations and explores the relation between associations and the broader socio-economic structures within which they exit.

Structuralists challenge the professionals’ claim to legitimate expertise, and seek to expose professions as self-conscious monopolies concerned primarily with their prestige and their members’ high incomes. As Larson puts it, “[p]rofessionalization is thus an attempt to translate one order of scarce resources – special knowledge and skills – into another – social and economic rewards.” Crucially, the scarcity of “special knowledge and skills” is often viewed
as artificial, due to “a gnawing suspicion that [professional] self-regulation generally manifests itself in anticompetitive restrictions.”

In this manner, the structuralists question the objectivity of scientific expertise, which is a cornerstone of the functionalist approach. According to these critics, professionals “barricade their expertise,” and make their knowledge “artificially abstruse and arcane,” in an effort to “mystify and exclude the laity.” Some scholars concentrate on the power wielded by professional organizations, while others are more interested in the power exercised by individual professionals. Another prominent critical stance is to underscore “autonomy” as the defining characteristic of a profession. As Melosh explains, professionals have an unusual level of independence in defining the scope and application of their expertise. For example, in the case of medicine “physicians claim not only the exclusive right to diagnose and treat illness, but also the right to control the division of labor in health services, to regulate related goods and services (such as the pharmaceutical industry), and even to name and control the social experiences of sickness and health.”


250 Ibid., 314-5. See also Johnson, Professions and Power, 41-61; Markowitz and Rosner, “Doctors in Crisis”; Illich et al., Disabling Professions; Klegon, “The Sociology of Professions”.

251 See, for example, Goode, “The Theoretical Limits of Professionalization”; Larson, The Rise of Professionalism; Freidson, Profession of Medicine.


253 Freidson, Profession of Medicine.

The critical models of professions look beyond the structure of work, and consider professions as “particular expressions and vehicles of dominant class and culture.” Through this lens issues of race and gender also become relevant alongside expertise. Melosh argues that women’s status as the “second sex” implies that “there can be no women’s profession.” We can identify female members of professions, but even our ways of speaking about them betray the anomaly of women in these positions as we mark the exceptional character of the “woman doctor,” the “woman lawyer” and the “lady professor.” Under this analysis, nursing could not qualify as a profession, since nurses’ autonomy is constrained by the dominance of the medical professional and most nurses were women in the twentieth century.

The distinction between functionalists and structuralists crucially revolves around whether technical expertise or social structure should be regarded as sources of professional authority. The structural critical stance can be thought about as yet another approach to defining professions, this time not according to traits, processes, or temporal sequences, but rather from a “power,” or a “monopoly,” perspective. In this context, professions necessarily involve a process of social closure and a consequent exploitation of the monopoly supply of knowledge-based skills and services. Striving for professional status is an attempt to keep out competing interest groups, and to maximize political influence and economic gain.

---

256 Ibid.
257 Ibid.
258 Ibid., 671. Critics of functionalism had hoped to replace the “list of attributes” of professions with a distinct historical and sociological description. But defining the essence of professions around notions such as “autonomy” (following Freidson, Melosh, and others) is arguably merely another candidate for a list of attributes.
260 According to Starr (The Social Transformation of American Medicine), critics of functionalism seek to provide a broad structural and historical description of the professions, in an effort to replace the previously unsuccessful attempts to provide a list of attributes.
In the 1990’s, scholars began to reevaluate the value of the professions. Key concepts under study have since included trust, risk, expert judgment, and institutional change in the professions.\(^{263}\) Scholars have also explored the relation between professionals and the state, and state support for the professions.\(^{264}\) Interestingly, there has been a return to Parsonian roots,\(^{265}\) and efforts have been made to underscore the positive role of professions in society again. As Evetts suggests, scholarship today is moving “away from market closure as the dominant paradigm” and instead involves “reappraisal and reassessment of professionalism as a normative value.”\(^{266}\) This is especially the case for Freidson, who in his latest work, argues that professions can be a desirable form of occupational market control that successfully provide complex services to the public.\(^{267}\) Scholars thus began to argue that the virtues of professionalism ensure that high standards of service are met, and a certain amount of market closure might be necessary to ensure that adequate education, training, and collegial professional relations are developed.\(^{268}\) Others have argued that the public interest and the professional interest may be pursued at the same time\(^{269}\) and that professional values encourage cooperation.\(^{270}\)

---


\(^{269}\) Mike Saks, *Professions and the Public Interest* (London: Routledge, 1995).

Here, we can see a fundamental ambivalence on display. Although scholars continue to write about professional organizations and the professionalization project with respect to their impact on “institutional fields”\textsuperscript{271} widespread disagreement continues over whether professionalization is good or bad. In order to explore this question, I now turn to discussing the normative issues underlying the debate.

1.3 Morality and the Professions

Much of the literature in the first half of the 20\textsuperscript{th} century praised professionals and their dedication to service and a professional service ethic.\textsuperscript{272} This approach is evident in the works of early contributors to the literature such as Brandeis and Flexner, and echoed by various theorists in subsequent years.\textsuperscript{273} I provide a brief overview of this literature below.

1.3.1 Professions as Inherently Moral Entities

Professionals sometimes distinguished themselves from trade and labor unions through ethics, so that while “the labour union exercises its monopoly for the sole purpose of gaining an economic higher ground ….[i]t is unthinkable that a professional secret should be used for

\textsuperscript{271} See, most recently, Daniel Muzio, David Brock, and Roy Suddaby “Professions and Institutional Change: Towards an Institutionalist Sociology of the Professions” \textit{Journal of Management Studies} 50, 5 (2013).

personal gain.”

In the legal profession, for example, Pound noted that the “spirit of public service” is a pivotal aspect of the profession, alongside “organization” and “learning,” while, in contrast, “the idea of gaining a livelihood is incidental.”

Similarly, the professions were exempted by the State Supreme Court of Washington from an excise tax on business activities in 1933, on the following grounds: “A profession is not a money getting business. It has no element of commercialism in it. True, the professional man seeks to live by what he earns, but his main purpose and desire is to be of service to those who seek his aid and to the community of which he is a necessary part.”

Thus, professional service, alongside professional expertise, marked a pivotal attribute of the professions, and continued to do so throughout the 20th century. As the chancellor of New York University put it in 1954, “prosperity and … happiness can both be attributed – insofar as we attain them – to the professions: To their growth, to their ever-increasing assumption of responsibility in providing for the needs and wants of the people.”

Commenting on the social role of engineers and architects in the inter-war period, Kohn noted that one socially significant effect of WWI was an impulse to “appraise the motives that have inspired various professional ideals,” since everyone shared an “almost universal desire for service, the sacrifice of one’s private interests to the common good.” He claimed that during the war, for once, “the money-

274 Brubacher, Teaching Profession and Practice, 11, 15.
making motive was laid aside.”

The work of all types of professionals should thus be equally bound by rules that aim to maintain professional service standards, and “equally free from the vulgar subordination of moral standards to financial interests.”

Those who see professions as having a “substantial moral commitment to the welfare of their clients” are described as like-minded people who share “an ethic of service to their clients” in addition to a certain level of arcane knowledge and skill. This ethic of service, which involves an altruistic concern for clients and a self-imposed duty of beneficence is described as “a central and necessary feature of the concept of a profession and distinguishes professions from most other groups of workers.” Tradespeople, in contrast, might act benevolently towards their customers, but will not put their customer’s interests ahead of their own interest to make a profit. Important theorists of professions have thus argued for a distinction between professions and trades. Despite the fact that professions were increasingly identified with technical expertise in the 20th century, some scholars continue to argue that professions need to have an “ethical dimension” as well as characteristics of technical expertise.

This stance on the relation between morality and the professions is not confined to the early 20th century. Consider Plato’s remark that “the physician as such studies only the patient's interest, not his own.” According to the Stoics, doctors had a special moral duty of

---

280 Ibid.
281 Ibid., 3
282 “Medicine, Profession, and Society,” 59.
283 Ibid.
284 Ibid.
285 See, for example, Freidson, Profession of Medicine; Larson, The Rise of Professionalism; Haskell, The Emergence of Professional Social Science.
beneficence, and in the first century AD, Scribonius Largus's physician was told never to “neglect love of humanity and all the duties it entails.” According to this account, as members of the medical profession, doctors have a moral duty of beneficence to their patients (as enunciated in the Hippocratic oath and its modern successors), and this moral duty is a pivotal aspect of the nature of the medical profession. Even today, after so many criticisms of professions and their moral role, many view professions as somewhat more ethically guided than markets and bureaucracies. Lewis and Maude, for instance, believe that “a moral code is the basis of professionalism.” The intimate connection between morality and professions continues, and is apparent in Richard Greenstein’s recent analysis of the role of professional regulation on professional ethics. He asks: “What if being a professional turned out to be inconsistent with ethical conduct? That would be strange. For when an occupation calls itself a profession, this declaration is normally taken to signal a commitment to high levels of both competence and ethical practice.”

1.3.2 Professions as a Means of Promoting Self-interest

Not everyone shares this optimistic assessment of the role of professions in society. Challenges to functional expertise and to the legitimacy of professional authority are closely connected with charges of instrumental morality and disguised promotion of self-interest. Many critics of the professions argue that the ethos of service orientation may have been a rhetorical

---

289 Ibid.
290 “Medicine, Profession, and Society,” 59.
strategy for professionals to deflect criticism of the power and prestige they acquired through professionalization.

The reaction of the critics of professionalization to the ethics of service ranges from “mild skepticism” to “curt dismissal,” and many share the “hostile conviction that professionals are wolves in sheep clothing, monopolists who live by the rule of caveat emptor, but lack the integrity to admit it.” Some have worried that while professionals want to sound altruistic, they in fact have motives of social closure and power. Critics have pointed out that professionals tend to reach a dominant social position with the support of sponsoring elite groups, for example, through the help of philanthropic foundations in the empowerment of the American medical profession. As Melosh explains:

In this view, commitment to service does not define which occupations become professions; instead, a claim to altruism serves to legitimate professional prerogatives after the fact. Professionals are no more or less disinterested than any other workers, but because they claim power in areas of intense cultural concern, such as sickness and health, they must justify their control and reassure their anxious clientele by seeming altruistic.

Melosh notes that ethical codes in the professions thus take on a special ideological role, and help to “justify professional autonomy by manifesting the profession’s commitment to service and providing the appearance of conscientious self-regulation.” The functionalist approach to thinking about professions has thus been taken by some to imply instrumentalism, whereby professionals use professional associations as tools for gaining social and economic

---

293 See Lierbermann, The Tyranny of the Experts; Markowitz and Rosner “Doctors in Crisis”; Illich et. al., Disabling Professions; Klegon, “The Sociology of Professions”.
294 Thomas Haskell “Professionalism versus Capitalism,” 181.
297 Ibid. J. Brown has examined the rhetorical strategies and language that professionals use to assert, legitimate, and maintain their dominance. See J. Brown, “Professional Language: Words that Succeed,” Radical History Review 34 (1986): 33-51.
privileges. Codes of ethics of professions, according to this stance, serve the function of legitimizing the profession and disguising self-interested behavior.298

Prominent functionalists such as Carr-Saunders and Wilson did not discuss morality, altruism, or ethics directly, but rather focused on the role of gentlemanly prestige, honor, and honesty in professional practice. Still, proponents of the professions have often been taken to have altruistic objectives.299 This title of altruism clashes with what critics have claimed were questionable aims of economic closure pursued by the professions. From this perspective, professionals organize into associations in order to exclude unqualified individuals, but more importantly to gain monopolistic market control of their occupational service. Explicit mechanisms such as entry prerequisites, institutionalized academic education, and on-the-job training and experience are some of the conditions for membership in the professional group. Once the regulatory guidelines of a profession have been sanctioned by the state in the form of a registration or a license, then professional monopoly is fully established and certified professionals have the exclusive right to service.300 Illich, for example, attacks the medical profession's “self image of benevolent caretaker,” and describes the medical profession as a “radical monopoly” where “social control” of the population by the medical system is “a principal economic activity.”301 The core assumption in the background of these criticisms is that both social and economic closure are morally suspect, and that morality is simply an instrumental tool used to divert attention from the questionable monopolistic goals of professionals.

301 Ivan Illich, Limits to Medicine (London: Marion Boyars, 1976), 43.
Criticism of the professions and their moral role in society is not unique to the second half of the 20th century. Characterizations of professionals as essentially self-serving seekers of wealth, power, status and exclusivity, whose declarations of concern for their clients are mere means to their self-seeking ends can be found throughout history.\textsuperscript{302} Consider Chaucer, who notes when describing the physician in his Canterbury Tales, that “gold in physik is a cordial, therefore he lovede gold in special.”\textsuperscript{303} The leveller Samuel Hartlib is quoted by Sir Douglas Black as follows: “The liberty of our commonwealth is most infringed by three sorts of men, priests, physicians, lawyers. The one deceives men in matters belonging to their souls, the other in matters belonging to their bodies, the third in matters belonging to their estates.”\textsuperscript{304} Similarly, Shaw held that “all professions are conspiracies against the laity.”\textsuperscript{305} In the mid-eighteenth century, opinion was hostile to professionals in England, and the writings of Dickens epitomized this criticism such that “to most men, barristers and public servants recalled Jarndyce and Jarndyce and the Circumlocution office.”\textsuperscript{306}

\subsection*{1.3.3 Conclusion}

Coming up with a synthetic amalgamation of the various efforts to define professions is a difficult task. In the next few chapters, I provide a novel theoretical approach that explains why professions exist and what role they play in the economy. For now, a broad, non-comprehensive summary of the various superficial characterizations discussed in this chapter is as follows:

\begin{footnotesize}
\begin{enumerate}
\item[\textsuperscript{302}] “Medicine, Profession, and Society,” 59.
\item[\textsuperscript{304}] D. Black, “Cui bono?” \textit{British Medical Journal} 2 (1977): 1109-1114.
\item[\textsuperscript{306}] Carr-Saunders and Wilson, \textit{The Professions}, 300.
\end{enumerate}
\end{footnotesize}
professionals are groups of individuals who i) have special, technical/intellectual skills and ii) who provide a service deemed crucial to the public interest. iii) Arguably, professionals share a special commitment to service and iv) independence/autonomy are essential to their work. v) Professionals develop and maintain formal specialized education. vi) There is disagreement about whether formal or informal professional associations, and formal or informal codes of ethics, are necessarily the next step in defining professions. What we do know is that as a group, at a later stage of development, professionals tend to come up with their own rules and regulations to continue their two-fold task of serving their own membership as well as their clients. vii) This two-fold role-orientation is reflected in the tendency of professions to strive for social and economic closure.

Economic closure, in this context, involves controlled access to the special privileges available to professionals, and setting standards for the education and certification of practitioners to guarantee their competence, thereby excluding others from the benefits of joining the group. Social closure involves building a reputation for prestige, which alongside the professionals’ service-orientated ethos, legitimates the higher pay and autonomous control of professionals over their work. On the one hand, professional associations promote the interests of their members through social and economic closure. But, on the other hand, professionals can claim to be serving a social purpose, since by creating a reputation for themselves, and communicating their potential for social utility to the public, they provide higher level,
specialized services along with increased availability of skilled services. The two-fold (self- and other-oriented) role of professionals is finally reflected in the morality/monopoly dispute regarding professions. While professional rules and guidelines promote a conception of professions as ultimately altruistic and service-orientated, at the same time, professional associations may function anti-competitively and manipulate the market for professional services.

As I have shown so far in this chapter, definitional uncertainty about the essence of professions has persisted over the past century. A consensus does not yet exist on the traits, historical processes, or features of professions, and critics have raised a number of criticisms of the process of professionalization and the morality of the role of professionals. Before moving on to resolve some of these confusions by proposing a definition for professions, I dedicate a short chapter to discussing the notion of professional morality in functionalist writings about professions. As I will show, both proponents and critics of professions routinely fail to make a key distinction between the demands of functional professional morality, on the one hand, and the demands of all things considered morality, on the other. I argue that professional morality according to the functionalist school requires individuals to promote the social function of the profession through a set of institutional guidelines. The motivations of professionals for following these institutional rules, as well as broader considerations about general all-things-considered morality are beyond the scope of the functionalist approach. Making this distinction between functional and all-things-considered morality will in turn set the stage for discussing the moral obligations of professionals later in the dissertation.
2 Functionalist Professional Morality

As I showed in the first chapter, scholars of professions often portray traditional functionalist writings as attempts to advance a normative thesis about professions. Some, such as the early proponents of professions, take functionalism to imply that professionals are by definition altruistic. Others, such as the structuralist critics of professions, take functionalism to imply that professionals are by definition egoistic and self-serving. But as I argue below, scholars of professions (both proponents and critics) have often failed to make a distinction between the demands of functional professional morality, on the one hand, and the demands of all things considered morality, on the other.

Professional morality, according to the functionalists, requires individuals to promote the social function of the profession through a set of institutional guidelines. The motivation of professionals for following these institutional rules is beyond the scope of the functionalist approach, as are broader considerations about general morality. Avoiding motivational controversies and attending to the role of social institutions in informing our moral lives is a focal point of the functionalist school of thought. In this chapter, I clarify the notion of professional morality as it was discussed by some of the early functionalist writers about the professions, and I emphasize the difference between this understanding of professional morality and all-things-considered morality.

---

1 See, for example, Durkheim’s *Suicide: A Study in Sociology*. 
Consider the following example. A surgeon who saves a person’s life fulfills the guidelines of professional morality by providing highly skilled technical services. So long as the surgeon practices medicine in adherence with relevant professional codes of ethics, it does not matter whether he became a surgeon to become rich or to serve human-kind, and this is not a question directly addressed by his institutional codes of ethics. Nor is his relationship with his spouse, his religious or political convictions, or his private accounting practice of relevance to the performance of his professional obligations. When we say that the surgeon abides by professional morality (in the functional sense), we do not mean he is moral in the broad, all-things-considered sense of the term, but are instead making a specific statement about his moral institutional life. Nevertheless, we are making an important normative judgment about this individual. Institutions are entrenched in the fabric of our daily activities. We occupy (and are affected by) social institutional roles when we go to work, visit the doctor, buy property, do our taxes, and mail postcards, for example. Professional morality reflects on one aspect of our moral lives, in contrast with all-things-considered morality which reflects on our amalgamated moral obligations and commitments. Still, professional morality imposes fundamental normative constraints on professional roles across multiple social spectrums.

Drawing a distinction between professional morality and all-things-considered morality is important, as the two realms involve different levels of theoretic analysis. While professional morality comments on institutional obligations arising through specific social roles, general morality reflects on obligations that supersede our institutional social roles. While general morality is a core subject of inquiry by philosophers, professional morality and its significance for thinking about moral obligations in the institutionalized modern social sphere has been cast to the periphery and received scant relative attention from moral philosophers. This is not to say that no research has been done on professional morality. Switching from the general
philosophical lens to the micro-lens of the scholarship surrounding professions and their role in society, legal scholars, medical ethicists and accountants (among others) have made important contributions to this field. However, the philosophical implications of these findings remain to some extent underappreciated. For this reason, in what follows, I discuss the nature of professional morality according to some of the earliest functionalist proponents of professions. (Later, in chapter 5, I return to this subject and reconsider professional morality based on my proposed definition of professions presented in the next chapter.)

To begin, we ought not to approach this study with a naïve or overly optimistic view of the positive role of professionals in society. Exploring professional morality as a research agenda need not imply that we ignore the morally problematic connotations of professionalization. Professionals often use their power, their institutional authority and autonomy to their advantage, for example. Making the distinction between professional morality and all-things-considered morality can in fact help us more clearly identify the normatively questionable aspects of the professional structure, and to focus on whether the social and economic closure imposed by professionals is morally justifiable. Proponents of professions have sometimes argued in support of the higher remuneration and the social prestige attached to professional work, since these rewards serve as an incentive for the sacrifices made while fulfilling the technical educational requirements, the lengthy professional pre-requisites, and the certification/licensing requirements for entering professions. The normative connotations of such a trade-off, however, must be carefully evaluated. We cannot gain access to a moral analysis of professional work so long as we uniformly confuse professional morality with all-things-considered morality, as many

---

scholars of professions have done. To make a gesture towards resolving this confusion, I now provide a brief examination of the work of three early functionalist theorists of professions.

While most theorists of the professions in the 20th century have avoided a systematic discussion of background moral assumptions about the professions, Durkheim, Tawney, and Parsons do discuss the role of morality in some detail in their writings. Durkheim argues that moral rules provide specific guidelines and duties for professions. These institutional rules guide individuals away from acting in their own self-interest and towards promoting the goals of the profession as a group. Tawney also discusses morality in the context of institutional settings that ensure individuals do not self-servingly cheat the system or derive personal profit from professional appointment. Finally, Parsons argues that it does not matter whether the individual’s goals or motivations are egoistic or altruistic, since professional ethics provides a service-oriented institutional structure that guides individuals towards promoting the profession’s social goals. I discuss these three positions in turn.

Some may argue that professional moral obligations must relate, at least to some extent, to the motivations of agents and their overall commitments as moral beings. In other words, some may disagree that a distinct and discernable line can be drawn between all-things-considered morality and professional morality. In my earlier example of the surgeon, we might want to extend different moral evaluations of the surgeon depending on his reasons for entering his profession (healing patients vs. attracting a certain type of spouse). My goal here is not to deny the importance of motivations or overall moral constraints, but rather to point out a misunderstanding shared among some scholars of professions with respect to the normative assumptions underlying the functionalist description of professions. As I shall demonstrate, functionalists were not trying to comment on motivations (and more broadly on all-things-
considered morality) – in fact they were trying to get around the theoretical difficulties surrounding an overall account of morality and moral motivations.

We may decide in the end to criticize functionalist scholars for not offering an adequately comprehensive description of the moral nature of professional life. We may question their choice not to reflect directly on the importance of having the right kind of motivations for professional work. But we still should not overlook the achievement of functionalists in providing an institutional solution that promotes social utility through professional morality. This is especially so if we take seriously the limitations posed by skepticism and the difficulty of altering rigid individual motivations.

2.1 Durkheim

According to Durkheim, the “science of morals” is based on “the study of moral facts,” and the facts consist of “rules of conduct that have received sanction.”\(^3\) Here, sanction is a “consequence of the act” of a person, where the act is not taken in isolation, but rather “from the conforming or not conforming to a rule of conduct.”\(^4\) Sanctions exist in penal, moral, and civil forms, and depend on the existing relation between an act and a regulation that governs the act’s “toleration or prohibition.”\(^5\)

Two sets of rules apply to these various sanctions. The first have “universal moral application,” and guide individual moral codes as well as rules governing our relation with other

\(^4\) Ibid., 2.
\(^5\) Ibid., 2.
people generally. As Durkheim puts it, in the face of “an identical moral consciousness,” these rules derive from our intrinsic nature as human beings and do not vary from one individual to another.\(^6\) A second class of rules depends on particular qualities that are not exhibited by everyone. Following Aristotle, Durkheim suggests that these duties vary, for example, based on age, or on greater or lesser degree of kinship, but also based on one’s role in society. For example, the duties of a citizen in an aristocracy differ from those of citizens in a democracy or a monarchy.\(^7\)

Durkheim is interested in morality as a system of rules that guides a group towards its purpose. These rules, which “constitute professional ethics,” entail a plurality of duties and a “plurality of morals”:

> as professors, we have duties which are not those of merchants. [Duties] of the industrialist are quite different from those of the soldier, those of the soldier from those of the priest, and so on…. We might say in this connection that there are as many forms of morals as there are different callings, and since, in theory, each individual carries on only one calling, the result is that these different forms of morals apply to entirely different groups of individuals. These differences may even go so far as to present a clear contrast…. The scientist has the duty of developing his critical sense, of submitting his judgment to no authority other than reason; he must school himself to have an open mind. The priest or the soldier, in some respects, have a wholly different duty. Passive obedience, within prescribed limits, may for them be obligatory. It is the doctor’s duty on occasion to lie, or not to tell the truth he knows. A man of the other professions has a contrary duty.\(^8\)

Thus, professional morality is defined within the specific parameters of a profession and its role in society. As Durkheim puts it, in the professional realm, we have to be “moral particularists,”\(^9\) since a plurality of morals operates within different vocations.

---

\(^6\) Ibid., 3.
\(^7\) Ibid., 4.
\(^8\) Ibid., 4-5.
\(^9\) Ibid., 5.
For Durkheim, professional ethics has some commonality with the domestic morals of the family, has less in common with civic morals (viz. the moral relation of the individual and the state), and is entirely absent in a conception of morality which governs “the relations of men as human beings.”\textsuperscript{10} The “distinctive feature” of professional morality is thus, that since the rules govern functions performed by specific groups of individuals, not everyone is in a position to evaluate whether professional duties have been performed adequately. Therefore, the rules are not of deep concern to the common moral consciousness.\textsuperscript{11} Precisely because society as a whole is not concerned with professional ethics, Durkheim argues that it is imperative for professionals to form strong associations that oversee the creation of the moral rules, and exercise authority over their application to members.\textsuperscript{12}

A core assumption in Durkheim’s work is that individuals are inherently egoistic, and that we need systems of sanctions and rules to guide us towards cooperative social living. As he puts it, “[i]t is not possible for a social function to exist without moral discipline. Otherwise, nothing remains but individual appetites, and since they are by nature boundless and insatiable, if there is nothing to control them they will not be able to control themselves.”\textsuperscript{13} Given the expansion of industrial and economic activity in his time, Durkheim worried that professional ethics would not apply to what was shaping up to be a mainly industrial society. He was concerned that the “amoral character of economic life” would amount to “a public danger,” since the economy was guided by “no rule except that of a clear self-interest,” which left no room for “disinterestedness, selflessness or sacrifice.”\textsuperscript{14}

\textsuperscript{10} Ibid.  
\textsuperscript{11} Ibid., 6.  
\textsuperscript{12} Ibid., 7.  
\textsuperscript{13} Ibid., 11.  
\textsuperscript{14} Ibid., 12.
Durkheim saw a necessary conflict between the interests of the individual and the interests of the group. As he puts it, “[t]he interests of the individual are not those of the group he belongs to and indeed there is often a real antagonism between the one and the other.”\(^{15}\) The function of moral discipline then is to provide the individual with a code of rules “so as not to damage collective interests.”\(^{16}\) Durkheim traces the history of professions to guilds, the essence of which he believed was the reproduction of a particular moral code.\(^{17}\) He argues that both the Roman guilds and the Medieval guilds, although different in structure, existed within a moral sphere and served a moral purpose.\(^{18}\) Similarly, he saw the source of moral activity of individuals in professional groups to be “a sense of this whole which they create by close association,” and an “adherence to some thing that goes beyond the individual and to the interests of the group he belongs to.”\(^{19}\) The moral aim of individuals that form associations, similar to the guilds of the Roman city and the Medieval times, is often “to associate, for the sole pleasure of mixing with fellows, and no longer feeling lost in the midst of adversaries.”\(^{20}\) This is why, according to Durkheim “the family is the political society in miniature,” and a close parallel exists between the family and the professional group.\(^{21}\)

Note that there is no mention of the interests of the client in Durkheim’s writing. There is also no mention of the interests of the individuals outside the group (e.g. people in competing professional/non-professional bodies) whose interests might be adversely affected by the profession. The focus of professional morality is on following rules that serve the interests of the

---

15 Ibid., 14.
16 Ibid., 14.
17 Ibid., 29-30.
18 Ibid., 17-23.
19 Ibid., 24.
20 Ibid., 25.
21 Ibid., 26.
professional group as a whole. The purpose of the rules of professional morality is to defend the group against the self-serving interests of the individual professionals within it.

How might Durkheim respond to the charge that professionals legitimate their authority by using their technical expertise as an excuse for claiming social prestige and inflated remuneration? We need to draw on broader macro-considerations about all-things-considered morality to answer this question and evaluate the role of professions in society. In the meantime, and in the absence of clear instructions regarding all-things-considered morality, Durkheim’s professional morality still gives us two levels of guidance. First, on a personal micro-level, professional morality guides the actions of the individual professional in her daily work, and instructs her to promote the interests of the professional group as opposed to her self-serving interests. Second, in so far as the presence of the professional group serves a social function or a positive social role, Durkheim’s professional morality provides a meso-level moral guidance that helps individuals indirectly serve broader social interests through adhering to the set of rules that guide professional associations. This meso-level moral guidance is more explicit in Tawney’s writings on professions, and I now turn to this discussion.

2.2 Tawney

Tawney defines a profession as a “trade” organized “for the performance of function.”22 Professionals are not just a group of individuals that make a living by doing the same work. A profession is furthermore not a group organized “exclusively for the economic protection of its

---

22 Tawney, The Acquisitive Society, 92.
members,” although this is normally among its purposes. Rather, a profession “is a body of men who carry on their work in accordance with rules designed to enforce certain standards both for the better protection of its members and for the better service of the public.” Morality, in this context, becomes relevant through the professional body of rules. Tawney includes duties towards the “protection of members” as a sufficient, though not necessary, condition of professional duties. It is duties towards “the better service of the public” that constitute a necessary condition.25

According to Tawney, all professions have two sets of rules: “rules which protect the interests of the community and others which are an imposition on it.” These rules form the “essence” of professional duties, according to Tawney. These resoundingly functional rules ensure the “competence of [the profession’s] members or the quality of its wares.” The rules deliberately prohibit certain kinds of conduct, which although profitable to the individual, “are calculated to bring into disrepute the organization to which he belongs.” Therefore, according to Tawney, the essence of professional responsibility is not an altruistic relation of care for the public, but rather the promotion of the reputation of the organization, which is of superlative functional importance. Through this notion of responsibility, the organization can also more successfully fend off competitors and grow. Through the organization’s success, the professionals will enjoy increased pecuniary as well as social rewards.29

23 Ibid.
24 Ibid.
25 Ibid.
26 Ibid.
27 Ibid., 93.
28 Ibid.
29 As we will see in the next chapter, the Harsanyi principle and the Revelation principle may be used to manage and enforce such a notion of responsibility.
Two sets of rules protect the interests of the professional group. One set is essentially “trade union regulations” which are designed to “prevent the economic standards of the profession being lowered by unscrupulous competition.”\textsuperscript{30} The other set of rules ensures that no professional has “any but a purely professional interest in his work, by excluding the incentive of speculative profit.”\textsuperscript{31} This second set of rules thus attempts to prohibit deriving private gain from professional title, and underscores notions of neutrality and service-orientation as core aspects of professions. As Tawney puts it, the object of these rules is to “impose on the profession itself the obligation of maintaining the quality of the service, and to prevent its common purpose being frustrated through the undue influence of the motive of pecuniary gain upon the necessities or cupidity of the individual.”\textsuperscript{32}

Morality is fundamental to Tawney’s definition of a profession. This is evident from his distinction between the concepts of industry and profession. He notes that the essence of industry is “the financial return which it offers to its share-holders.”\textsuperscript{33} Morality in industry entails satisfying just this function. The essence of profession, on the other hand, is quite different:

though men enter it for the sake of livelihood, the measure of their success is the service which they perform, not the gains which they amass. They may, as in the case of a successful doctor, grow rich; but the meaning of their profession, both for themselves and for the public, is not that they make money but that they make health, or safety, or knowledge, or good law.\textsuperscript{34}

For Tawney, the\textit{ essence of the professions} is thus not to make money, but rather to provide a\textit{ service}. Service might be called a moral obligation in this context, since the functional

\textsuperscript{30} Tawney, \textit{The Acquisitive Society}, 93.
\textsuperscript{31} Ibid., 93.
\textsuperscript{32} Ibid., 93-94.
\textsuperscript{33} Ibid., 94.
\textsuperscript{34} Ibid.
definition of morality is the set of rules that guide groups of individuals towards the successful achievement of the group’s purpose or essence.

Tawney explains that while professionals “depend on” their profession as a source of income, they do not consider any conduct which increases their income acceptable.\(^{35}\) Hence, “while a boot-manufacturer who retires with half a million is counted to have achieved success, whether the boots which he made were of leather or brown paper, a civil servant who did the same would be impeached.”\(^{36}\) Other examples of following the ethos of service orientation include: doctors who refrain from practices which, despite earning them a large fee, would not serve the patients’ best interest; scholars and teachers who refuse to make money by deliberately deceiving the public, “as is done by makers of patent medicines, however much the public may clamor to be deceived;” judges and public servants who do not “sell justice for money;” soldiers who put service before private inclination, and even before “the reasonable preference of life to death.”\(^{37}\) Still, Tawney warns against idealizing the “professional spirit” and notes that maintaining a professional standard requires safeguards. After all, “every country has its traitors, every army its deserters, and every profession its blacklegs.”\(^{38}\) But even a standard that is often abandoned still has at its core a moral rule or guideline. As Tawney puts it:

> the meaning of a profession is that it makes the traitors the exception, not as they are in industry, the rule. It makes them the exception by upholding as the criterion of success the end for which the profession, whatever it may be, is carried on, and subordinating the inclination, appetites and ambitions of individuals to the rules of an organization which has as its object to promote the performance of function.\(^{39}\)

\(^{35}\) Ibid.
\(^{36}\) Ibid.
\(^{37}\) Ibid.
\(^{38}\) Ibid., 95.
\(^{39}\) Ibid. The kind of “industry” Tawney has in mind here sounds like the completely unregulated market economy devoid of any functional organizational rules or boundaries.
Admittedly, Tawney’s definition of the essence of professions as service makes the distinction between professional morality and all-things-considered morality less clear. For Tawney, following the rules of the profession for the most part seems to match up with all-things-considered morality in the end, “making traitors an exception”. Still, both Durkheim and Tawney provide a functional picture of professional life, with morality as the essence of professions. Both authors are keen to point out the contrast between, the “vulgarity”\textsuperscript{40} and the “boundless and insatiable” nature of individual appetites,\textsuperscript{41} on the one hand, and on the other hand, the “disinterestedness, selflessness or sacrifice” of morality, which is free of “undue influence of the motive of pecuniary gain upon the necessities or cupidity of the individual.”\textsuperscript{42}

This dichotomy between self- and other-oriented types of individual motivation is crucial to the notion of morality as the set of rules that promote the interests of the group.

I now discuss Parsons’ functionalist stance on morality and the professions. Parsons underscores that it is not self-interest, but rather the prohibition on breaking of the functional rules of the organization that is the focus of professional ethics. Given Parsons’ clarity on this issue, and given his monumental impact on the literature on professions (he is cited in almost every sociological contribution to the subject,) it is surprising that so many scholars have managed to miss the role he affords to morality in his writings about professions.

\textsuperscript{40} Ibid., 96.
\textsuperscript{41} Durkheim, \textit{Professional Ethics and Civic Morals}, 11.
\textsuperscript{42} Ibid., 12.
2.3 Parsons

Parsons introduced the functionalist school of sociology to America. He sought to demonstrate that institutionally, professions are similar to other occupational groups, and that the altruistic characterization of motivations for professions is false.

Parsons discussed the commonplace view that there is a sharp distinction between professions on the one hand, and the “modern economic order,” “capitalism,” “free enterprise,” and the “business economy” on the other. Business tends to be characterized by “the high degree of free play it gives to the pursuit of self-interest.” In contrast with business, professions are often typified by “disinterestedness.” Allegedly, professionals do not engage “in the pursuit of personal profit,” and rather focus on providing services to patients or clients, or promoting “impersonal values like the advancement of science.” It is in this context that professions are viewed as “atypical” and the remains of “mediaeval guilds,” and this is the reason some worried that they would disappear in an increasingly commercialized world.

Parsons notes that in traditional thinking about human action, “the most basic of all difference in types of human motivation has been held to be that between ‘egoistic’ and ‘altruistic’ motives.” This classification is then applied to two different spheres of activity, whereby the businessperson “egoistically” pursues her self-interest, and the professional

---

43 See Parsons “The Professions and Social Structure.”
44 Still, contributions by theorists of all stripes often wrongly claim that Parsons and functionalists somehow promoted the idea of professionals as altruists and that functionalists are responsible for the fact that “capitalist occupation” masks itself as motivated solely by self-less altruism.
46 Ibid.
47 Ibid.
48 Ibid.
49 Ibid.
50 Ibid., 36.
“altruistically” serves others’ interests regardless of her own.\textsuperscript{51} Business and profession thus are popularly taken to exemplify “the most radical cleavage conceivable in the field of human behavior.”\textsuperscript{52} Parsons blames this dichotomy between egoistic and altruistic motivations on the “predominantly economic and utilitarian orientation of thought.”\textsuperscript{53} He points out that the dominant business economy proposes that we live in an “acquisitive society” where “economic men” care only for themselves, and where professionals rise beyond “these sordid considerations, devoting their lives to ‘service’ of their fellow men.”\textsuperscript{54} The business professionals are thus expected to advertise their product. Unlike the doctor, business people are not expected to sell to customers “regardless of the probability of their being paid.”\textsuperscript{55} Conformity with the professional pattern, Parsons notes, is said to take away the opportunity to “gain financial advantages” from the professional.\textsuperscript{56}

Parsons then asks ironically, “is it not obvious that [the doctor] is ‘sacrificing’ his self-interest for the benefit of others?”\textsuperscript{57} And his answer is: no. He argues that we in fact have little assurance about the motivations of doctors and business people alike. What we do know is the institutional structure within which these individuals work. Parsons thus introduces a distinction between individual motivation and institutional structure, and argues that “the acquisitiveness of modern business is institutional rather than motivational.”\textsuperscript{58} He points out that self-interest is “seldom … decisive in motivation” for a businessperson. A businessperson’s orientation is

\begin{itemize}
\item \textsuperscript{51} Ibid.
\item \textsuperscript{52} Ibid.
\item \textsuperscript{53} Ibid.
\item \textsuperscript{54} Ibid., 43.
\item \textsuperscript{55} Ibid.
\item \textsuperscript{56} Ibid.
\item \textsuperscript{57} Ibid.
\item \textsuperscript{58} Ibid, 36.
\end{itemize}
instead toward “a total comprehensive situation extending over a considerable period of time.” 59

For this reason, the difference between the businessperson and the professional is “rather in the ‘definitions of the situation’ than in the typical motives of actors as such.” 60

Thus, functionalism becomes an important tool for thinking about professions, since it provides a theory for stipulating the situational and institutional backdrop in which individuals work. Parsons is in effect suggesting that we leave behind the dichotomy between egoism and altruism of motivations of individuals, and that we look to situations in which individuals work to evaluate the moral implications of their work. The suggestion is neither that professionals inherently share the same motivations as business people (as the revisionist critiques of functionalism began suggesting in the 1960’s), nor that professionals are self-less altruists and therefore morally superior (as many proponents of professional ethics have suggested). Instead, professional morality according to Parsons is concerned with following institutional rules as opposed to adopting altruistic or egoistic motivational attitudes.

As Parsons elaborates in his “Motivations of Economic Activity,” the codes of ethics of most professional groups imply that “the essence of professionalism consists in a series of limitations on the aggressive pursuit of self-interest,” as exemplified by the codes of medical ethics that forbid advertising of services, and require doctors to treat patients regardless of the probability of getting paid. 61 But as he clarifies, “it does not follow that, in adhering to the code as well as they do, medical men are actually acting contrary to their self-interest in a sense in

59 Ibid., 43.
60 Ibid. As Parsons explains, the phrase “definition of the situation” refers to a fundamental function of institutions that defines the situation for action (234). In this context, Parsons cites W. I. Thomas The Unadjusted Girl (Boston: Little, Brown & Co., 1927), Introduction.
which business men habitually do not." Instead, a principal component of the difference between commercialism and professionalism is institutional, rather than motivational.  

Crucially, in both business and the professions, “the self-interest of the typical individual is on the whole harnessed to keeping the institutional code which is dominant in his own occupational sphere.” For example, while it is true that by advertising and breaking the ethics codes of the profession, an individual physician could reap an immediate financial advantage. But if the medical institutional structure is working well, then it is doubtful that this behavior will “from a broader point of view” be in the professional’s personal self-interest. Acting against professional codes would draw a reaction from the physician’s professional colleagues, and also from the public, thereby proving injurious to the individual’s professional standing. This, in effect, is “the underlying control mechanism” in the professions.

To summarize, Parsons thinks the essence of business and professions are different: profit versus service-orientation. Still, he insists that just because their institutions are different, it does not follow that the motivations of individuals in business and professions are also different. In fact, both institutions strive towards objective accomplishments and a sense of recognition/reputation. Professions share core functional institutional features with business and government (rationalism, functional specificity, and universalism), although people’s misunderstanding of professionals as altruists has blurred this similarity in the past. Parsons’ conclusion in the discussion of motivations and institutions is thus as follows:

---

62 Ibid., 63.  
63 Ibid.  
64 Ibid. Italics mine.  
65 Ibid., 64.  
66 Ibid., 65.  
67 Ibid., 45-46.
The typical motivation of professional men is not in the usual sense ‘altruistic,’ nor is that of business men typically ‘egoistic.’ Indeed there is little basis for maintaining that there is any important broad difference of typical motivation in the two cases, or at least any of sufficient importance to account for the broad difference of socially expected behavior.\(^{68}\)

These statements are in sharp conflict with the common understanding of professions as synonymous with morally superior, benevolent, and altruistic behavior.

### 2.4 Some Criticisms

One criticism to be considered, with respect to the moral obligations of professionals, is that even if professionals satisfy the demands of functional morality towards the clients or customers of the profession, there remain other important demands of morality (e.g. moral obligations towards adversely impacted third parties) which need attention and which functionalism neglects.\(^{69}\) The functionalist might reply to this objection as follows. The demands of professional morality are not exhaustive. We have personal or civic moral rules to abide by at the same time as our professional roles, and depending on the situation, the latter might be trumped by the former. While functional moral rules guide the professional, this does not mean that professionals should forget about their non-functional duties. The core question that remains unanswered here is: how do we determine the threshold past which one set of moral demands trumps or supersedes the other?

A second core criticism of functionalist professions is that professional economic closure dampens competition, and creates economic monopolies. Economic closure is morally suspect, given the emphasis that professions place on service ethics. Finally, a third core structuralist

---

\(^{68}\) Ibid.  
\(^{69}\) See Greenstein, “Against Professionalism.”
criticism of the functionalists is that from the social perspective, professional social closure is a classist, elitist movement, and that again, there is something morally problematic here given the functionalist emphasis on serving social order and supporting social solidarity. The problem with these criticisms is that they often assume that service ethics is in fact an all-things-considered ethics. But this need not be the case. Instead, the functionalist service ethic is a very specific set of rules that aims to ensure that professionals promote the interests of the professional group whose function is to then provide a service.

Clarifying the functional notion of morality helps resolve some of the confusions surrounding the definition of “profession.” Professions have been praised and criticized, as inherently altruistic on the one hand, and self-consciously opportunistic on the other. But as the traditional functional theorists of professions would point out, professional ethics in fact involves a set of institutional rules that guide professions towards the goal of service-orientation. The essence of professionalization in this context would be to create appropriate professional ethics rules for the function of the specific professions. But how might we determine “appropriateness”? In the background of most classic functional contributions to the literature is the idea that the appropriate role of professions is to provide positive social utility or service.

As Adam Smith noted:

> We trust our health to the physician … our fortune and sometimes our life and reputation to the lawyer and attorney. Such confidence could not safely be reposed in people of a very mean or low condition. Their reward must be such, therefore, as may give them that rank in the society which so important a trust requires.”

Some degree of social and economic closure might thus be necessary insofar as it rewards professionals in return for their dedication to service. Such a set-up might even make

---

professionals immune to corruption to some extent by paying them well enough. We need to
determine the threshold past which the pursuit of economic and social power hinders the
professionals’ capacity to provide social utility. We might, on the other hand, decide that such a
threshold is impossible to reach and that professional work is necessarily socially injurious
despite the claims of functionalist scholars. In the next chapter, I continue this debate over the
socially beneficial nature of professions through a discussion of economics and the role of
information asymmetries in the professional service market.
3 Information Asymmetries, Regulation, and the Professions

In chapter one, I noted that a survey of attempts to define the professions reveals uncertainty and confusion in the literature. Given the deficit of philosophical contributions to this topic, I have so far explored perspectives from sociology and political economy in my survey of the various definitions and stipulations of the fundamental features of professions. In this chapter, I turn my attention to an economic analysis of the professions and draw on the literature in information economics to explore the essence of professionalism.

The professional institutional structure is controlled, protected, and implemented through professional regulations. Professionals draw on regulations (in the form of registration, certification, and licensure) to appeal to the authority of the state in order to define the function and mechanism of professional associations. Going back to Weber’s discussion of various forms of regulation of the market, regulation may refer to self-regulation or private regulation by professional associations, and also to governmental or public regulation of the professions. Two main lines of argument populate the literature on the nature of professional regulation. On the one hand, proponents of “public interest theory” argue that regulations aim to promote and facilitate professional contributions to the public interest in the face of market failures. On the other hand, supporters of “private interest theory” argue that regulations do not promote the public interest, and are, in fact, merely a decoy for enhancing the benefits accrued by

---

professionals through professional membership. Mirroring the morality versus monopoly debate discussed in chapter 1, at the heart of the debate over professional regulation is a dispute over normative models and the implicit morality of the professions. Professions and the professional regulations that govern them are designed to be beneficial for the public interest, according to some, and to be detrimental to the public interest, according to others. This is why analyzing the dispute between public and private interest theories of regulation is crucial for my study of the nature of professions.

This chapter is divided into two parts. In Part I, I discuss the nature of professional regulation, and provide an overview of the public and private interest theories of regulation. As I will show, some private interest theorists of regulation deny the existence (and the problematic nature) of market failures. I evaluate the claims of these critics in Part II, where I provide a basic introduction to information asymmetries and discuss some common game theoretic solutions to moral hazard and adverse selection problems. As I will show, on the one hand, evidence suggests that professions serve a social function by reducing inefficiencies and by creating the conditions of trust necessary for transactions that would otherwise fail to take place. A profession’s success in fulfilling these functions depends on how well its regulatory design responds to information asymmetries. On the other hand, professionalization seems to allow professionals to take advantage of the conditions of market failure in order to maximize rents and professional gains at the expense of the public. Here, professional regulation is an instrument for anti-competitiveness and for the promotion of the interests of members of professional associations. Thus, a descriptive understanding of professional regulations reveals that they may serve a social or an anti-social function.

2 Rents are returns in excess of the professional’s opportunity costs. Opportunity costs are the cost of a product measured in terms of the value of its best alternative (when that alternative is not chosen).
Recall from the introduction, that my task is to explore a normative standard for professionalization in the first few chapters of the dissertation. Such a standard cannot be merely prescriptive, and has to be informed by descriptive features of professions. In this chapter, continuing on my work in chapter 1, I develop an empirically-informed understanding of what professionals do and how professional regulation works. This understanding in turn lays the groundwork for a normative model of professions, developed in the next chapter, which provides a reconstructive account of the normative purposes already implicit in the practices of professionals and the regulations that guide them.

3.1 Professional Regulation

By discussing professional regulations in Part I of this chapter, I in effect continue the analysis that started in chapter 1 about the nature of professions and their normative role in society. I have explored perspectives from sociology, law, and political science in my survey of the various definitions and stipulations of the fundamental features of professions. In this chapter, I turn my attention to an economic analysis of the professions.

The economic analysis of professions necessarily involves discussing regulations. For many theorists, the distinguishing factor between professions and other occupations is the presence of regulations. Professional regulations may be thought about as a set of (formal and informal) rules and norms for governing professional social and economic interactions. These

---

3 The term “regulation” is used differently across various disciplines. Professional regulations are often defined more descriptively as a set of statutes (in the form of registration, certification, and licensure) through which professionals appeal to the authority of the state in order to define the function and mechanism of professional association.
rules outline the institutional structure that overlooks professional activity. An economic study of
the nature of professions and the essence of professionalization can benefit from a study of
professional rules and regulations. As Decker and Yarrow have explained, “[i]t is a central
proposal of economics that market rules exert a considerable influence on the behavior of
market participants and on economic performance.” The rules of the game in game theory are
similarly a key factor for determining the conduct of players, which in turn impacts the outcomes
of games.

The link between market rules and market performance has also been discussed in the
literature on industrial organization, where the market structure encompasses barriers to entry,
concentration among suppliers, and the general regulatory structure governing the market. In the
economic literature on the study of regulation, the emphasis is on exploring the impact of
different rules and regulations on market performance. What role or social function
professionals play in society is thus a downstream consequence of the kind of regulatory norms
they operate under. This is why a study of the essence of professionalism and the social role of
professionals would not be complete without an analysis of the kinds of rules and regulations
that govern the work of professionals.

We may distinguish between three levels of regulatory control in the professions:
registration, certification, and licensing. The first level, registration, requires individuals to list

---

4 Christopher Decker and George Yarrow, “Understanding the Economic Rationale for Legal Services Regulation,” Regulatory Policy Institute Report for the UK Legal Services Board (2010), 5
their names in an official register as a condition for engaging in certain activities.\(^8\) For example, The Canadian College of Professional Counsellors and Psychotherapists is a regulatory body that overviews the registration system for counsellors and psychotherapists practicing in Canada. The second level, certification, requires involvement of a government agency to certify that individuals possess certain skills, although uncertified individuals are not prevented from practicing. For example, in most states in the United States, both certified and uncertified accountants are allowed to practice, but only certified accountants can use the title CPA (Certified Public Accountant).\(^9\) The third level, licensing, requires that individuals obtain a license from a recognized authority in order to engage in an occupation. Not only does the license demonstrate competence, the passing of required tests, and the fulfillment of educational requirements, it also subjects non-licensed practitioners to fines and/or jail sentences.\(^10\) For example, physicians and surgeons need a state sanctioned license to practice medicine in Canada and the United States.

What kind of justification do proponents of these regulations provide? Registration may be justified as a device for taxation, as a standard setting device for academic training and professional experience, or as a control mechanism for fighting fraud.\(^11\) In comparison to registration, justifying certification and licensure has proved more controversial. Critics argue that certification is something the private market generally can do for itself through the use of private certification agencies that certify the competence of service providers.\(^12\) Other avenues for certifying the quality of a product (or service) include private certification arrangements,

---

\(^9\) Ibid.
\(^10\) Ibid., 145.
\(^12\) Friedman, *Capitalism and Freedom*, 146.
private industrial testing laboratories, consumer testing agencies, and Better Business Bureaus, alongside educational degrees, brands, and retail chains.\textsuperscript{13}

In contrast with certification, which communicates the completion of certain prerequisites and requirements for holding a particular professional title, entry into a licensed profession, and transactions among buyers and sellers of licensed services, require the permission of the state.\textsuperscript{14} From a legal perspective, professional licensing statutes outline the tasks and functions of the occupation, prescribe that these functions may not be legally performed except by licensed practitioners, and describe the procedures for acquiring a license.\textsuperscript{15} Licensing statutes delineate the kind of behavior and activitieslicensees may engage in, and legally forbid all non-licensed persons from engaging in these activities.

There are, broadly speaking, two sets of theories regarding professional regulation: the public and private interest theories of regulation. The debate between these two camps is of pivotal importance to the development and operation of contemporary professions, and numerous authors have contributed to this literature in the past fifty years.\textsuperscript{16} As we shall see, private

\textsuperscript{13} Ibid.
\textsuperscript{14} Rottenberg, Occupational Licensure and Regulation, 2.
\textsuperscript{15} Ibid.
interest theorists worry that regulations impose anticompetitive restrictions on the market for professional services, despite arguments by public interest theorists that justify regulations in the name of the public interest. In both cases, professionals draw upon the state’s authority in order to enforce regulations. This is the case even when a profession self-regulates, since a self-regulated profession is governed in part by itself and in part by the government. The regulatory structures in the UK, North America, and Europe, reflect different stances on these theories of regulation. I begin by providing a brief summary of the public interest theory of professional regulation.

### 3.1.1 The Public Interest Theory of Professional Regulation

In general, professional service regulation is subject to two kinds of classic economic concerns about efficiency and equity. With respect to equity, the main issue is the distribution of economic resources. Access and affordability are key problems, and there are also concerns with issues of fairness such as protection against abuses of power. These issues are important, and as I showed in chapter 1, they have been well documented by structuralist critics (discussed

---

17 In 2003, the Director General for Competition at the European Commission began an investigation into the effects of regulation in the professions. The main task of the investigation was to explore the impact of professionalization on competitiveness, and to discover whether current regulation serves the interests of consumers. See European Commission, Competition and Professions: http://ec.europa.eu/comm/competition/sectors/professional_services/overview_en.html (Accessed 28 September 2011).


earlier in chapter 1). My focus in this chapter is on efficiency. The general public policy concern about efficiency has to do with maximizing the value of total production or service. Information asymmetries affect quality of service, broadly construed, and the structure of supply of professional services can directly impact service quality and professional competition. As I show below, responding to these concerns about the efficiency of professional service is a key rationale for the public interest theory of regulation. Ultimately, I shall argue that from a normative perspective, rules that govern professions with the goal of increasing relative efficiency help fulfill the social purpose of professions as trust-preserving institutions that facilitate mutual benefit in society.

Part of the first fundamental theorem of welfare economics states that, given certain strict assumptions, perfect competition will lead to a Pareto-optimal equilibrium. Such an equilibrium is efficient because it is impossible to make one person better off without making someone else worse off. When the conditions of perfect competition change, however, economic efficiency is reduced. This economic inefficiency is called a market failure. According to the public interest theory of regulation, government regulations can remedy market failures and raise social welfare. In the context of professions, the public interest is often invoked as the main impetus for professional regulation.

---

20 Perfect competition refers to a market structure where i) there is an abundance of sellers of a product; ii) no one seller has control over market prices; iii) sellers supply identical products; iv) buyers have perfect information; v) there are no costs to transacting; vi) the same resources and technology are available to everyone; vii) there are no barriers to entry or exit; and also viii) no externalities in production or consumption.


23 See Posner, “Theories of Economic Regulation.” For further discussions, see McGraw “Regulation in America: A Review Article,” and Stigler, “The Optimum Enforcement of Laws.”

Government regulation of economic sectors typically takes place for three reasons: (i) to address market structure imperfections, such as the presence of a natural monopoly; (ii) to deal with externalities (uncompensated economic effects on third parties) which regulators hope to eliminate or internalize; and (iii) to influence imperfect information markets (involving, say, a lack of information), which may lead purchasers to make non-optimal decisions that could have been avoided in the presence of more information. I explain each of these three scenarios in turn below. Depending on the kind of market failure, intervention might involve changing liability rules, taxes, subsidies, or some form of regulation. Of course, government failures (in contrast with market failures) occur regularly, and we need to be attentive to the limits to government regulations as well. For now, the question is: why do market failures persist in the absence of professional regulations? In this context, I briefly discuss (i) market structure imperfections, (ii) externalities, and (iii) imperfect information markets in turn below.

(i). The “market structure” refers to a number of service providers that provide identical, homogeneous services. A market structure imperfection occurs when one firm has control over certain products or services in the market. In a natural monopoly, for example, one firm can efficiently serve the entire market demand at a cost that is lower than any two or more smaller firms. In the case of professions, a cluster of service providers take the place of the solo dominant firm, and thus the correct terminology is a “cartel” as opposed to a “monopoly.” In comparison to a monopoly, a cartel is made up of a group of firms offering the same services, and exercising control in the market. Professions are cartels since although each member of a professional organization enters into transactions with various clients, no single member is large

---

enough to qualify for the title of monopoly. Aside from cartels, market structure imperfections in the professional market arise, to some extent, because of the heterogeneity of talents of would-be professionals.\textsuperscript{27} In any group of individuals, the most talented and promising tend to have the best remunerative alternatives and the highest reservation prices (i.e. the cost of their services is the highest). Thus, without regulation, a professional market might attract a large pool of un-/under-qualified entrants, especially if other industries or professions affect rents in such a way as to lure the talented away.\textsuperscript{28}

(ii). Externalities in this context are the broad social welfare losses that result from inadequately rendered professional services. For example, when an ailing education system leads to deficiencies among educators and produces students who operate below their potential, externalities are posed to the whole society. The same is true when an innocent person is convicted due to sub-par legal procedures.\textsuperscript{29} The problem of externalities is especially prevalent in professions such as engineering, architecture, accounting, and auditing, where a mistake in construction, design, or calculation can have serious negative effects on society at large.\textsuperscript{30} If transaction costs\textsuperscript{31} are low, we can internalize the externality through adding a risk premium to the price of services. The problem is that transaction costs are often too high in the market for professional services. Taxes are another solution to negative externalities, whereby the external costs are charged in the form of a tax.\textsuperscript{32}

\textsuperscript{27} Ibid., 7.
\textsuperscript{28} Ibid.
\textsuperscript{29} Ibid.
\textsuperscript{30} Philipsen, “Regulation of Liberal Professions and Competition Policy,” 206.
\textsuperscript{31} Transaction costs are the costs incurred in making an economic exchange. A risk premium is the return in excess of the risk-free rate of return that a professional service is expected to yield. See the Coase Theorem: Coase, “The Problem of Social Cost,” 1.
\textsuperscript{32} Philipsen, “Regulation of Liberal Professions and Competition Policy,” 206.
(iii). Information asymmetries affect professional services since most clients lack the highly specialized training to adequately evaluate the nature of services rendered by professionals. Because of this lack of expertise, purchasers have difficulty judging service quality with confidence, and an air of uncertainty surrounds the layperson as she enters the professional-service market. But efficiency losses occur not just because clients are unable to evaluate professional services. Crucially, efficiency losses also occur when the lack of knowledge incurs dead-weight losses that arise from lost transactions between clients and service providers, who might have come together if it were not for the lacking environment of trust that surrounds information asymmetries. Similar to the problem of imperfect market structure, externalities and the lack of information stems to a large extent from the heterogeneity of the professional service-providers. Given the extensive impact of the heterogeneity of talents problem, a core goal of professional regulation, according to public interest theorists, is “the provision of homogeneous services, coupled with the assurance of some minimum acceptable target level of professional competence.”

Regulators typically deal with market imperfections by controlling supplier entry and also the conditions of sale, which in turn grants them a level of control over prices. Externalities are usually mitigated through regulations that influence the kind of products and services on offer, along with their production processes. While economists prefer to use taxes and subsidies, regulators often ban a particular service or product outright instead. As for imperfect

---

34 Ibid., 8.
35 I discuss dead-weight losses and information asymmetries in more detail later in this chapter.
37 Ibid.
38 Ibid., 6.
39 Ibid.
information markets, regulators tend to resort to government-induced advertising or labeling campaigns. Although these efforts bring about sporadically increased public awareness of the information issues, they often fail to alter buyer or seller behavior.\footnote{Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 7.} In these respects, regulation is promoted in the name of the public interest.

Still, the presence of information asymmetries and externalities need not always signal an opportunity for improvement through public regulation. Jumping straight from the presence of market failure to a conclusion about regulatory intervention is not necessarily warranted. It may in fact be more costly to eliminate market failures than to leave them be – some information asymmetries and externalities are arguably a necessary feature of markets. Eliminating information asymmetries may mean abolishing all divisions of labor in market activities, and fighting externalities at any cost can be quite inefficient. Instead, the appropriate guidance provided by regulations reveals more/less efficient ways of governing professions and maps out how institutions and organizations may be developed so as to improve efficiency in the face of the division of knowledge among different parties and the costs of market transactions.\footnote{See Coase, “The Nature of the Firm” and Coase “The Problem of Social Cost.” See also H. Demsetz, “Information and Efficiency: Another Viewpoint,” \textit{Journal of Law and Economics} 12 (1969): 1-22.}

A considerable body of empirical work points to the severe limitations of public interest theory for understanding how regulations work in practice.\footnote{See, for example, the survey of different regulatory contexts by Joskow and Noll, who argue that the positive findings of public interest theories are for the most part inconsistent with available evidence: Paul L. Joskow and Roger G. Noll, “Regulation in Theory and Practice: An Overview,” in ed. Gary Fromm, \textit{Studies in Public Regulation} (MIT Press, 1981):1-65.} I survey some of these limitations in my discussion of private interest theory below.
3.1.2 The Private Interest Theory of Professional Regulation

The private interest approach argues against the purported merits of professional regulation, and was developed based on public choice theory, capture theory, and the “Chicago” theory of regulation.44 In general, proponents of the private interest theory of regulations argue against allowing the state to influence professional activity. Friedman viewed professional licensure as a special case of a more general and widespread phenomenon, “namely, edicts that individuals may not engage in particular economic activities except under conditions laid down by a constituted authority of the state.”45 Friedman noted that the medieval guild system was the ultimate realization of licensure at the hands of the state, and that the overthrow of the guilds was “a sign of the triumph of liberal ideas” and “an indispensable early step in the rise of freedom in the Western world.”46 He drew extensively on the rhetorically charged, critical writings of Walter Gellhorn, which he called the best brief survey of professions.47 Consider the following excerpt:

As long ago as 1938 a single state, North Carolina, had extended its law to 60 occupations. One may not be surprised to learn that pharmacists, accountants, and dentists have been reached by state law as have sanitarians and psychologists, assayers and architects, veterinarians and librarians. But with what joy of discovery does one learn about the licensing of threshing machine operators and dealers in scrap tobacco? What of egg graders and guide dog trainers, pest controllers and yacht salesmen, tree surgeons and well diggers, tile layers and potato growers? And what of the hypertrichologists who are licensed in Connecticut, where they remove excessive and unsightly hair with the solemnity appropriate to their high sounding title?48

45 Friedman, Capitalism and Freedom, 137.
47 Friedman, Capitalism and Freedom, 139.
Thus, the control of professions by state law was frowned upon by Friedman and Gellhorn. As Stigler puts it, the problem of regulation in this context involves discovering when and why a profession can “sue the state for its purposes” in an effort to boost its own profitability.\textsuperscript{49} I discuss some of the arguments in support of this stance below.

\textit{Market Allocation Mechanism \& Anti-Competitiveness}

To discredit state regulation, proponents of the private interest theory point to market allocation mechanisms and anti-competitiveness as first-order and second-order impacts of professionalization. In order to analyze these impacts, we first need a more extensive understanding of professions from a market perspective.

The standard definitions of professions offered in the economic literature are strictly “supply-oriented.”\textsuperscript{50} That is, professions are often defined in terms of the kind of services they provide, and the role which professionals play as suppliers. An industry (or a profession) is often defined as “an occupation or vocation requiring training in the liberal arts or the sciences and advanced study in a specialized field;” a professional is in turn “one who has an assured competence in a particular field or occupation.”\textsuperscript{51} But this class of definitions fails to recognize the demand for professions and professional services.\textsuperscript{52} We need to know whether individuals who demand a particular kind of service at a given price, would substitute that service for

\textsuperscript{50} Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 4.
\textsuperscript{51} The American Heritage Dictionary of the English Language.
\textsuperscript{52} Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 4. Demand elasticity is a measurement of the quantity of a service demanded, given a change in its price. Cross-elasticity of demand measures the responsiveness of the demand for one product to changes in the price of another product. Supply elasticity is the measure of the responsiveness the quantity supplied of a service, when the price of that service changes. Cross-elasticity of supply measures the responsiveness of the supply of one product to changes in the price of another product.
another kind of service in the case of a rise in service prices. In other words, we need to know whether the service is “cross-elastic in demand with respect to other services.”\textsuperscript{53} We also need to know whether individuals who do not yet belong to a specific profession, are ready within short notice to provide the services of a given profession when demand for professional services rises.\textsuperscript{54} In other words, we need to know whether service is “cross-elastic in supply with respect to other services.”\textsuperscript{55} Standard definitions of professions tend to leave out information about cross-elasticity in supply and demand. But knowledge about the availability of alternative sellers and buyers helps distinguish between a professional service market and the service market of other occupations.

For example, the service provided by taxi-drivers is cross-elastic in demand, since demand for taking taxis drops in response to a drop in the cost of owning and operating personal vehicles by the public. The service of taxi drivers is also cross-elastic in supply, since supply of taxi-drivers drops in response to a drop in the cost of public and private transportation. In contrast with the services of taxi-drivers, the services of physicians are usually not cross-elastic in demand. We do not witness an overall rise in the number of patients in doctors’ offices based on the price of services of alternative medicine practitioners, for example. The services of physicians are also generally not cross-elastic in supply, since we do not see a change in the number of practicing physicians based on changes to the cost of alternative medicine. A patient’s belief about medicine, more than the price of medicine, seems to impact the demand and supply of medical services.

\textsuperscript{54} Ibid.
\textsuperscript{55} Ibid.
But why do people have certain beliefs about the services of traditional medical practitioners? How are these beliefs formed? Paying attention to the demand characteristics of professional services helps answer some of these questions. The high level of specialization of professionals, their lengthy period of formal training, and their required practical experience in the field provides the “perception” among prospective clients that professional services are not easily replaced by non-professionals.\(^{56}\)

Market demand curves for professional services thus tend to be relatively price-inelastic, due to a perceived lack of substitutes and the mandatory nature of services offered.\(^{57}\) Let me elaborate. Price-inelasticity of demand implies that the quantity of professional services demanded is not responsive (or elastic) in response to a change in the price of professional services. The concept of perception is of utmost importance in this context, since purchasers of professional services ordinarily cannot do without those services (the services are “mandatory”: e.g. healthcare), and yet few clients can evaluate specialized professional services and their results.\(^{58}\) This means that the distinction between different professional service markets are to some extent “illusory”; professionalization thus acts as “an effective market allocation mechanism” by allowing for the possibility of narrow distinctions within the service market.\(^{59}\) It is naturally in the interest of the professions in question to maintain such illusions. This is the “first-order” impact of professionalization.\(^{60}\) Since professionals are not a homogeneous group, and since talents cannot be transferred or shared, purchasers are confronted by “differentiated professional services… whether real or imagined.”\(^{61}\) For example, the distinction between a

\(^{56}\) Ibid.

\(^{57}\) Ibid., 6.

\(^{58}\) Ibid.

\(^{59}\) Ibid.

\(^{60}\) Ibid., 5.

\(^{61}\) Ibid.
chiropractor and an orthopedist is not just in the difference in their training experience, nor in each group’s claim to differentiation from the other. Rather, the two professionals ultimately enjoy separate service markets because their patients view their services as imperfect substitutes.

Beyond serving as a market allocation mechanism, a “second-order” impact of professionalization is that it restricts competition within each narrowly defined service market. Regulation contributes to both the market allocation mechanism and the restricted competition effects of professionalization. But how might professionalization restrict competition? This is a long-standing problem in the history of professions. Smith warned about the impact of professional regulation on competitiveness as follows:

…the policy of Europe occasions a very important inequality in the whole of the advantage and disadvantage of the different employments of labour … [first] by restraining the competition in some employments to a smaller number than might otherwise be disposed to enter into them. The exclusive privileges of corporations are the principal means it makes use of for this purpose … [which] … necessarily restrains the competition … to those who are free of the trade. To have served an apprenticeship … is commonly the necessary requisite for obtaining this freedom.

Similarly, Mill observed the impact of professionalization on competition and in turn on wages:

… there are kinds of labour of which the wages are fixed by custom, and not by competition. Such are the fees or charges of professional persons: of physicians, surgeons, barristers, and even attorneys. These, as a general rule, do not vary, and though competition operated upon those classes as much as upon any others, it is by dividing the business, not, in general, by diminishing the rate at which it is paid.

---

62 Ibid.
63 Ibid., 6.
Competition can happen on at least three levels in the professions. Consider the example of physicians. A physician may compete with other physicians in terms of excellence, innovation, reputation, etc. She may also compete with members of other professions, e.g. alternative medicine practitioners who belong to separate professional bodies, but who provide similar services. Finally the physician may compete with non-professional groups, e.g. yoga medical practitioners (who have not yet professionalized) in providing services to clients. All three of these levels of competition are affected by the regulatory bodies governing the professionalization of physicians. As we will see below, critics see the regulatory boards of professions as mechanisms for restricting competition.66 Common strategies among these boards include “prohibiting competitive bidding, price and service advertising, and other forms of direct and indirect price and service competition.”67

Market allocation mechanism and anti-competitiveness (the first and second-order characteristics of the professions) lie in the background of most criticisms raised by private interest theorists of professional regulation. I discuss a number of these criticisms below. Through these arguments, private interest theorists have shown that professional regulations can be self-serving, and designed to promote benefits for professionals.68

Restrictions to Entry & Cartelization

What kinds of regulatory policies might professions seek from the state? Control over entry, the power to affect substitutes and complements, price-fixing policies, and direct monetary

66 Blair and Rubin, Regulating the Professions, vii.
67 Ibid.
subsidies are some common requests. Control over entry and price-fixing are particularly striking in this context. Three popular strategies for controlling entry to the professions are restricting admissions to professional schools, licensing examinations (supervised by the profession or the state), and advertising restrictions. As Benham argues, professional licensure is “characterized by controls over entry.” The entry barriers create “windfall gains (rents),” making the prospective rents (and the threat of their loss) a major impetus for licensure. Regulatory bodies in the professions tend not to use price-fixing policies, but they do discourage price competition. For example, standard professional fees are charged for well-defined services, and advertising is prohibited in certain professions. But this does not mean that professionals all charge the same price, or that purchasers receive uniform services for the same price.

Thus, a common conclusion among private interest theorists is that professional licensing is a device for limiting entry into occupations, “thereby serving as the first step towards cartelization.” The cartelization of professional services involves creating professional associations made up of service providers in an effort to exert restrictive influence on the market for professional services. Common policies among cartelized service providers are market allocation mechanisms and regulated prices. (Note that there is a distinction between cartels and monopolies, since although each member of a professional organization enters into

72 Ibid.
73 Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 8. For example, physicians and surgeons are not allowed to advertise in Canada, but lawyers are.
transactions with various clients, no single member is large enough to qualify for the title of monopoly.\(^76\)

Why might professionals seek to cartelize? According to Becker, the gain to firms from colluding is “positively related to the elasticity of their marginal cost curves” and is “inversely related to the elasticity of their collective demand curve.”\(^77\) Since professionals offer their services on a units-of-time basis, they offer a fixed number of service units. This means that professions have a “perfectly elastic marginal cost curve until or about the capacity point and a perfectly inelastic marginal cost curve thereafter.”\(^78\) That is, each professional chooses a minimum price, or “rate”, below which services are withdrawn, and at which price all clients are served, to the extent that the professional’s time allows. Professions further have inelastic demand curves since, as I explained earlier with the example of taxi drivers versus doctors, demand for professional services does not shift (is not responsive/elastic) in response to changes in the price of services. Thus, given their elastic individual marginal cost curves and their inelastic demand curves, professionals have a strong financial incentive to collude.\(^79\)

Cartelized markets also typically exhibit higher costs or slower entry to the market, both of which are attractive features of the professions for their existing members.\(^80\) Cartels are permitted to exist so long as the public and the government view the professional cartel as providing beneficial services at fair prices.\(^81\) The problem is that although, in theory, cartel members are supposed to have equal abilities, in fact, their services are usually heterogeneous

\(^{76}\) Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 9
\(^{79}\) Ibid.
\(^{80}\) Posner, “Theories of Economic Regulation,” 344.
and difficult to measure.\textsuperscript{82} Attempts to set up and strengthen cartels by the professions include campaigns to secure the initial passage of a licensing statute for an unlicensed occupation, placing constraints on joining existing licensed occupations, striking down exemptions that permit unlicensed practitioners to operate, and broadening the scope of professional practices assigned to licensed individuals.\textsuperscript{83}

\textit{The Power of Special Interests}

A related problem with professionalization is that it arguably allows the state to assign power to members of professions through regulation.\textsuperscript{84} Professional competence may thus become less relevant in such a system, while “the personalities of the members of the licensing board and the mood of the time” may determine decision-making.\textsuperscript{85} Friedman dismisses justifications of professional regulation that appeal to guarding the public interest, and draws attention to the power of producer groups that tend to be more politically concentrated than consumer groups. As he explains, “the public interest is widely dispersed. In consequence, in the absence of any general arrangements to offset the pressure of special interests, producer groups will invariably have a much stronger influence on legislative action and the powers that be than will the diverse, widely spread consumer interest. . . .”\textsuperscript{86} So although regulated policies claim to be in the public interest, they can have the effect of “curbing competition and promoting the interests of the [regulated] bodies.”\textsuperscript{87}

\begin{flushleft}
\textsuperscript{82} Ibid., 11
\textsuperscript{83} Rottenberg, \textit{Occupational Licensure and Regulation}, 5.
\textsuperscript{84} Friedman, \textit{Capitalism and Freedom}, 141.
\textsuperscript{85} Ibid., 141.
\textsuperscript{86} Ibid., 143.
\textsuperscript{87} Ibid.
\end{flushleft}
This effect is described by the capture theory of regulation, according to which regulation is created in response to demands of various interest groups that compete with each other to maximize the respective incomes of their members. Arguably, regulatory boards and bodies seek these policies, by “utilizing the professional group’s influence and political clout to bend the public will.” According to this line of reasoning, regulation comes about as a result of competition for political influence among different interest groups. The strong influence of interest groups on politicians might override the public interest, since professional associations are smaller, better organized, and single-issue-oriented in contrast with the general public. The lobbying efforts of professional interest groups are often successful since the public’s information costs are too high to evaluate the rent-seeking behavior of the professions. Interest groups thus exhibit rent-seeking behavior and seek to influence political decisions, often at the expense of social welfare. Professional regulations incur a social cost, because they inevitably create a “monopoly” position for one group at the expense of the rest of the public. Echoes of this criticism can be heard throughout the private interest theory of regulation literature.

Equally problematic is the fact that licensure assigns an extensive level of control to the current members of the licensed occupation. Typically, state licensure entails the creation of regulatory boards, which make decisions regarding entry-requirements (educational or otherwise), and establish codes of conduct. Boards consist primarily, and often exclusively, of members of the licensed profession. The worry is that members of these boards may, overtly or

---

90 Philipson, “Regulation of Liberal Professions and Competition Policy,” 207.
92 Philipson, “Regulation of Liberal Professions and Competition Policy,” 207.
93 Friedman, Capitalism and Freedom, 148.
94 Consider for example Rottenberg, Benham, and Blair and Rubin, discussed earlier.
not, substitute the public interest with their economic self-interest in their decision-making.\textsuperscript{95} As Blair and Rubin put it “many of the boards have become –if they were not from the outset – essentially insulators, whose basic purpose is to shelter those regulated from the demands (and the incentives) of competition.”\textsuperscript{96} Even when board members are not members of the occupation that is seeking licensure, they are often appointed by those occupational associations.\textsuperscript{97} The economic interests of professional board members may also influence decisions about admission requirements and definitions of licensure standards.

Ultimately, as Horowitz puts it, we’d “have to be more than a little naïve to overlook the self-serving aspects of many of the goals of these self-regulating bodies.”\textsuperscript{98} A pivotal problem for regulations and policies that seek to remedy market failures in the professions is that they often create the illusion of homogeneous service, as opposed to actually attending to the problem of heterogeneity of service providers. By perpetuating this illusion of homogeneity, self-regulatory bodies establish rules that in effect serve the financial interests of members of the profession, through “creating additional market imperfections and externalities, and inhibiting the diminution of uncertainty and imperfect information.”\textsuperscript{99} How do professional regulations create “additional market imperfections?” I discuss this effect in the next subsection.

\textit{Inefficiency}

Private interest theorists worry about the prevalence of inefficiencies caused after the creation of professional regulations, which purported to resolve inefficiencies in the market for

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{95} Blair and Rubin, \textit{Regulating the Professions}, vii.
\item \textsuperscript{96} Ibid.
\item \textsuperscript{97} Gelhorn, \textit{Individual Freedom and Government Restraints}, 140-1, and Friedman, \textit{Capitalism and Freedom}, 140.
\item \textsuperscript{98} Horowitz, “The Economic Foundations of Self-Regulation in the Professions,” 7.
\item \textsuperscript{99} Ibid., 8.
\end{itemize}
\end{footnotesize}
professional services in the first place. Instead of having a positive impact, regulations arguably limit the services offered by professionals, and at times make it difficult for specialists to communicate information effectively to consumers and potential clients.\textsuperscript{100}

Thus, professional restrictions might dissipate rents. Consider for example, the competition among entrants into professions, which imposes further costs prior to entry into the profession, or the time and training requirement during the period of apprenticeship after entry into the profession.\textsuperscript{101} Even if licensing was an attempt towards cartelization, as some have argued, these attempts fail to a large extent, since “licensure in the professions differs in several respects from the most efficient form of cartelization.”\textsuperscript{102}

Of course, inefficiency and its consequences for social welfare are not the only variable to consider in evaluating professional regulation. Still, inefficiency of professional regulations is problematic given the fact that many regulations are justified in the first place because they seek to improve efficiency. The question here is whether the inefficiencies given rise to by professional regulations are more extensive than the pre-existing efficiencies that the regulations attempted to remedy, and to what extent the new inefficiencies relate to (or offset) the original inefficiencies.

\textit{Quality Maintenance & Legitimization}

Another interesting feature of professional regulation involves spending considerable resources on demonstrating the public benefits of regulation. There is widespread evidence that

\textsuperscript{100} Blair and Rubin, \textit{Regulating the Professions}, vii.
\textsuperscript{101} Weingast, “Physicians, DNA Research Scientists, and the Market for Lemons,” 83.
professionals would like their regulation to appear “under the guise of maintenance of quality.” Historical examples include such efforts by goldsmiths, plumbers, barbers, horseshoers, undertakers, and optometrists. Successful professionalization is thus closely related to creating the public impression that professionals provide pivotal services well. Some professional groups are more successful than others in securing regulatory control, because of their ability for undertaking activities and expending resources that persuade their own members and the public of their legitimacy.

Along similar lines, critics have doubts about the professionals’ claims regarding serving the public interest and using licenses to keep out the incompetent. Surely, the argument goes, the purpose of licensing cannot be to improve the quality of service, since practitioners, as opposed to their clients, are the systematic promoters of licensing. Suspicious of the claim that professional licensure is beneficial for the public interest, Friedman argues that if the public interest was at stake, then the public, as opposed to the producers, would be behind the movement to push for legislative protection. As he puts it:

> In the arguments that seek to persuade legislatures to enact such licensure provisions, the justification is always said to be the necessity of protecting the public interest. However, the pressure on the legislature to license an occupation rarely comes from the members of the public who have been mulcted or in other ways abused by members of the occupation. On the contrary, the pressure invariably comes from members of the occupation itself. Of course, they are more aware than others of how much they exploit the customer and so perhaps they can lay claim to expert knowledge.

One might object that the public’s absence from the campaign scene need not be taken as proof of a stance against licensure. Purchasers can also display their interests through using (or

---

104 See Gellhorn, Friedman, and Benham discussed earlier in this chapter.
106 Friedman, *Capitalism and Freedom*, 140.
refusing to use) the services of licensed professionals. Consider, for example, the consumer movements to purchase non-genetically modified food or fair trade products at grocery stores. It also seems plausible that there might be cases where the initial impetus for regulation comes from the public, but is then co-opted by the private interests of the profession at a later time (or vice-versa). These various interests might be pushing and interacting with one another in the process of regulation for any given product. For example, in the case of fair trade coffee, although the movement began with the consumers’ hope to assist independent coffee farmers around the world, the final result arguably excludes farming operations that are too small to afford the costs of being regulated and monitored by the bodies in charge of doling out the fair-trade stamp. Still, the consistent absence of market failure arguments from the position of consumer groups, and the prevalence of these same arguments in the position of professional industry groups does raise some serious concerns.

**Denying Market Failures**

Finally, some theorists claim that the economic arguments regarding market failure are more or less fabricated, in order to support the self-serving legitimizing efforts of the professions that lobby for licensure. Benham puts this point as follows:

… a common activity in twentieth-century America has been the cultivation of the public perception of “market failures” or externalities. Almost all licensed occupations have claimed they will successfully cope with undesirable market failures. Frequently there has been little or no evidence in support of the proposition that such externalities exist or that the proposed solution will improve the situation. The absence of systematic evidence in support of such claims has, however, never appeared to act as a deterrent.\(^\text{107}\)

---

\(^{107}\) Benham, “The Demands for Occupational Licensure,” 17. Critics on the other side have made similar claims about the lack of adequate empirical evidence for the problematic nature of regulations. See discussion in Philipsen “Regulation of Liberal Professions and Competition Policy.”
Benham blames economists for having “rarely given any systematic attention” to the impact of market failures on professions, and goes so far as to classify research regarding public benefits of professions as merely legitimizing efforts on the part of regulated organizations.\textsuperscript{108} He recounts the example of American barbers in 1901, who sought legislation to keep out barbers with lower standards on the grounds of public health.\textsuperscript{109}

Many argue in favor of professional regulations, because without them, non-expert clients would not be able to detect and police the heterogeneity of talents of professionals.\textsuperscript{110} The worry is that despite the professional association’s promise of homogeneity of services, professional talent is not uniform. But Friedman calls this reasoning “strictly paternalistic,” since it implies that ordinary individuals are “incapable” of choosing service providers such as physicians and barbers, because they lack the specialized skills to discern experts from mediocre practitioners.\textsuperscript{111} This is the kind of argument that Friedman claims he encounters most often from proponents of licensure. Surely, he argues, professional clients are capable of performing enough market research to discern the difference between services of professionals, without needing the meddling hand of the government. Private industrial testing laboratories, consumer testing agencies, educational certificates, brands, chains, repeat purchases, and community information all can serve as remedies for operating in non-regulated markets.\textsuperscript{112} In effect, critics like Friedman question the prevalence and problematic nature of information asymmetries in the professional service market.

\textsuperscript{108} Benham, “The Demands for Occupational Licensure,” 17.
\textsuperscript{111} Friedman, \textit{Capitalism and Freedom}, 148.
\textsuperscript{112} Ibid. See also Rottenberg, \textit{Occupational Licensure and Regulation}, 7. Social media also has an important role in this context.
Other critics, such as Rottenberg, view any positive evaluation of professional regulation as a false script spread by practitioners whose interests are served by regulation. Rottenberg is aware that in markets with asymmetric information, the forces of competition will inevitably lead to the survival of groups that offer services at the lowest quality and the lowest price. He also concedes that this is the reason why some argue that the state “must act as the agent of consumers” and ban incompetents from practicing professions in environments of asymmetric information. Still, he notes that any “casual observer” can report that in unregulated markets for food, education, health care, etc. “relatively high-quality goods and services are available (at relatively high prices).” This observation alone leads Rottenberg to conclude that “probably market processes operate in such a way that the assumption of information asymmetry is rarely fulfilled.” He views the claims of economists about the perils of market failures as more or less imaginary, and explains: “it does not appear that competitive markets in the real world serve consumers as badly as the informational asymmetry model suggests they do.” Repeatedly purchasing the same products, and seeking advice from “the experience of kinfolk, friends and neighbors,” is said to serve as more than adequate remedies for operating in a non-regulated market.

The question at hand now is whether there are reasons, despite all the problems pointed out by private interest theorists, why professional regulations might be desirable. The public interest theorists of regulations would argue that there is. While resolving the dispute between private and public interest theorists of regulation is beyond the scope of my dissertation, clarifying the

114 Ibid., 6.
115 Ibid., 7. In contemporary context in Canada, all of these markets are heavily regulated, but Rottenberg is writing in the United States, and presumably referring specifically to unregulated parts of these markets.
116 Ibid.
117 Ibid. (Italics mine)
118 Ibid.
nature and role of market failures can be insightful in this context. In Part II of this chapter, I provide a brief overview of information economics, in order to illustrate that an outright denial of market failures is almost certainly off the mark.

3.2 Information Asymmetries: A Backgrounder

The rise of the literature on information economics in the latter half of the twentieth century has had a profound impact on thinking about economic and social phenomena. Still, as we have seen in Part I of this chapter, the prevalence and problematic nature of market failures are not always recognized. This is worrisome in part because market failures due to information asymmetries have an important impact on professions and the work of professionals. I begin by providing a broad overview of the emergence of information economics and its consequences for mainstream economic thought. I then present some game theoretic frameworks and models of information economics, and discuss some strategies for contract design in the context of information asymmetries. This discussion allows me to explore some of the ramifications of information asymmetries for efficiency in the market for professional services. Ultimately, in the next chapter, I formulate the essence of professionalism as an institutional structure, which when directed by appropriate norms, can create Pareto improvements by allowing transactions to occur despite information asymmetries in the market for professional services.

Three core assumptions in modern neo-classical economics (specifically in the “invisible hand theorem”) are as follows: a) the existence of perfect competition, b) the possession of
perfect information, and c) the lack of monopolies in the market. These basic assumptions simplify economic analysis and provide superior conditions for theorizing. So long as these assumptions are accounted for, the presumption has been that we can accurately theorize about efficiency, viz. the key currency of economic research. In this section, I explore the nature of the assumption about perfect information in particular, and I discuss some of the difficulties in taking the lack of information into account when the assumption does not hold.

A select number of historical figures in economics, including Smith, Marshall, Weber, and Mill, were aware of the complex role of information, although they did not conceive of this role as problematic. The hope was that economies with almost perfect information would look similar enough to the economies with perfect information, and thus that idealized models would be sufficient so long as information was not too imperfect. Mainstream economic theory (as embodied in the competitive general equilibrium theory formalized by Arrow and Debreu), presumed that the fundamental theorems of welfare economics and the tools that provided insight for economies with perfect information would also hold for economies with imperfect information. But information is imperfect, obtaining information can be costly, and there are important asymmetries of information. Mainstream theorists were aware of these problems of course, but had hoped they did not make a big difference. But the extent of

122 According to general equilibrium theory, the economy is a collection of economic agents that further their interests by making supply and demand decisions over commodities, labour types and assets. The key focus of general equilibrium theory is on the existence and efficiency properties of competitive equilibrium. See W. D. A. Bryant, General Equilibrium: Theory and Evidence (World Scientific Publishing, 2010).
information asymmetries is affected by the actions of individuals as well as firms, and information asymmetries can have a significant impact upon the outcome of the market.\textsuperscript{125} Furthermore, “close enough” to perfect information turns out to make a big difference.

In 1945, Hayek argued that the central problem of economics was a problem of knowledge (or information), although he focused narrowly on the problem of information about scarcity.\textsuperscript{126} Hayek showed that designing an efficient economic system entailed finding “the best way of utilizing knowledge initially dispersed among all the people.”\textsuperscript{127} In this context, the decentralized price system\textsuperscript{128} led to the efficient allocation of scarce resources. Even though the preferences of individuals and the technologies of the firms were unknown, resource allocation was taken to be Pareto efficient,\textsuperscript{129} since prices conveyed the relevant information.\textsuperscript{130} But arguably, Hayek focused too narrowly on the problem of information about scarcity. Similarly, the mainstream economics of the nineteenth and early twentieth century focused too narrowly on how the price system solved the problem of resource allocation once and for all. What was left out was an approach that determined how the market system responded to a continual barrage of new information.\textsuperscript{131}

As I show below, in addition to information about scarcity, other problems of information include selection problems, where the characteristics of the items being transacted are unknown. Consider, for example, the fact that employers want to know the productivity, strengths, and

\begin{flushleft}
\textsuperscript{125} Stiglitz, “The Contributions of the Economics of Information …,” 1441.  
\textsuperscript{127} Ibid., 520.  
\textsuperscript{128} In a centralized economy, a central authority (e.g. a government) is in charge of economic planning and regulation. In contrast, a decentralized system is a market economy where autonomous economic entities make decisions about their individuals prices, investments, production, and distribution.  
\textsuperscript{129} A Pareto efficient improvement is one that improves the utility of one member of a group, without making any other member any worse off than she or he already is.  
\textsuperscript{130} Stiglitz, “The Contributions of the Economics of Information …,” 1446.  
\textsuperscript{131} Ibid., 1449.
\end{flushleft}
weaknesses of their workers; investors want to know the potential return on various assets in
which they may invest; and insurance companies want to know the likelihood that the various
people they insure might have an accident or get sick. In the professional service market,
prospective clients want to know the quality and standard of services they are about to purchase
from service providers. The selection problem is exacerbated in this case since information for
evaluating services is not always available, and clients are often not specialized enough to
intelligently evaluate the quality of professional services. Furthermore, the nature of many
professional services complicates the relationship between quality of service and utility. For
example, we pay for the services of a physician, regardless of whether those services make us
feel any better, and we pay for the services of an attorney, whether or not she can save us from
going to prison.\footnote{See the discussion regarding the nature of professional services in section 3.1 in this chapter.} Also, the quality of service is often discernable only after it has been
provided, further complicating the evaluation of service quality prior to entering into contracts
under conditions of uncertainty.

In addition to selection problems, another set of information problems is \emph{incentive} or \emph{moral
hazard} problems. Here, the focus is on information about behavior after the transaction has taken
place. For example, employers want to know how hard their workers work after they have been
employed, insurers want to know what care their insured take to avoid an accident, and lenders
want to know what risks their borrowers may take.\footnote{Stiglitz, “The Contributions of the Economics of Information …,” 1447.} Similarly, professional service buyers want
to know how hard service providers are working, and what kind of actions they are taking in
providing them with high-quality services.\footnote{Not everyone agrees that moral hazard problems arise in the market for professional services. As we shall see
later in this chapter, this is because professional services are often thought not to have a perfectly elastic supply
curve. Influential thinkers such as Akerlof (1979) and Leland (1980) view professional quality as an adverse
selection problem, but Shapiro (1986) and others also underscore moral hazard problems.}
clients lack the technical skills to evaluate the service provider’s quality of service. This means that service providers may neglect to exert the required level of effort in providing the services they are contracted to perform, and, consequently, they may not practice due care in providing services to clients.

I provide a discussion of some of the relevant concepts in information economics below. After providing a background discussion of information asymmetries in the general case, I elaborate on the implications of information asymmetries for the professional service market. But first, I explain the relation between information asymmetries and the problem of efficiency.

3.2.1 Efficiency

Among the various forms of market failure, information asymmetries are the most important to consider in thinking about the professions, since the possession of specialized knowledge or skills is central to the traditional definition of professions. The connection between information asymmetries and efficiency is worth expanding on in this context. The literature prior to the development of information economics attempted to establish the efficiency of market economies by focusing on special cases. For example, theorists focused on a single consumption good or a single accident, where each individual purchased all of her insurance from a single insurer.\(^{135}\) In the market for professional services, the focus used to be on one particular service, where the professional client purchased all her services from one professional service provider. The problem is that such cases occur under unique conditions, and thus do not serve as an appropriate gauge for the efficiency of markets in general. The actions taken by one

service provider can affect service prices in general and, thus, the profits of other service providers – both in that same profession, and in competing professions. These effects give rise to market failure in the presence of information asymmetries such as moral hazard.\footnote{Arnott and Stiglitz, “The Welfare Economics of Moral Hazard,” 101.} 

We may formalize the notion that an economic system should maximize consumer welfare and exploit all opportunities to benefit some individuals without penalizing anyone else. This is called the efficiency criterion.\footnote{Donald E. Campbell, Incentives: Motivations and the Economics of Information. 2nd ed. (Cambridge: Cambridge University Press, 2006), 23.} Efficient and inefficient outcomes may be defined as follows: “Outcome A is efficient if it is feasible and there is no other feasible outcome B that gives everyone at least as high a payoff as A and gives at least one individual a strictly higher payoff than outcome A. An outcome is inefficient if it is not efficient.”\footnote{Ibid.} In an efficient economic system, individual production and consumption activities are coordinated such that all opportunities to increase welfare are taken without reducing anyone’s wellbeing. In an inefficient economic system, equilibriums exist that could be improved upon and doing so could make some people better off without adversely affecting anyone else.\footnote{Ibid.}

Formalizations of Adam Smith's notion of the invisible hand have been the central propositions in welfare economics for a long time. But the informational assumptions underlying the theorems were generally not made explicit until the last decade of the twentieth century. Greenwald and Stiglitz\footnote{Bruce C. Greenwald and Joseph E. Stiglitz, “Externalities in Economies with Imperfect Information and Incomplete Markets,” The Quarterly Journal of Economics 101, 2 (1986): 229-264.} showed that when information is imperfect (or markets are incomplete) – which is almost always the case – competitive markets, in general, fail to be constrained Pareto
efficient. In a constrained Pareto efficient market, informational or institutional constraints help us improve market outcomes for individuals in a group without making at least one of those individuals worse off. The first fundamental theorem of welfare economics thus describes the singular set of circumstances under which the economy is Pareto efficient. Greenwald and Stiglitz listed some of the features of real markets that make them inefficient:

There is not a complete set of markets; information is imperfect; the commodities sold in any market are not homogeneous in all relevant respects; it is costly to ascertain differences among the items; individuals do not get paid on a piece rate basis; and there is an element of insurance (implicit or explicit) in almost all contractual arrangements, in labor, capital, and product markets. In virtually all markets there are important instances of signaling and screening. Individuals must search for the commodities that they wish to purchase, firms must search for the workers who they wish to hire, and workers must search for the firm for which they wish to work.

Although each of these is a small element, the cumulative effects may turn out to be substantive. A model that draws on information economics demonstrates that in all of the above circumstances, Pareto improvements are possible. Greenwald and Stiglitz’s results thus provided a negative answer to what had long been a goal of economic research, that is, to find general conditions in which the competitive economy is still constrained Pareto efficient although a complete set of markets is nonexistent.

Now that I have discussed the relevance of efficiency in thinking about information asymmetries, I turn to explaining some of the important concepts in information economics, such

---


143 Ibid.

144 Ibid., 260

145 Ibid., 250. Earlier studies had shown that there could be multiple equilibria with an incomplete set of markets, some of which Pareto dominated others. But ultimately, it Greenwald and Stiglitz demonstrate that, in general, and except for some special unique cases, every equilibrium is Pareto inefficient.
as principal agent relationships, and information problems such as moral hazard and adverse
selection.

3.2.2 Principal-Agent Relationships & Games of Asymmetric Information

Principal-agent relationship arises when “one economic agent – the agent – takes an action
that affects another economic agent – the principal.” According to Milgrom and Roberts,
agency relationships occur in “situations in which one individual (the agent) acts on behalf of
another (the principal) and is supposed to advance the principal’s goals.” Principal-agent
relationships exist all over the place: e.g. between managers and the firm owners, between
managers and their subordinates, between the electorate and their elected officials, between
elected officials and members of their bureaucracies, and between service providers and their
clients. In the context of the professional service market, the principal is the client and the
agent is the service provider. Within a game theoretic framework, the principal and the agent are
two “players” in a game, who interact with each other via a contract. The players make “binding
contracts” at a certain stage in the game, viz., the principal commits to paying the agent an
agreed sum in return for the agent’s commitment to pursue the principal’s specific objectives.

146 Dutta, Strategies and Games: Theory and Practice, 293.
147 Paul Milgrom and John Roberts, Economics, Organization, and Management (Upper Saddle River, NJ: Prentice
Hall, 1992), 170. For more on agency theory see Daniel M. Hausman and Michael S. McPherson, Economic
Analysis and Moral Philosophy (Cambridge: Cambridge University Press, 1996); Jean-Jacques Laffont and David
149 Eric Rasmusen, Games and Information: An Introduction to Game Theory. 2nd ed. (Oxford: Blackwell, 1997),
166.
A typical principal-agent game can be represented by the game tree in Figure 3.1:

Figure 3.1. The Principal-Agent Game

Agency theory formalizes the risks posed by agents when they act according to interests that conflict with the interests of their principals. Here is how the game works: Suppose that the agent can only choose between two actions and that each action can have three different consequences. The principal moves first on the tree. A “move” represents a given action by a player. When the principal offers the agent a contract, for example, this can be represented by a move on the part of the principal, shown by a line in the game tree that originates from the principal. At the root of the game tree, the principal decides what kind of incentives to offer the agent. The principal offers incentives in the form of w and w’, or W and W’ in this particular

---

game. W, for example, might represent the compensation package provided by a manager to an employee (where the employee is the agent and the manager is the principal).

The agent makes the next move on the tree, by either accepting or rejecting the principal’s offer. Here, based on the incentive scheme, the agent chooses to exert high or low effort, denoted by $e_H$ and $e_L$. Effort in this context includes decisions about how hard to work, what projects to select, what quality service to provide, and even whether to remain in the contract with the principal. The agent’s action is unobservable to the principal, and since the agent chooses her action after the principal, the agent’s decision nodes form the second stage of the tree.\(^{151}\)

The transaction’s profitability depends on the agent’s actions. Profit levels may be good (g), medium (m), and bad (b), such that $g>m>b$. Still, the profit determination is uncertain, and uncertainty is modeled via the fictional player Nature.\(^{152}\) Nature is a pseudo-player who takes random actions at specified points in the game with specified probabilities. It is sometime useful to include individuals called pseudo-players in the game, whose actions represent a state of the world.\(^{153}\) For example, in our game tree, we can model the possibility of a recession as a move by Nature. With probability $x$, there will be a serious recession, with probability $y$, there will be a mild recession, and with probability $z$ (where $x+y+z=1$), there will be no recession. This random move by Nature means that the model yields more than just one prediction. In response to a serious recession in Figure 3.1, the agent might decide to work harder and produce profit level g, and so on for profit levels m and b. There are thus “different realizations of a game depending on the results of random moves.”\(^{154}\) The profit determination in this case is uncertain, i.e. every one

\(^{151}\) Ibid., 295.
\(^{152}\) Ibid.
\(^{153}\) Rasmusen, *Games and Information*, 11.
\(^{154}\) Ibid.
of the agent’s actions leads to a “probability distribution over profits.” Here, a one-to-one relationship does not exist between the agent’s actions and the firm’s profit because of uncertainty. This means that when the principals see the profit level \( g \), she cannot infer whether the agent exerted high or low effort.

The principal sets the agent’s conditional wage at \( w_g \) if he observes profit of size \( g \), and so on for profits \( m \) and \( b \). Since the principal may not be able to observe (or adequately evaluate) the agent’s efforts, the principal can only reward the agent on the basis of profits. Thus, the net profit for the principal is \( (g - w_g) \) whenever gross profits are \( g \), \( (m - w_m) \) when profits are medium, and \( (b - w_b) \) when profits are bad. The agent’s utility then depends both on her compensation and the effort that she exerts. If she is paid \( w_m \) and she has taken action \( e_H \), then her total payoff is \( u(w_m) - d_H \), where \( u(w_m) \) is the utility from compensation \( w_m \), and \( d_H \) is the disutility of \( e_H \). A basic assumption here is that the agent prefers more money to less and is risk averse, and that the disutility of high effort is greater than low effort, that is \( d_H > d_L \).

We may model the professional service market using this same game tree for the client (the principal) and the service provider (the agent). The game begins when the client offers the service provider a contract that the service provider accepts. The service provider might then exert high or low levels of effort in carrying out her duties towards the client, and arrive at different results depending on random variables such as events and conditions in the professional service market.

---

156 Ibid., 296.
157 Ibid.
158 Ibid.
Consider the relationship between a lawyer and her client as an example. The client offers the lawyer a contract, which she may or may not accept. If she accepts the contract, then she may exert high or low effort at her work (although this effort level is unobservable by the client). Suppose that the lawyer’s case for her client is affected by the results of a similar case currently in court. The outcome of the alternative case is unknown (Nature), but would have important implications for how the lawyer will argue for her case. The state of the world determined by Nature then determines the lawyer’s “profits,” or her overall ability to provide services to the client, as bad, medium, or good. Here, the lawyer’s “payoff” is the utility of her rewards from working on the case, minus the disutility of her efforts. Her rewards may include fees she receives from her client, but also the reputational effects of her performance. Her disutility might involve time and effort exerted over research and strategizing leading up to going to court. The client’s payoff is then the service she receives from the lawyer (good, medium, or bad) minus the fees she pays to the lawyer. The profit outcome might be in monetary terms (e.g. if the client is suing someone), in personal terms (e.g. when a rape victim gets her assaulter to go to jail), in social justice terms (e.g. when a statute results in greater equality for minority groups), or a combination thereof.

As a modeling convenience, in most principal-agent games, it is assumed that the order of moves allows the principal to make a take-it-or-leave-it offer to the agent for the job. Thus, the agent has little bargaining room if he has to compete with other agents. The take-it-or-leave-it offer helps simplify the analysis so we can avoid getting caught up in a bargaining sub-game.\textsuperscript{159} Still, professional service providers (agents) are often characterized in the critical regulation literature as having more bargaining power than their clients (principals), which makes it

\textsuperscript{159} Rasmusen, \textit{Games and Information}, 170.
difficult for their clients to choose among different suppliers. This characteristic is arguably enforced through the high level of required training and experience, and the “esoteric”\textsuperscript{160} nature of professional knowledge, which makes the services of service providers relatively rare and in high demand. Furthermore, the bargaining hierarchy in professional relationships can be viewed as problematic, since many service providers (e.g. doctors, lawyers, and teachers) offer what we take to be basic and fundamental services that are needed by every member of society.

The principal-agent model is an example of a relationship where the parties are in danger of having asymmetric information. In a game involving asymmetric information, the information sets of the players differ in ways relevant to their behavior, or differ at the end of the game. The essence of asymmetric information is that some player has useful private information, viz. information that is different and not worse than any other player’s.\textsuperscript{161} In Figure 3.1, the principal moved first, and was unaware of the action the agent would choose to take at the second node. This implies that the agent observes her own move, and the principal is only capable of possibly deducing the nature of the agent’s action from observing the outcome.\textsuperscript{162}

In this context, the principal can be defined as the uninformed player who has the coarser information partition, while the agent is the informed player who has the finer information partition.\textsuperscript{163} Player i’s information set $w_i$ is the set of different nodes in the game tree that he knows might be the actual node, but between which he cannot distinguish by direct observation.\textsuperscript{164} The information partition represents the different positions that the player knows he will be able to distinguish from each other at a given stage of the game, thereby carving up the

\textsuperscript{160} See the discussion of the structural turn in defining professions in chapter one.
\textsuperscript{161} Rasmusen, Games and Information, 47.
\textsuperscript{162} Ibid.
\textsuperscript{163} Ibid., 166.
\textsuperscript{164} Ibid., 40.
set of all possible nodes in the subsets called information sets.\textsuperscript{165} Consider the following two information partitions as an example:

Partition I: (J1), (J2), (J3), (J4) vs. Partition II: (J1), (J2), (J3, J4)

A combination of two or more of the information sets in a partition, which reduces the number of information sets and increases the number of nodes in one or more of them, is a coarsening. The ultimate refinement is for each information set to be a “singleton,” containing one node as in partition I above. This is called a game of perfect information. Thus, partition II is coarser, and partition I is finer. A finer information partition is the formal definition for “better information,” although information quality is defined independently of its utility to the player. In other words, it is possible for a player’s information to improve and yet his equilibrium payoff (i.e. the payoff he receives in equilibrium) to fall.\textsuperscript{166}

In a game of asymmetric information, the information sets of the players differ in ways relevant to their behavior, or differ at the end of the game. Such games have imperfect information, since information sets that differ across players cannot be singletons.\textsuperscript{167} In a game of perfect information, each information set is a singleton. In a game of imperfect information, one player is unaware of actions taken by other players (including nature), either in the past or contemporaneously.\textsuperscript{168} Professional services are games of imperfect information, where the client (principal) is unaware of the service provider’s (agent) moves (e.g. effort levels). The physician thus has private information and finer information partitions in comparison to her patient. For example, a patient may be unaware of how well her doctor is trained with respect to

\textsuperscript{165} Ibid., 42.
\textsuperscript{166} Ibid., 43.
\textsuperscript{167} Ibid., 47.
\textsuperscript{168} Dutta, \textit{Strategies and Games}, 309.
the specific illness the patient is being treated for. Assuming that the doctor knows the proper
route to diagnosing and healing the patient, the patient may still not know whether the doctor is
taking all necessary actions en route to proper treatment.

In the context of principal-agent relationship, striving for efficiency involves finding a
contract that maximizes welfare given the informational constraints. In this context, two types of
contracts are possible: i) A first best contract achieves the same allocation as the contract that is
optimal when the principal and the agent have the same information set and all variables are
contractible; ii) a second best contract is Pareto optimal given information asymmetry and
constraints on writing contracts. The difference in welfare between the first-best world and the
second-best world is the cost of the agency problem.¹⁶⁹

_Moral Hazard_

A principal-agent problem has a moral hazard component if the principal cannot simply
force the agent to act according to the principal’s interests.¹⁷⁰ Arrow discussed the problem of
moral hazard, or incentives, in his Yrjo Jahnsson lectures.¹⁷¹ He explicated the issue in the
context of insurance, where individuals who are insured against a risk have inadequate incentives
to take actions to avoid that risk. A person who has insurance coverage will have less incentive
to take proper care of an insured object than a person who does not.¹⁷² The agent’s interests and
the principal’s interests are thus not necessarily congruent, and the principal cannot always
monitor the agent’s activities. If we could assume perfect information, then the insurance

¹⁶⁹ Rasmusen, _Games and Information_, 175.
¹⁷⁰ Dutta, _Strategies and Games_, 293.
¹⁷² Dutta, _Strategies and Games_, 305.
contract would stipulate the actions that were to be taken by the policyholder and the company, and there would be no moral hazard or incentive problem. But actions are at best imperfectly observed. For example, insuring firms can require (and monitor) the installation of sprinklers in commercial buildings, but it is not possible to monitor every instance of a fire hazard (e.g. the incidental smoking habit of an employee).\textsuperscript{173} Similarly in the case of car insurance, firms cannot monitor how carefully their insured clients drive.

Since it is in no individual policyholder’s interest to exert better care, an opportunity may exist for government interventions to lead to Pareto improvements, making everyone better off.\textsuperscript{174} As Arnott and Stiglitz proved in 1992, economies where moral hazard problems exist have numerous forms of potential inefficiency, and are essentially never constrained Pareto efficient. This analysis has cast serious doubts on the relevance of the fundamental theorems of welfare economics, and on the potential for the efficient decentralizability of the economy.\textsuperscript{175}

Applied to the professional service market, moral hazards arise when the service provider (agent) has inadequate incentive to work hard enough, provide the right quality of service, or even to remain in a contract with the client (principal). Here, the agent’s interests and the principal’s interests are incongruent, and the client cannot adequately monitor or evaluate the service provider’s activities. As a result, the client may simply prefer to do things herself, or forgo the service, because she cannot trust anyone else to do the job well. If we could assume perfect information, then the service contract would stipulate the actions that were to be taken by the client and the service provider, and there would be no moral hazard or incentive problem. Since actions are imperfectly observed, we may need monitoring and incentive pay solutions as a

\begin{footnotesize}
\begin{itemize}
\item\textsuperscript{173} Stiglitz, “The Contributions of the Economics of Information …,” 1453.
\item\textsuperscript{174} Greenwald and Stiglitz, “Externalities in Economies with Imperfect Information and Incomplete Markets,” 230.
\item\textsuperscript{175} Arnott and Stiglitz, “The Welfare Economics of Moral Hazard,” 91-92, and 115.
\end{itemize}
\end{footnotesize}
condition of remaining in the contract and ensuring the receipt of payment. For example, doctors might receive income supplements from pharmaceutical companies for prescribing a medical solution that does not unambiguously serve the long-term wellbeing of patients. Consider the widespread use of pacemakers for heart patients, despite research that shows the device some times keeps patients’ hearts beating long after the rest of the body gives up. Some would argue that being kept alive in a vegetative state, and being forced to do so without having indicated such a preference ahead of time, is worse than the alternatives. Patients for the most part lack the ability and expertise to monitor and evaluate their physician’s services.

Monitoring and incentive pay are two strategies developed as a solution to address moral hazard.\textsuperscript{176} A principal can offer a variety of incentive schemes—e.g. wage-based, franchise-based, and wage-plus-bonus schemes.\textsuperscript{177} In general, these incentive schemes suggest that in order to elicit hard work, we need to give bonuses for good results: the higher the profits, the larger should the size of the agent’s bonus be.\textsuperscript{178} A key recent finding is that it is dangerous to focus incentives on easily observable variables, as they may divert attention from other less easy to measure but nevertheless more important variables. Consider, for example, the importance of performance in reading and writing, in comparison to cognitive skills in education.\textsuperscript{179} Cognitive skills, while less observable, may be more important than reading and writing skills, which are easily observable. Still, it is often challenging to focus incentive schemes so as to capture less observable effects and to reward students that display them.

\begin{flushleft}
\textsuperscript{177} Dutta, \textit{Strategies and Games}, 297-9.
\end{flushleft}
Professional service markets are susceptible to both moral hazard, and adverse selection problems. In the next section, I elaborate on the impact of adverse selection.

Adverse Selection

The earliest contribution to the literature on adverse selection was presented in the context of income tax system design. Mirrlees\textsuperscript{180} discussed the design of an income tax system with the purpose of maximizing social welfare. In this context, if the government knows everyone's ability, it should levy lump sum taxes on everyone, and there is no need to be concerned about distortions. But the government cannot know everyone’s ability. All it can do is make inferences about individual ability based on observed income.\textsuperscript{181} Mirrlees’ problem arises whenever one party, e.g. the government, attempts to design a set of self-selecting contracts – such as an income tax schedule – with the goal of maximizing an objective such as social welfare.\textsuperscript{182} The problem of self-selection, as it is called today, involves the process through which individuals reveal information about themselves through their choices.\textsuperscript{183} The self-selection problem has been recast in terms of Pareto optimality.\textsuperscript{184} Here the question is, based on the observable information structure, how can we maximize the welfare of one type of individual, given stipulated levels of welfare for everyone else?\textsuperscript{185} The most difficult cases, which still remain unresolved, arise in the

\textsuperscript{181} Stiglitz, “The Contributions of the Economics of Information…,” 1450.
\textsuperscript{182} Ibid.
\textsuperscript{184} Given an original allocation of goods among a number of individuals, a Pareto improvement involves changing the allocation of goods such that at least one individual becomes better off without making any other individual worse off. Pareto-efficiency or Pareto-optimality are achieved when no further opportunities for Pareto improvements exist.
\textsuperscript{185} Stiglitz, “The Contributions of the Economics of Information…,” 1451.
analysis of competitive equilibria in economies in which individuals have limited information and know this to be the case.

George Akerlof stimulated an entire field of research when he introduced his model of the market for shoddy used cars (“lemons”) in 1970. Akerlof explained that adverse selection arises because car quality is better known to the seller than to the buyer. In agency terms, the principal contracts to buy a car from the agent, but the quality of the car, whether high or low, is non-contractible. The seller in the used car market has private information about his car’s type before agreeing to the contract, and the prospective buyer does not know about this private information related to the characteristics of the car. This means that sellers might overpay for a low-quality used car, because they lack the information that would reveal inferior value prior to purchasing the car. More problematically, knowing about this risk of overpaying, fewer buyers might decide to enter contracts in the used car market.

As a result, high-quality used car sellers might find fewer buyers and eventually hesitate to advertise in the used car market, resulting in even further efficiency losses and deadweight losses. A deadweight loss is an efficiency loss that occurs when equilibrium for a good or service is not Pareto optimal. Deadweight losses occur when buyers whose marginal benefit outweighs their marginal cost are not entering into contracts. In the market for lemons, this leads to lost transactions, because buyers do not purchase high-quality used cars even when they are offered at lower prices.

---

187 Ibid., 494.
188 Ibid.
Akerlof focused on economic models where trust plays a pivotal role. As he explained, beyond formal written contracts, “informal unwritten guarantees are preconditions for trade and production,” and business suffers when informal guarantees are indefinite. While this type of uncertainty was explored by game theorists (e.g. via the Prisoner’s Dilemma) prior to 1970, it was Akerlof who explored the implications for the traditional Arrow-Debreu approach to uncertainty. The core strength of Akerlof’s short and influential article is its analysis of the relation between quality and uncertainty in the theory of markets. He argued that the interaction of quality differences and uncertainty can explain a number of important institutions of the labor market, including for example, government intervention in medical insurance and the licensing of professionals.

Let us unpack his market for lemons argument. Akerlof proposed an explanation for the large price difference between new cars and cars that have recently left the showroom, as follows. He assumed, for the sake of clarity, that there are only four properties of cars: new, used, good, and bad. A new car might be a good car, and a used car may be bad, or a “lemon.” In this market, individuals tend to buy cars without knowing whether it is a good car or a lemon, although they know that the probabilities of ending up with one or the other is q versus (1-q), where q is the proportion of good cars produced.

Sometime after I purchase a used car, I will acquire knowledge of whether my car is good or a lemon. An asymmetry in information now exists, since, if I decide to sell my car, I will have more knowledge about its quality in comparison to a potential buyer. Still, since the buyer cannot

---

189 Ibid., 500.
190 Ibid.
191 Ibid., 488.
192 Ibid., 489.
differentiate between good cars and lemons, he will pay the same amount for each. More owners of bad cars will want to resell them, so the incidence of lying in the used car market will be higher than in the new car market. And the value of a new car is clearly higher than the value of a used car. Yet, as Akerlof puts it, “the bad cars sell at the same prices as good cars since it is impossible for a buyer to tell the difference between a good and a bad car.” Buyers thus cannot differentiate the good or bad quality of the car prior to selecting the car for purchase. Eventually, “good cars may be driven out of the market by the lemons,” in so far as the owner of a good car is no longer able to enter a transaction to sell his car at the appropriate price in a market for lemons. This phenomenon is sometimes referred to as a “lost transaction.” Eventually, we end up with a higher than usual fraction of cars in the used car market being lemons. The average quality of used cars falls, alongside a fall in the price of used cars, since lemons get sold and resold. This is called “adverse selection.” Depending on the grades of goods in a market, even worse pathologies than lost transactions in the market for lemons may come about. The bad quality can drive out the not-so-bad, which in turn can drive out the medium-bad, which itself can drive out the not-so-good, and ultimately even the good, so that eventually the market is no longer in existence. Akerlof’s paper represented the first attempt to produce a partial equilibrium model, where the market price affected the quality of the good offered, and in turn consumer demand. Akerlof emphasized the possibility that little or no trade might take place in

---

193 In this context we are not considering used, rare car models that might fetch higher prices than an average new car.
195 Ibid.
196 Ibid., 493.
197 Ibid., 498.
198 A partial equilibrium model takes only a restricted range of data about the market into consideration while holding all other factors constant. For example, we might inquire about how the price of one particular item changes based on information about the quality of that product, while holding all else fixed.
equilibrium, resulting in a demand curve that looked markedly different from standard demand curves.\textsuperscript{199}

Interesting applications of this phenomenon to the insurance market are widespread. Why do people over the age of 65 have difficulty buying medical insurance in the United States? In other words, why doesn’t the price of medical insurance “rise to match the risk”?\textsuperscript{200} Akerlof’s answer is information asymmetries. Individuals, as opposed to insurance companies, have more knowledge about their own medical needs. As the price of medical insurance rises, individuals who are more certain of their eventual need for medical attention will increasingly purchase insurance, as opposed to healthier individuals who are at lower risk. As a result, the “average medical condition of insurance applicants deteriorates as the price level rises.”\textsuperscript{201} The insurance applicants have information about their health and their potential cost of care, but this information is not available to insurance companies who are selling and pricing insurance. Akerlof also points out that insurance companies “do their own ‘adverse selection’,” because group insurance, offered through employment often “picks out the healthy,” since health is a “precondition for employment.”\textsuperscript{202} Thus, insurance companies can leave out those who need medical insurance the most.

The prevalence of information asymmetries (on the part of patients and insurance companies) can provide an argument in favor of government regulation of medical insurance. From a cost-benefit perspective, government intervention in providing medical insurance may pay off, because although each individual in the market might be willing to buy insurance and

\textsuperscript{200} Akerlof, “The Market for ‘Lemons’,” 492.
\textsuperscript{201} Ibid.
\textsuperscript{202} Ibid., 494.
pay the expected cost of her medical expenses, private insurance companies cannot afford to sell policies in this manner. Such a strategy would “attract too many lemons.”\textsuperscript{203} Lemons in this context are patients whose real medical condition is bad but who cannot be distinguished from more healthy patients by the insurance company.

Another useful application of Akerlof’s model is in evaluating the costs of dishonesty. In markets where quality may be misrepresented, and in which goods are sold dishonestly, the purchaser’s problem is to accurately identify quality. The major cost of dishonest dealing, in addition to cheating the purchasers, is that it drives out honest dealings from the market. Similar to the market for lemons, the entire market might be driven out of existence as a result of the lemons or dishonest dealers. As Akerlof puts it, “there may be potential buyers of good quality products and there may be potential sellers of such products in the appropriate price range; however, the presence of people who wish to pawn bad wares as good wares tends to drive out the legitimate business.”\textsuperscript{204}

At first glance, a simple solution to the dishonesty problem might seem apparent. Once customers realize that “bad wares” are being pawned as “good wares,” won’t they abandon that seller and flock to another? Can’t customers return to the legitimate businesses that provide good wares, and even better, tell all their friends to do the same, drawing more and more customers away from pawn artists? The answer is no – at least not under the background assumptions in Akerlof’s model. If consumers were capable of overcoming information asymmetries on their own, then we would not end up with markets for lemons. Abandoning bad sellers might end up being costly for consumers as well. People could take used cars to a mechanic to check out, for

\textsuperscript{203} Ibid.  
\textsuperscript{204} Ibid.
example, but that would cost extra. Transaction costs like this make switching between service providers inefficient.

The problem is that the nature of services in these markets is such that opportunities for repeat purchase and consumer research are limited, and consumers require the help of (private or public) centralized planning to acquire relevant information. This is why Akerlof indicates a role for a number of “counteracting institutions”, including “guarantees,” “brand names,” “chains,” (e.g. hotels and restaurants) as well as “licensing” of the professions, all of which work to counteract the “effects of quality uncertainty.”

Applied to the market for professional services, adverse selection arises because the service provider, unlike the client, knows the quality of service she can provide. In agency terms, the client (principal) contracts to purchase a service from the service provider (agent), but the quality of the service, whether high or low, is non-contractible. The service provider has private information about the quality of his work, his talents, his educational background, etc., before agreeing to the contract. The prospective client, on the other hand, does not have access to this private information about the service provider. Even if the client could get access to this information ahead of time, her lack of expertise in the field might deter her from adequately evaluating the information. Moreover, since the provision of such services does not occur, in general, on a regular basis, clients have a limited chance to learn through repeat buying, and the seller’s reputation also has a limited impact. As a result of this problem, the overall quality of

\begin{footnotesize}
\begin{enumerate}
\item Ibid., 500.
\item Ibid., 499.
\item Rasmusen, Games and Information, 224.
\item For a more recent discussion of this problem see Philipsen, “Regulation of Liberal Professions and Competition Policy,” 205. To make matters more complicated, professional services are often “experience goods” and “trust goods (or credence goods).” See P. Nelson, “Information and Consumer Behavior,” Journal of Political Economy 78 (1970): 311-29; M. Darby & E. Karni, “Free Competition and the Optimal Amount of Fraud,” Journal of Law and
\end{enumerate}
\end{footnotesize}
service by service providers may start to decline. Further, the information asymmetries may lead to a lack of an environment of trust where clients and professionals interact with each other, eventually affecting (and lowering) client demand for professional services.

Lost transactions are a problem in this context, since the providers of skilled professional service may no longer be able to enter into transactions and provide their services at the appropriate price in the market for professional services. These lost transactions constitute deadweight losses or an allocative (or productive) inefficiency, since the equilibrium for a given professional service is not Pareto-optimal. Here, individuals that acquire specialized professional education or technical apprenticeships incur higher marginal costs than marginal benefits. Some examples might help clarify this point further.

Suppose I am in a position to develop a drug that is superior to other comparable products currently on the pharmaceutical drug market. Doing so, however, will require significant investment, and so the drug will have to cost more than what is currently available on the market. If I am confident that I can signal reliably to the public about the superior quality of my product, then there is hope that there will be a market for my drug even though it is priced higher than competing products. But due to information asymmetries, customers may not be able to discern the superior quality of my product, and so I’ll have to sell my product for the lower going price of inferior “lemon” products. At this point, I realize that for the price at which the average lemon product is selling, I will not be able to afford to produce my superior quality product. This means that I may hesitate to develop the product in the first place. The superior quality drug may thus

not get invented, because there is no way to distinguish it from counterfeits or lower quality drugs. The more abstract problem, here, is that certain transactions won’t occur as a result of prevalent information asymmetries.

The money and time invested in drug research are akin to the money and time investments people make by going to law school and medical school. If customers cannot discern the superior quality of the end product, however, then there is no point in investing and developing these products. Transactions will not occur unless service providers are able to reliably signal and distinguish the superior quality of their services. In the absence of guaranteed signaling opportunities offered through professionalization and the creation of professional associations, efficiency losses are incurred as a result of lost transactions.

Although Akerlof mentioned that his research had implications for professional licensing, he did not elaborate on these implications. Following Akerlof’s model, Leland showed that information asymmetric professional markets benefit from minimum-quality standards, such as licensing. Leland explained that high-quality service-providers do not find it advantageous to provide their services if the rewards they receive reflect the cost of lower average quality service. Because these service-providers withdraw from the market, the price and quality of overall service declines. For Akerlof, the worst-case scenario in asymmetric information markets was that only “lemons” would be offered for sale. Extending this argument, Leland showed that while asymmetric information markets need not “degenerate to the lowest-quality level,” still,

---

inefficiencies will necessarily prevail in competitive equilibrium, and the quality of service will be too low.\footnote{Leland, “Minimum-Quality Standards and Licensing in Markets with Asymmetric Information,” 267.}

Shortly after Leland, Shapiro made the influential contribution of reconceptualizing licensure as a kind of regulation that requires minimum levels of human capital investment by professionals.\footnote{Shapiro, “Investment, Moral Hazard, and Occupational Licensing.”} Shapiro argued that licensure raises professional training levels, and thereby alleviates the loss of utility associated with the provision of high-quality services in information asymmetric markets. The role of certification in this context is to provide consumers with information about professionals' training levels. Shapiro argued that both licensing and certification benefit consumers who value quality highly, at the expense of those who do not.\footnote{Ibid., 843-5.}

In circumstances where the investments of sellers are not observable, Shapiro argued that licensing can raise total surplus.\footnote{Ibid.} Unlike Leland, who argued in favor of certification, Shapiro showed that certification can be Pareto inferior to licensing\footnote{Recall that while certification communicates the completion of certain prerequisites and requirements for holding a particular professional title, entry into a licensed profession and transactions among buyers and sellers of licensed services require the permission of the state.\footnote{Shapiro, “Investment, Moral Hazard, and Occupational Licensing,” 862.} See Morris M. Kleiner and Alan B. Krueger, “The Prevalence and Effects of Occupational Licensing,” British Journal of Industrial Relations 48, 4 (2010): 676–687; Morris M. Kleiner “Occupational Licensing: Protecting the Public Interest or Protectionism?” Upjohn Institute Policy Papers, Policy Paper No. 2011-009; Elizabeth Graddy “Interest Groups or the Public Interest –Why Do We Regulate Health Occupations?” Journal of Health Politics, Policy and Law 16, 1 (1991): 25-49.} (or even to a laissez-faire policy), because it can lead to excessive investment as a signaling device.\footnote{Ibid.} The debate over the Pareto-efficiency of certification versus licensure continues to this day.\footnote{Ibid. 862.}

Resolving moral hazard and adverse selection problems is a difficult task. In the next two sections, I model a number of standard game theoretic problems of information, and provide some contract design solutions for resolving these problems.
3.3 Modeling Information Asymmetries: The General Case

I now elaborate upon, and model moral hazard and adverse selection problems as instances of information asymmetries. Games of asymmetric information may be divided into five categories: (i) moral hazard with hidden action, (ii) moral hazard with hidden knowledge, (iii) adverse selection, (iv) signaling, and (v) screening.217

(i) Moral hazard with hidden action is a game in which two players begin with symmetric information when they agree to a contract, but then one of the players takes an action unobserved by the other. In the context of the principal-agent model, the employer knows the worker’s ability but not her effort level.218 In (ii) moral hazard with hidden knowledge, two players begin with symmetric information when they agree to a contract. Nature then makes a move observed by only one of the players, and that player takes some action, which is possibly a simple report of Nature’s move.219 Information in both these types of games is complete but uncertain. “Certain” information entails that no moves are taken by Nature after any player has moved in the game. “Complete” information entails that Nature does not move first, or that her initial move is observed by every player.220

Some examples of moral hazard with hidden action are as follows:221

---

217 Rasmusen, *Games and Information*, 166.
218 Ibid., 168.
219 Ibid., 166.
220 Ibid., 47.
221 Ibid., 168.
1) An insurance policyholder (agent) may not exercise due care to avoid theft of her belongings, thereby posing a moral hazard risk to her insurance company (principal).

2) An insurance policyholder (agent) may drink and smoke and take part in other activities with health risks, thereby posing a moral hazard risk to her insurance company (principal).

3) A sharecropper (agent) may put in minimal or inadequate farming effort, thereby posing a moral hazard risk to the plantation owner (principal).

4) A landlord (agent) may neglect her share of the upkeep of the building without the tenant’s knowledge, thereby posing a moral hazard risk to his tenant (principal).

Some examples of moral hazard with hidden knowledge are as follows:222

1) Shareholders of a company may not know the riskiness of a company president’s investment decisions. The company president (agent) thus poses a moral hazard risk to the shareholders (principal)223

2) The FDIC may not have knowledge of the safety of loans provided by a bank. The bank (agent) thus poses a moral hazard risk to FDIC (principal).

(iii) Adverse selection models have incomplete information, since Nature moves first and chooses the agent’s type, based on the agent’s abilities to perform the task at hand.224 Here, the worker knows her own ability from the start, but the employer does not.225 Some examples of adverse selection are as follows:226

1) An insurance applicant (agent) may not reveal to her potential new insurance company (principal) that she is infected with the HIV virus

2) An employer (principal) might not be able to discern the skill level of a potential new hire (agent).

Ibid.

223 Note that in this example, the shareholders of the company observe the actions of the company president – they just do not know how risky those actions are because they do not have access to certain facts about the world.

224 Rasmusen, Games and Information, 167.

225 Ibid., 168.

226 Ibid.
If in addition to knowing her own abilities from the start, the agent can send a “signal” to the principal about her type, and if he sends the signal before the principal offers the contract, the model is signaling. If the signal is sent after the contract has been offered, the model is screening. Some examples of (iv) signaling and (v) screening are as follows:

1) An employer (principal) might discern the skill level or education level of a potential new hire (agent)

2) A buyer (principal) may evaluate the durability or warranty of a product promised by the salesperson (agent)

3) An investor (principal) may make a judgment about a stock issuer based on the stock value and the percentage of stocks retained by the issuer (agent)

Consider the game trees in Figure 3.2 as representations of principal-agent problems corresponding to the five categories of asymmetric information discussed so far. In each model, the principal (P) offers the agent (A) a contract, which he accepts or rejects. In some, Nature (N) makes a move or the agent chooses an effort level, message, or signal. The (a) and (b) moral hazard games are games of complete information with uncertainty. The principal offers a contract, and after the agent accepts, Nature “adds noise” to the task being performed. Adverse selection models, games (c), (d), and (e) have incomplete information, therefore Nature moves first and picks the type of the agent.

---

227 Ibid., 167.
228 Ibid., 168.
229 Ibid., 167.
230 Ibid., 166.
231 Ibid., 167.
3.3.1 On Contract Design

I have so far mentioned a number of information asymmetry problems and suggested that we might be able to resolve these problems through regulatory interventions that tweak contract
design. If we think about the problem of regulation in the language of principal-agent relationships, the question becomes: what is the ideal contract between the principal and the agent? I begin by considering some developments in information economics that helped theorists model games of incomplete information (such as those between professional service providers and their clients).

*The Harsanyi Doctrine*

According to the old definition of a game of complete information, all players know the rules of the game. Otherwise the game is one of incomplete information. This is the reason why, until 1967, game theorists spoke of games of incomplete information only to point out that they could not be analyzed. But the old definition is inadequate since the game itself is ill defined if it does not specify what the players’ information sets are.\(^{232}\)

According to the new definition, in a game of incomplete information, players do not know some relevant characteristic possessed by their opponents, which may include payoffs, available options, and beliefs.\(^{233}\) In three papers published in the *Journal of Management Sciences* in 1967 and 1968, Harsanyi proposed a generalization of Nash equilibrium that is appropriate for an incomplete information game. A Nash equilibrium is the most popular solution concept for strategic form games. A set of strategies and (the corresponding payoffs) constitute a Nash equilibrium, when no player can benefit by changing her strategy so long as the other players keep their strategies unchanged. In other words, the strategy vector \(s^* = s_1^*, s_2^*, \ldots s_N^*\) is a Nash equilibrium if \(\pi_i(s_i^*, s_{-i}^*) \geq \pi_i(s_i, s_{-i}^*)\) for all \(s_i\) and all \(i\). That is, each player \(i\), in playing \(s_i^*\), is

\(^{232}\) Ibid., 48.

\(^{233}\) Dutta, *Strategies and Games*, 312.
playing a best response to the others’ strategy choice. Here, $\pi_i$ is the payoff function for player $i$, $s_i$ is a strategy for player $i$, and $s_{-i}$ is a strategy vector for the other players. The * indicates that the strategy has a payoff that is higher than any other strategy. $s_i^*$ is a “dominant strategy” in the sense that it is a best strategy to play provided the other players play the strategy combination $s_i$.

In order to gain a better understanding of Nash equilibrium, consider the following coordination game between players with different preferences. Suppose a merger occurs between two firms that decide to operate as one larger firm. The new entity has to choose a common computer system to replace the two firms’ old systems, $I$ and $A$. Even though each firm prefers to continue using the system it used prior to the merger, both firms will be better off if they use the same system than if they continue to use different systems. To determine the Nash equilibrium in this game, we need to check each action pair (pair of numbers in a box) to determine whether each player’s action maximizes her payoff given the other player's actions. Consider Figure 3.3:

Figure 3.3: The Nash Equilibrium Solution for Strategic Games – An Example.

<table>
<thead>
<tr>
<th></th>
<th>Player 1</th>
<th></th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$I$</td>
<td>$A$</td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td>2,1</td>
<td>0,0</td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>0,0</td>
<td>1,2</td>
<td></td>
</tr>
</tbody>
</table>

234 Ibid., 64.
Consider action profile \((I,I)\). Here, neither player can increase her payoff by choosing an action different from her current choice. This means that this action profile is a Nash equilibrium. In \((I,A)\), given player 2’s actions, player 1 can choose A as opposed to I, thereby receiving a payoff of 1 rather than 0. At the same time, given player 1’s actions, player 2 can increase her payoff by choosing I rather than A. This action profile is thus not a Nash equilibrium. \((A,I)\) is not a Nash equilibrium either, since given player 2’s actions, player 1 obtains a payoff 2 as opposed to 0 by choosing I as opposed to A. Similarly, given player 1’s actions, player 2 can increase her payoff by choosing A rather than I. So neither player has an incentive deviate from \((I,I)\). Through similar reasoning, we can show that \((A,A)\) is a Nash equilibrium, since neither player can increase her payoff by choosing an action different from her current one. This game, therefore, has two pure strategy Nash equilibria, \((I,I)\) and \((A,A)\).\(^{236}\)

Harsanyi pointed out that any game that had incomplete information could be remodeled as a game of complete but imperfect information without changing its essentials. This is done by simply adding an initial move in which Nature chooses between sets of rules. In the transformed game, all players know the new meta-rules, including the fact that Nature has made an initial move unobserved by them.\(^{237}\) Consider Figure 3.4, “Follow-the-leader III”, to help in visualizing the Harsanyi transformation.\(^{238}\)

\(^{236}\) Ibid.
\(^{237}\) Rasmusen, Games and Information, 49.
\(^{238}\) Ibid., 50.
Smith and Jones are depicted as S and J in this game. Suppose that Jones does not know the game's payoffs precisely. He does have some idea of the payoffs, however, and we may represent his beliefs by a subjective probability distribution. He places a 70% chance on the game being game (a) in Figure 3.4, a 10% chance on game (b) and a 20% chance on game (c). In reality the game has a particular set of payoffs, and Smith knows what they are. This is a game of incomplete information since Jones does not know the payoffs. It is a game of asymmetric information, since when Smith moves, Smith knows something Jones does not, and it is a game
of certainty, since Nature does not move after the players do. But the game cannot be analyzed in the form shown in Figure 3.4, since Nature does not have a place in the model.

We may approach the game using the Harsanyi transformation and remodel it as in Figure 3.5, “Follow-the-leader III: After the Harsanyi Transformation.” Here, Nature makes the first move and chooses the payoffs of games (a), (b), or (c) in accordance with Jones' subjective probabilities. Smith observes Nature's move but Jones does not. When Nature moves, especially if she affects the strategy sets and payoffs of both players, it is often said that Nature has chosen a particular "state of the world". In Figure 3.5, Nature chooses the state of the world to be (a), (b), or (c).

Recall that a player’s information set at any particular point of the game is the set of different nodes in the game tree that he knows might be the actual node, but between which he cannot distinguish by direct observation. A name for this kind of information set, based on the appearance of these diagrams, is “cloud.” In Figure 3.5, we can see that nodes J2, J4, and J6 are in the same cloud, and J1, J3, and J5 are in a separate cloud. This means that Jones can tell the game has reached one of these two clouds, but he cannot “pierce the fog” to tell exactly which node has been reached. All players begin the game with the same beliefs about the probabilities of the moves Nature will make. This modeling assumption is known as the Harsanyi doctrine. So long as we follow this doctrine, the model cannot reach a situation where two players possess exactly the same information, but disagree as to the probability of some past or future move of

---

239 Ibid.
240 Ibid., 51.
241 Ibid.
242 Ibid., 40.
243 Ibid., 41.
The trouble with games of incomplete information was that we could not easily model them. The Harsanyi doctrine allows us to remodel games of incomplete information as games of complete but imperfect information.\textsuperscript{245}

Figure 3.5: Follow-the-leader III (After the Harsanyi Transformation)

\textsuperscript{244} Ibid.

\textsuperscript{245} Ibid., 51. Recall that in a game of imperfect information, one player is unaware of actions taken by other players, either in the past or contemporaneously (Dutta, Strategies and Games, 309.)
The Harsanyi doctrine thus allows us to model the professional services market. The professional-client relationship is a game of incomplete information, since the client does not know the type, payoff, or available options of the service provider. Still, the Harsanyi doctrine allows us to model this game as one of complete but imperfect information. Nature moves first, and since both the service provider and client know and have the same belief about the probabilities of Nature’s move, the game becomes one of complete information. Since service providers continue to hoard private information, information is imperfect, and the clients find themselves in Jones’ position in Figure 3.5. They can tell the game has reached one of the clouds, but they cannot “pierce the fog” and know which of the nodes has been in fact reached. Professional services are thus games of imperfect information, where the client (principal) is either unaware of, or aware but incapable of evaluating the actions taken by the service provider (agent). The Harsanyi modeling assumption here is that both service providers and their clients begin the game with the same beliefs about the probabilities of Nature’s moves, e.g. about the service provider’s type, viz. the quality of the service provider’s work and experience. This could be drawn simply from background knowledge of the frequency of various types in the population.

**Pooling vs. Separating Equilibria**

Now that we have a better understanding about modeling in professional services, I discuss two kinds of typical equilibria. In general, to resolve information asymmetries, we can give service providers incentives to do the right thing. Absent such incentives, we need to provide information refinements for clients, so that their information partitions are not perpetually coarser than that of the service providers. In this context, we want to try to fight pooling equilibria and try to create separating equilibria. In a pooling equilibrium, all service providers
send the same message, which means that the client does not receive information about the
difference between the service providers. In a separating equilibrium, the service provider sends
different messages, allowing the client to learn each service provider’s type, and discern
information from the message. I discuss these equilibria in turn for both moral hazard and
adverse selection scenarios.

\[ a. \text{ Moral Hazard} \]

In moral hazard with hidden knowledge, the agent, but not the principal, observes a move
made by Nature after the game begins. Although information is symmetric when the contract
begins, information becomes asymmetric later. The principal views agents identically, but knows
that agents develop private types midway through the game. The principal’s goal, then, is to
provide agents with the right incentives to disclose their types.\[246\] The pooling and separating
equilibria are helpful tools in this context, and they are introduced through the following
production game and a discussion of mechanism design.

The players in this production game are the principal and the agent, and the order of play is
as follows:

1. The principal offers the agent a wage contract of the form \( w(q, m) \)
2. The agent accepts or rejects the principal's offer
3. Nature chooses the state of the world \( \theta \), according to probability distribution \( f(\theta) \). The
agent observes \( \theta \), but the principal does not
4. If the agent accepts, he exerts effort \( e \) and sends a message \( m \), both observed by the
principal
5. Output is \( q(e, \theta) \)

\[246\] Rasmusen, *Games and Information*, 195.
Here, $q$ is the monetary value of output, $e$ stands for effort, $m$ for message, and $U$ is the utility function of the agent, which decreases in effort and increases in wage. $V$ is the principal’s utility, which is the difference between output and wage.$^{247}$

The payoffs of the game are as follows:

1. If the agent rejects the contract, $\pi_{agent} = \bar{U}$ and $\pi_{principal} = 0$
2. If the agent accepts the contract, $\pi_{agent} = U(e, w, \theta)$ and $\pi_{principal} = V(q - w)$

The reservation utility $\bar{U}$ is the minimum for which the agent will accept the job.$^{248}$ The principal would like to know $\theta$ so he can tell which effort level is appropriate. Ideally, she would succeed in employing an honest agent who would choose $m = \theta$, but alas, “in non-cooperative games, talk is cheap.”$^{249}$ Since the principal cannot trust the agent, she must instead design a contract that either provides incentives for truthfulness or takes the agent’s lying into account. The mechanism design literature addresses this question. Here, the principal needs to implement a mechanism to extract information from the agent.$^{250}$ In hidden-knowledge moral hazard models, the principal tries to make different actions attractive under different states of the world, so the agent's choice depends on those hidden states. If this were a hidden-action moral hazard model, the principal would try to construct a contract that will induce the agent to take the single appropriate action.

If all types of agents pick the same strategy in all states, the equilibrium is pooling. Otherwise, the equilibrium is separating.$^{251}$ A model might have multiple Nash equilibria, some

---

$^{247}$ Ibid., 169.
$^{248}$ Ibid.
$^{249}$ Ibid., 196.
$^{250}$ Ibid.
$^{251}$ Ibid.
pooling and some separating. A single equilibrium can include several contracts, but if it is pooling, the agent always uses the same strategy, regardless of type.\textsuperscript{252} In a principal-agent model, the principal’s goal is to design the contract so as to achieve separation, unless the incentives turn out to be too costly. An equilibrium is \textit{fully revealing} if the agent's choice of contract always conveys his private information to the principal. Between pooling and fully revealing equilibria are the imperfectly separating equilibria, also called semi-separating, partially separating, partially revealing, or partially pooling equilibria.\textsuperscript{253} The principal then attempts to maximize his profits subject to two constraints: (1) incentive compatibility (the agent picks the desired contract and actions), and (2) participation (the agent prefers the contract to his reservation utility). In a model with hidden knowledge, the incentive compatibility constraint is often called the self-selection constraint, since it inspires different types of agents to pick different contracts.\textsuperscript{254}

Applied to the professional service market, where moral hazard prevails, the principal (the client) tries to make different actions attractive for the agent (the service provider). If all service providers pick the same strategy, the equilibrium is pooling, otherwise, the equilibrium is separating. When thinking about remedying moral hazard in a professional-client relationship, our goal is to design a contract that achieves separation, i.e., a contract that allows the client to observe and monitor the different actions and effort levels of the service provider. In a fully revealing equilibrium, the service provider’s action always reveals his “private information,” viz. the information that the client is not in a position to monitor, to the client. The client then attempts to maximize his profits subject to two constraints: (a) incentive compatibility (the

\textsuperscript{252} Ibid., 197.
\textsuperscript{253} Ibid.
\textsuperscript{254} Ibid.
service provider accepts the desired contract and actions), and (b) participation (the service provider prefers the contract to his reservation utility\textsuperscript{255}). Following Mirrlees, the incentive compatibility constraint is the self-selection constraint, since it inspires different types of service providers to pick different contracts.\textsuperscript{256}

\textit{b. Adverse Selection}

Signaling is a way for an agent to communicate his type under conditions of adverse selection. Spence\textsuperscript{257} was the first to show how signaling could emerge as a solution to the asymmetric information problem introduced by Akerlof. The signaling contract specifies a payoff or a wage that depends on an observable characteristic, i.e. the signal, which the agent chooses for himself after Nature chooses his type. If the agent chooses his signal before the contract is offered, he is signaling to the principal. If he chooses the signal afterwards, the principal is screening him.\textsuperscript{258} See representative game trees (d) and (e) in Figure 3.2.

When the hidden characteristic is a quality variable, and quality can be either good or bad, producers of high-quality products have an incentive to signal their quality. But in order for the signal to be credible, the low-quality supplier should not be able to gain by transmitting the same signal as the high-quality supplier. Since signaling typically consumes resources, it is substantially more costly for the supplier of the low-quality commodity to transmit the same signal as the supplier of the high-quality commodity. This means that under the right conditions, the additional cost of strong signaling dissuades those who have only low-quality items to sell.

\textsuperscript{255} In other words, the service provider decides that she is better off entering that contract, as opposed to doing nothing, given the circumstances.
\textsuperscript{256} Rasmussen, \textit{Games and Information}, 197.
\textsuperscript{258} Rasmussen, \textit{Games and Information}, 249.
By sending a weaker signal, the low-quality producer in effect reveals his type. However, because the signal imposes real costs on the individual and on society, truthful revelation comes at a price. That is, the resources consumed in signaling by high-quality producers sometimes do not provide any direct utility. In fact, there is a range of signals consistent with equilibrium, and often the same result could have been obtained with a lower investment in signaling. Sometimes everyone invests in signaling but the signaling doesn’t distinguish the high quality products/services from those of low quality.  

Firms with introductory offers at very lower prices also attempt to signal, but there is a danger that the consumer will interpret the low price as evidence of low quality (as opposed to producer confidence). Banks signal their trustworthiness through the size of their edifices, guarantees signal the firm's confidence in the quality of its products, and owner-managers of firms signal their confidence that the firm is not overvalued by restricting the number of their own shares that they sell. Sellers of goods and services have had to think not just about how to act to make profit, but also about how their actions will be viewed and interpreted. Thus, it is clear that the inferences and beliefs of clients (and other outside players) do matter.

The standard example is educational pre-requisites and certifications. Here, the signal is the amount of training that an individual has undergone prior to applying for a position. Training is expensive and individuals who know themselves to be innately intelligent and hard working, relative to others, often find the process less taxing and they are more likely to graduate and be certified. Thus, educational institutions contend that the population of graduates contains a

---

260 Ibid., 292.
disproportionately high number of individuals who are intelligent and hardworking and therefore productive.

Educational prerequisites can thus be used to sort workers into $H$ types, who have relatively high ability and productivity, and $L$ types, who have relatively low ability and productivity. Here, $H$ types are those who get high marks in the training program, while $L$ types are those who do not perform well in the program. High quality ($H$ type) workers can generate substantially more profit than low-quality workers ($L$ types), since $H$ types are innately more productive. Although individuals make choices in their lives that help to determine their personal type, at the time an employer makes a hiring decision (or the profession decides to admit an applicant), the worker’s type has already been determined. This means that the employer faces a hidden characteristic problem. To attract $H$ types we can offer a higher wage: $W_H > W_L$, where $W_H$ is a wage paid to $H$ types and $W_L$ is the wage paid to $L$ types. But we cannot expect the $L$ types to truthfully reveal themselves and claim the lower $W_L$. Since production takes place in settings that make it difficult to identify the contribution of individuals in the short run, it is often not possible to directly separate the $L$ types from the $H$ types at the time of hiring them.

Now suppose it costs $H$ types $C_H$ dollars to graduate from college and it costs $L$ types a higher amount, $C_L$. $C_H < C_L$ since $L$ types may need extra semesters to graduate and they have to work much harder to get admitted and graduate from college. Thus, in spite of their preference for anonymity, it is possible to induce the $L$ types to identify themselves: the firm could pay a

---

262 Campbell, Incentives: Motivations and the Economics of Information, 291.
263 The payment to the worker is called wage here, but it is in fact the present value of expected compensation including benefits over the lifetime of the job.
264 Campbell, Incentives: Motivations and the Economics of Information, 291.
265 Ibid.
salary $W_H$ to applicants with graduation certificates, and a salary $W_L$ to workers without a certificate. So long as $W_H - C_L > W_L$, the $L$ types will not pay the cost $C_L$ necessary to obtain a certificate. It would be more advantageous to obtain the lower salary $W_L$ without the additional education required to qualify for the higher wage. If $W_H - C_H > W_L$ then $H$ types will incur the cost $C_H$ of obtaining a graduate certificate, obtaining the higher net salary $W_H - C_H$. Both these conditions hold so long as $C_H < W_H - W_L < C_L$. Assume $W_H = 1000$ and $W_L = 600$, then $C_H < 400 < C_L$ is required for signaling to reveal a worker’s characteristic. This gives a range of equilibria, many of which will be inefficient because the signaling could be done at lower cost to the individual and to society.\footnote{Ibid.}

We thus have two kinds of equilibria in this scenario. In a \textit{pooling} equilibrium, the two types of workers get the same amount of education, and each receives the same pay, namely the weighted average marginal product,\footnote{The marginal product of labor is the change in output from employing an additional unit of labor} weighted by the proportion of each type in the workforce. In a \textit{separating} equilibrium, the $H$ type gets more education than the $L$ type, and the firm pays more to the workers with the higher level of education.\footnote{Campbell, \textit{Incentives: Motivations and the Economics of Information}, 294.} A separating equilibrium must satisfy the self-selection constraint, that is, each type must find it advantageous to send a signal that is different from the one transmitted by the other type.\footnote{Ibid., 296. Spence (1973) was the first to show how signaling could emerge as a solution to the asymmetric information (hidden characteristic) problem introduced into the literature by Akerlof (1970). In separate contributions, George Akerlof, Michael Spence, and Joseph Stiglitz showed that the presence of asymmetric information in real-world markets required a new way of modeling economic exchange. These authors were awarded a Noble prize in economics in 2001 for their analyses of markets with asymmetric information.}

This discussion is applicable to the relationship between professional service providers and their clients as well. Assume that low-quality service providers will not directly reveal their identity, and only individual $i$ knows the risks posed by $i$ to the client. Due to a lack of
experience, lack of natural talents, or some other quality deficiency, service provider i poses a risk to her clients, unlike other service providers who do not share those quality deficiencies. There are only two possible equilibria, one in which every service provider charges the same for services, and one in which the different types charge different prices. A pooling equilibrium is one in which the same contract is obtained by both risk categories. A separating equilibrium is one in which the $H$ types are separated out, by providing the $L$ types with a contract that is less desirable to the $H$ types, than the contract designed for them.

The Revelation Principle

The principal may offer a contract that induces the agent to be untruthful in equilibrium. Thus, it is possible for the principal to take lying into account in the design of the contract. The revelation principle states that for every contract $w(q, m)$ that leads to lying (that is, to $m \neq \theta$), there is a contract $w^*(q, m)$ with the same outcome for every $\theta$, but with no incentive for the agent to lie. Here, $q$ is the monetary value of output, and $w(q, m)$ is a wage contract. In many contracts, it is profitable for the agent to send false messages since, when the state of the world is $a$, he receives a reward $x_1$ for the true report of $a$ and $x_2 > x_1$ for the false report of $b$. The principal can offer the agent a contract that gives her a reward of $x_2$ regardless of whether she reports $a$ or $b$, which would lead to exactly the same payoffs for each player while giving the agent no incentive to lie. According to the revelation principle, a contract with no lying can always be

---

271 Ibid., 312-3. Thus, one of the downsides of professionalization is that it may sometimes block strategies for wage and contract design by enforcing uniform prices for services that are not always uniform.
272 Rasmusen, *Games and Information*, 198.
found by imitating the relation between states of the world and payoffs in the equilibrium of the contract with lying.\textsuperscript{273}

For example, if we are worried about people with an annual income of $70,000 lying on their income taxes, we can charge them the same tax as the ones who make $50,000. Similarly, the mother who agrees never to punish her daughter if she tells her all her escapades will never hear any untruths.\textsuperscript{274} While the revelation principle implies that a truth-telling equilibrium exists, it does not guarantee that the equilibrium is unique. It is possible for the equilibrium to be a weak Nash equilibrium in which the optimal contract gives the agent no incentive to lie but also no incentive to tell the truth. If agents can derive the slightest utility from telling the truth, then truth telling becomes a strong equilibrium. If the agent’s utility from telling the truth is really significant, it should be made an explicit part of the model. Once the utility of truth telling is higher than a certain threshold, in fact, agency problems and the costs associated with them disappear.

Thus, although the literature on agency theory focuses for the most part of providing material incentives as a solution to information asymmetries, this observation regarding the value of truth-telling points towards non-material incentives (e.g. personal conceptions of the good, loyalty to a team or to a principal, trust and collegiality, belief in the purpose of an organization, etc.) for resolving inefficiencies arising from information asymmetries in principal-agent relationships. As Rasmusen puts it, “this is one reason why morality is useful in business.”\textsuperscript{275} Professional morality thus can have an important role in designing efficient contracts between service providers and their clients. I elaborate on this point in the next two chapters. Still, as we

\textsuperscript{273} Ibid.
\textsuperscript{274} Ibid.
\textsuperscript{275} Ibid., 199.
have seen, morality is one element alongside wage levels, educational requirements, quality standards, etc., all of which constitute the structure of professional associations.

So far, I have explained the mechanics of information asymmetric markets, and discussed typical problems arising under information asymmetry, namely moral hazard and adverse selection. I explored the Harsanyi doctrine, the pooling and separating equilibria, and the revelation principles as strategies for overcoming information asymmetries. Now that we have a better sense of the nature of information asymmetries and strategies for responding to them, we can see that denying market failures, as some theorists have done, is not feasible. Still, as we saw in the discussion of private interest theory in Part I, professional regulation can give rise to a number of serious issues (e.g. anti-competitiveness, cartelization, inefficiency, etc.).

3.3.2 Conclusion

There is a dispute between public and private interest theories of regulation. Each stance points to the descriptive practices of groups who draw on professional trappings to argue for an opposing position. Professional regulations operate as the rules that guide and govern professional activity, and appear to be used to raise or decrease Pareto efficiency in the market for professional services.

The descriptive picture of the work of professionals reveals evidence that can support two contrasting implicit moralities of the professional market. Professions might be cartelization mechanisms that promote the interest of their members, as the ultimate purpose they were designed to fulfill. The descriptive findings of private regulation theorists provide some damning evidence for this stance. In that case, the fact that professional mechanisms arise to respond to
information asymmetries when markets fail, and the fact that those institutional responses are Pareto-improving and save lost transactions, are lucky accidents or random deviations from the implicit private interest normative characterization of professions.

Conversely, the evidence presented in this chapter could support a different implicit professional morality: that professions are Pareto-improving, trust-creating, trust-preserving institutional mechanisms that strive to promote mutual benefit and save lost transactions in the market for professional services. The empirical findings of public regulation theorists provide evidence for such an implicit morality. In this case, the cartelizing, rent-seeking evidence regarding professional associations are distortions of professionalism.
4 The Mystery of Professions

In this chapter, I describe trust as the “mystery of professions” which distinguishes professional associations from other private economic entities. I begin by making the descriptive background assumption that the nature of professions and professional regulations is, at least to some extent, socially beneficial. Based on this assumption, I present a normative model of professions that defines professions as trust-creating, trust-preserving institutional structures that respond to information asymmetries, with the goal of increasing Pareto efficiency in the market for professional services. I then compare my definition of professions with some of the dominant proposals reviewed in my first chapter, and show that the information asymmetry approach has superior explanatory power in comparison to alternative definitions.

4.1 Professions as a Private Economic Entity

Broadly construed, the market is a social sphere where it is not always possible for economic actors to trust each other. Markets are designed to create social welfare despite the absence of the relationships of trust that guide private interactions with family members, friends, religious groups, or national communities. Contracts, for example, typically presuppose a deficit of trust (or a lack of trust altogether, according to neoclassical economics). To be sure, many contracts would be rendered useless if business partners could trust each other to follow through on the terms of their interaction without having signed on the dotted line.
Take, for example, the experience of buying a used car. Buyers know that trust is significantly diminished in this environment, and for the most part, take it for granted that sellers are seeking only their own interest. Sellers are not trusted to voluntarily reveal the history of accidents, all the relevant defects of the car, or typical issues with the car’s brand. Sellers are also assumed to be seeking the highest sale price, regardless of whether the car is the right purchase for buyers. This is a structural feature of the used car market. Similar to the market for used cars, lawyers offer their services in the private market for goods and services. Yet, client interactions with lawyers are markedly different from those in the used car market – we generally have an expectation that we can trust our legal counsel. Such an expectation of trust applies to many professional services. The question is: why isn’t every market transaction like buying a used car? This is the mystery of professions. What needs to be explained here is the persistence of high-trust behavior among certain private economic actors but not among others.

According to Fukuyama, trust is the “expectation … of regular, honest, and cooperative behavior, based on commonly shared norms,” on the part of members of a given community. “Regular,” here means predictable, and Fukuyama adds that regularity is a sufficient condition, since in many societies the expectation is that people will regularly cheat. Such cheating would lead to a “deficit of trust” since it is “regular but dishonest.” The commonly shared norms are the element that makes this definition more technical. For my purpose, shared norms are important for professionalism. It is not the fact that the doctor promised to cure me or that I signed a contract with the engineer that makes me trust them. Rather, the norms guiding the behavior of professionals are shared, such that we come to have certain standardized expectations about how these interactions are supposed to go. Crucially, the norms guiding the

2 Ibid.
professions are typically supposed to help us trust that professionals will serve our interests as clients.

From the doctors’ white coats and charts, to the lawyers’ suites, desks, and file folders, there are various controlled circumstances that professional settings recreate on purpose, in order to underscore the presence of regular and expected norms that promote the interests of clients. The norms are cooperative in the sense that they are mutually beneficial, providing specialized services to clients, while earning social and economic gains for professional service providers. Of course, the emphasis on trust varies across professions, e.g. between doctors, accountants, lawyers, nail salon operators, plumbers, auto mechanics, and pet shop aquarium tank operators. We don’t trust all professionals in the same way or to the same extent. But the core values represented by professionalism are to a large extent shared, such that, at least by definition, we are supposed to trust that our interests are looked after when we consult legitimate professional services.

The mystery of professions, then, is that they seem to involve more trust than other private economic organizations. So why do suppliers of professional service go out of their way to cultivate trust, even though the private economic realm is designed to function with a deficit of trust? The scope of the mystery does not require a general theory of social trust, but rather, a focused explanation for why we rely on trust to explain interactions with professional economic entities, but not with other private actors in our everyday market interactions.

As I showed in detail in my first three chapters, the cooperative feature of professions and the professional commitment to promoting the public interest has been supported descriptively, and has received endorsement from public interest theorists of professional regulation. However, there is considerable disagreement over the empirical evidence supporting this view. As
Waddington has suggested in his discussion of professional ethics, “it is perhaps time we were a little less trusting” of the professions and the professional claim of being moral. Along the same line, the descriptive account of cartelizing professionals and the private interest theory of regulations discussed in previous chapters warn about believing the claims of professionals to serve the public interest. Resolving the empirical (ongoing) debate over the public and private theories of professional regulation and the nature of professional work is beyond the scope of this dissertation. Still, since my goal is to provide a reconstructive normative model of professionalization, I needed to engage with the descriptive debates that explain how professions work.

At this stage, I introduce a core assumption in order to help guide the remainder of my analysis in the dissertation. The assumption is as follows: at minimum, professions serve a social function. I do not deny that empirically speaking, individuals that call themselves professionals and are members of professional associations may fail to provide socially beneficial services. Still, I assume that going back to the guilds and throughout the history of the professions, professionals provided socially beneficial services, at least to some extent, even if merely as an extension of being part of an institutional model that was originally designed to be socially beneficial. And viewing trust as a descriptive feature of professions has some intuitive appeal. Consider some of the most highly populated professional associations: doctors, surgeons, teachers, engineers, and lawyers – they all work within broadly speaking trust-worthy institutional settings. Within my reconstructive model, making this descriptive assumption about the socially beneficial nature of professions corresponds to a certain implicit morality of the professions and a certain normative model. My analysis in the previous three chapters provided

---

relevant empirical evidence, about what professionals do and what they have historically done, that go towards supporting this stance.

But why do professional associations rely on trust? Why not stick to contracts, and try to design more intricate performance-pay incentives within existing contracts, for example? I propose that trust is a governance mechanism, in so far as it governs and influences professional conduct. Governance can be thought about as the set of institutions that coordinates an organization’s activities, and establishes and monitors rules and guidelines. Professional ethics codes are one example of a trust-creating governance mechanism, that is, of an institution that facilitates and coordinates professional behavior, for example, through imposing an ethos of service on professional service providers. The price mechanism (i.e. the market) is another social institution, which governs professional behavior, for example, through competition and reputation-building. Governmental intervention is yet another social institution, which can influence professional action, for example, through imposing provincial or federal regulations for professionals.

Which governance mechanism is superior? One way to choose is based on the relative costs of the governance mechanisms, i.e. in terms of their transaction costs. The lower the transaction costs are for a given governance mechanism, the more resources can be freed and used for other purposes. Therefore, economic decision-makers are likely to choose the governance mechanism with the lowest transaction costs. For example, a doctor-patient relationship can be governed using a variety of institutional governance mechanisms. One governance mechanism may involve imposing an ethics guideline (enforced though collective

sanctions) that induces doctors to act in the best interest of their patients. Another mechanism may be to pass a law that obliges doctors to promote their patients’ welfare (although this may impose hidden costs on patients). A third option is to leave things to the market, and expect competition and reputational information to ensure that doctors act in the interest of their patients (although this may impose undue costs on patients given their lack of technical expertise).

As Arrow has shown in his work on institutional economics, non-market governance institutions tend to arise in the presence of market failures. As he puts it: ”. . . when the market fails to achieve an optimal state, society will, to some extent at least, recognize the gap and nonmarket institutions will arise attempting to bridge it.”6 Trust-creating professional institutions are one such non-market oriented institutional governance structure. For example, since the lawyer-client relationship is subject to information asymmetries, the Canadian Bar Association imposes a Code of Professional Conduct as a core governance structure. The information asymmetries are too extensive for outsiders (whether they are lawyers from other Bars or laypeople) to properly police or enforce such codes of conduct – this is why the Canadian Bar Association is in charge of detecting violations and implementing the codes. In this scenario, the code of conduct imposes lower relative transaction costs in contrast with other governance institutions, and works in conjunction with them. The presence of ethics governance mechanisms, going back to the Hippocratic oath in the case of physicians, creates trust in the services of professionals.

---

Trust and trust-creating professional mechanisms such as codes of ethics can thus be thought about as a non-market oriented institutional governance mechanism that guides professional behavior in the presence of market failures. As I explained in the previous chapter, market failures occur when market prices deviate from marginal social costs due to the presence of monopolies, externalities, and/or transaction costs. These market failures may occur within principal-agent hierarchies, despite market solutions and government intervention. Government institutional mechanisms can fail to appropriately correct market failures, for example, by misdiagnosing problems or being overly costly to implement. Market failures have been written about extensively, and a growing literature documents government failures in institutional governance. But in contrast to market and government failures, there is a deficit of attention to the role of trust and institutional mechanisms like codes of ethics in institutional economics. Trust is an attractive governance mechanism whenever it imposes lower relative transaction costs than alternative institutional structures.

In the next section, I elaborate on the implications of defining professions in terms of trust-preserving institutional mechanisms.

---

7 See Thomsen “Business Ethics as Corporate Governance.”
4.2 Professions as Trust-preserving Institutional Mechanisms

I begin by reminding the reader about the nature of typical market failures arising from information asymmetries in the market for professional services.

4.2.1 Professionalization as a Response to Moral hazard & Adverse Selection

In the market for professional services, the effort level exerted by professional service providers (agents) cannot always be adequately monitored or evaluated by non-specialist clients (principals). (See Figure 4.1 and 4.2) The professional and her client thus have symmetric information when they agree to the contract, but Nature makes a move unobserved by the client, and information becomes incomplete. Recall that Nature is a pseudo-player who takes random actions at specified points in the game with specified probabilities. A move by Nature is called a state of the world. In this context, the state of the world dictates that certain information about the professional’s effort levels and decision-making will be hidden from the client. This state of the world is observed by the professional. The professional then makes a move based on Nature’s move. For example, she may exert a low level of effort.

Figure 4.1:

Moral hazard hidden information

[Diagram showing the interaction between the client, professional, and Nature, with arrows indicating the flow of effort and decision-making.]
Since individual service providers may not have the incentive to exert high effort, regulatory interventions can lead to Pareto improvements and enhanced welfare by imposing incentive and monitoring solutions. Insofar as professionals work to build a reputation for themselves as high-quality service providers, we may characterize professional reputation as a form of incentive-building exercise. Other incentive-building elements include standards of practice, and avenues for disciplining service providers and for receiving client (or third party) evaluation. Codes of ethics and institutional norms can also be drawn on as non-market solutions for responding to moral hazard. These are some of the ways that the structure of particular professions is designed as a response to moral hazard problems in the market for professional services with the goal of better serving the public interest. I elaborate further on these features of professions below.

The problem of information thus arises because the effort level and service quality of professionals (agents) cannot always be adequately monitored or evaluated by non-specialist clients (principals). The professional’s quality of service, effort levels, and decision-making will be hidden from the client. In the case of moral hazard, the core problem is that service providers

---

may not have the incentive to deliver the right kind of services to clients. In this scenario, there are different routes to resolving the deficit of incentives. The standard information theoretic solution to moral hazard problems is to impose incentives and monitoring solutions. Left to the market and the price system, professional reputation could provide such incentives and lead to Pareto improvements.

For example, as a chiropodist, I might rely on customer reviews and social media ratings to build a reputation for myself. This reputation-building exercise provides me with the right kind of incentives, so as to keep effort levels and standards of service high, even when my clients are not in a position to evaluate the quality of my services. In my capacity as a chiropodist, I might enter into a special agreement with a pharmaceutical supplier to provide an injection that costs less than competing products, passes existing safe-guarding tests, and yet poses potential long term health problems. Clients may not have access to relevant information to evaluate the use of this injection, and even when things backfire and word gets out about the problems with the injection, it might take a while for the reputational impact of this malpractice to really impact my business.

Another solution, following the three-part governance model I laid out at the outset of this chapter, is for the government to update legislation and ban the aforementioned injection. Recognizing that chiropodists have an incentive problem and might be prone to using this product as a cost-cutting measure (assuming a private health care system), the government might ban the commercial use of this product. While this solution boosts social welfare, it only does so to some extent, since drug regulatory agencies are always one step behind in discovering entrepreneurial products and strategies that save businesses money but pose harm in one way or another.
A third solution is to draw on ethics as a solution to moral hazard. Specifically, professional mechanisms such as codes of practice\textsuperscript{11} and peer evaluations are examples of professional institutions that give clients an assurance that they can trust their service providers. It is possible to think of circumstances where the questionable injection is used by some service providers, without adverse affects on the reputation of the operator, at least during the preliminary period of use. The legislative bans on such a product might be in progress, but as it often does with regulation, it might take a prolonged period of time before the product is banned from use. In such a scenario, it is conceivable that codes of practice that expressly dictate the obligation to promote the interest and long-term safety of clients, in conjunction with peer evaluation by other professionals who have market knowledge about products and ingredients might discourage chiropodists from taking advantage of moral hazard scenarios. In contrast with other private economic actors, professionals have excelled at drawing on this final trust-preserving governance mechanism.

Consider another example. Teachers are the most populated profession in Ontario, with over 234,000 members.\textsuperscript{12} When it comes to making decisions about mathematics curriculum, for example, they may have an incentive to choose adjustments based on the ease of grading, regardless of the overall impact on the students. The students and their parents, on the other hand, may not be in a position to properly evaluate the curriculum adjustments. What kind of governance structure should we use when it comes to resolving this moral hazard problem?


\textsuperscript{12} Ibid., “Number of Members in Regulated Professions.”
Teachers in one province, say Ontario, might consider performance-ranking competition (with other provinces, e.g. Alberta) as their key guidance for deciding about curriculum. If there is evidence that Ontario students rank higher than Alberta students under the new scheme, then arguably the market governance structure would not be critical of (and would in fact support) lighter grading load for teachers under the new scheme. Beyond competition, there is specific provincial legislation (the Ontario College of Teachers Act\textsuperscript{13}) guiding curriculum development. The act specifies the frequency and structure of adjustments permitted to curriculum, and may include safeguards that ensure an overall positive impact on student learning. Finally, as professionals, teachers have access to a set of trust-creating institutional structures that go beyond market solutions and government responses – the “ethical standard”\textsuperscript{14} for members of the teaching profession. For example, the ethics standards of “care” require teachers to have “compassion, acceptance, interest and insight for developing students' potential.”\textsuperscript{15} Such a standard of care may or may not lead to the highest math test-scores, but is a governance mechanism in curriculum design within the teaching profession.

Beyond moral hazard, professional services are also prone to another information problem: adverse selection. Here, the client (principal) contracts to receive a service from the professional (agent), but the quality of the service is non-contractible. While the professional knows her own abilities from the start, the client does not have this information. If the professional can send a “signal” to the client about her type, and if she sends the signal before the principal offers the contract, then the model is signaling. Again, the three governance mechanisms (market, \textsuperscript{13}“Ontario College of Teachers Act,” 1996, chapter 12: \url{http://www.e-laws.gov.on.ca/html/statutes/english/elaws_statutes_96o12_e.htm} (Accessed Feb 22, 2013).
\textsuperscript{15}Ibid.
government, and non-market – e.g. ethics) may be drawn upon to respond to the information asymmetries and make Pareto-improvements, independently or in conjunction, depending on relative transaction costs.

Since individual professionals may not signal their type to clients, regulatory intervention can lead to Pareto-improvements and enhanced welfare, by imposing certification or minimum-quality standards (such as licensing) on professionals. Professionalization, and the whole process of forming professional associations, can thus be explained as a self-selection, i.e. signaling process, whereby professionals try to communicate their type to clients prior to entering into contracts with them. The complex and onerous educational and experience requirements of becoming a professional, alongside standards of practice and codes of ethics, can also be viewed as attempts to signal high professional quality to the public. Clients are often not in a position to properly evaluate which professional services are adequately specialized and trustworthy. The market solution to this problem is to allow professionals to practice without any entry requirements. The quality of the services would establish a reputation for each service provider, and competitors in the market would push upwards on the standards of practice. This is a governance mechanism that works for occupations and business entities to a large extent. Another governance mechanism is to draw on government legislation that imposes certain duties on all accountants and chiropodists with penalties for failure to abide. Provincial and federal legislation prohibit individuals from providing services without the appropriate pre-requisites and without the required examination scores. Finally, these professional bodies use education and rigorous testing and examination procedures as a trust-creating mechanism that responds to adverse selection.
Examination and testing are core non-market professional institutional structures that respond to adverse selection problems. In Ontario, for example, most professions require applicants to successfully pass an exam; 19 professions require applicants to successfully pass more than one exam; of this 19, 4 require applicants to successfully pass 4 exams; and 8 require some applicants (typically from outside the province) to successfully complete an eligibility exam first, in addition to all the other exams. These eligibility exams are designed to determine the applicant's readiness to write the registration/certification/qualifying exam or to meet regulatory requirements.\textsuperscript{16} Testing and entry/eligibility examinations usually address adverse selection problems, while licensing, minimum-quality standards, and peer-evaluation are typically used to police quality of service after professionals and clients enter into an agreement, to deal with moral hazard information problems. For example, Ontario chiropodists are required to take two exams – a “registration” exam that “combines theory, jurisprudence, ethics, and professional practice”, alongside an “objectively structured clinical exam” and a “clinical evaluation”. Chartered accountants have to pass four sets of tests: “uniform evaluation,” “core-knowledge exam,” “school of accountancy exam,” and a “chartered accountancy reciprocity exam.”\textsuperscript{17} These exams are designed to ensure that the right kind of candidates with the desired set of characteristics enter the profession.

\textsuperscript{17} Ibid.
4.2.2 The Definition of Professions

As I have shown in the previous section, each of the non-market/government/trust governance structures poses a different solution to the problem of information. I noted at the outset of this chapter that the mystery of professions was that they drew on trust, unlike other private economic actors that operated with an assumed deficit of trust. Professionals typically draw on trust since given the applicable information problems, this governance mechanism has proven transaction cost-minimizing. Professions can thus be defined as trust-creating, trust-preserving institutional structures that are guided by norms and regulations to respond to information asymmetries, with the goal of increasing Pareto efficiency in the market for particular services. In the face of moral hazard and adverse selection information problems, the professional institutional structure augments the service provider’s incentives and moral commitments to provide quality service, and communicates information about service quality to clients.

As we have seen in the previous chapters, however, some have used the title “profession” and drawn on the professional institutional structure for anti-competitive, rent-seeking purposes that benefit the professional membership at the expense of overall Pareto efficiency in the professional service market. According to my approach, this is a distortion of professionalization, and a deviation from the socially beneficial purpose of professions in society, since the key normative essence of professionalization is the creation and preservation of an environment of trust, and a cooperative ethos that guides professionals towards serving a positive social function.

Recall from chapter 1, the list of characteristics that have been identified over the course of the past century as definitive features of professions. I now provide a comparison between my
definition and some of the alternative definitions provided in the literature, in an effort to demonstrate the explanatory attractiveness of my approach. The most popular characteristic on the list was the possession of special, technical, and intellectual skills, via a specialized body of knowledge (whether implicit or explicit). But in comparison to my definition, specialized knowledge remains a superficial trait. The underlying problem in the professions is information asymmetries, since structural features make it difficult to adequately observe or evaluate the work of professionals. Even if technical skills were universally available and shared by doctors and patients alike, a host of moral hazard problems related to quality evaluation and effort-policing would still arise that prevent non-professionals from assessing due diligence. Experts with specialized knowledge ultimately have finer information partitions than clients and independent parties.

Some might argue that professional governance mechanisms such as educational prerequisites and testing distinguish between the skilled and the unskilled, and that technical skills are thus ultimately the defining feature of professions. But while examinations control the quality of professional services by keeping out the unskilled and the under-qualified, the underlying reason for doing so can be better explained via information asymmetries. Technical skills are only part of the story. Because of adverse selection information problems, professional associations do not have access to relevant information that lets them successfully distinguish between low-level and high-level candidates. Professional tests and the educational prerequisites are designed as they are because those mechanisms weed out candidates with lower levels of commitment and talent. But further, and more importantly, educational prerequisites inculcate the norms of the profession and the ethos of service-orientation in the new professional candidates. These values and norms are important, because of the prevalent information asymmetries that professionals might otherwise take advantage of for their own benefit.
Another popular defining feature of professions is autonomy. Unlike ordinary occupations, a professional designation earns practitioners institutional and self-regulatory autonomy. Only professionals (as opposed to non-specialist government bodies or the public at large) are typically allowed to make decisions about the nature of professional work or provide professional evaluation and guidance. The intricate educational pre-requisites required for entering into many professions are also autonomously designed by professionals themselves, unlike public education and general university degrees that are usually connected with municipal, provincial, and federal educational mandates.

But autonomy is also not an intrinsically valuable trait. In so far as autonomy is a typical characteristic of the structure of professional work, it becomes noteworthy through its role in information asymmetries surrounding professions. The moral hazard problems surrounding professional work make it difficult for non-professionals to adequately appreciate the intricacies of professional work, institutional design, and regulatory needs. But although professionals exercise autonomy, they are still open to scrutiny. As I have shown, trust and professional ethics governance mechanisms have historically played an important role in reducing information asymmetries in the professions. Although non-professionals do not judge this process, professional ethics regimes involve extensive peer-evaluation by other professionals. Professional ethics has traditionally been viewed as a mechanism for ensuring that professionals provide competent service and are trustworthy. My earlier discussion of professional ethics by Carr-Saunders and Wilson, in chapter 1, illustrated this point. However, empirical evidence of the development of ethics codes in the professions suggests that peer-relationships among professionals are at least as important as the relationship between professionals and clients. Going back to Percival’s canonical text from 1803 (and other early writings in the nineteenth
century on this subject)\textsuperscript{18} the focus of professional ethics has been to a large extent on the rules that regulate relationships among physicians.\textsuperscript{19} Thus, although professionals have autonomy, underlying this autonomy is a long-standing systematic process of peer-control, as part of a system of professional ethics that ultimately responds to information asymmetries.

Among the prominent traits for defining professions, authority or power has been a front-running contender. Professionals draw on the rational and scientific features of their technical knowledge to exert legitimacy and authority. Power is often a target of criticism, since inflated professional remuneration and social stature are staked through professional authority. But the underlying worry here is not the mere fact that professional make a good living or hold socially desirable positions. Recall Smith’s view, discussed earlier in the dissertation: “We trust our health to the physician … our fortune and sometimes our life and reputation to the lawyer and attorney. Such confidence could not safely be reposed in people of a very mean or low condition. Their reward must be such, therefore, as may give them that rank in the society which so important a trust requires.”\textsuperscript{20} Smith conjectured that attractive pay and social status were what we necessarily need to offer individuals we have to trust with performing some of the most vital services to the community. The worry from critics of professional power, instead, is that individuals might act against the implicit morality of the professions, and fail to make mutually beneficial Pareto-improvements through their professional activity. Thus, again, professional power is not intrinsically problematic. The real problem is that by drawing on professional power, practitioners might get away with or legitimize their failure to respond appropriately to information asymmetries.

\textsuperscript{18} See a review of this literature in Waddington “The Development of Medical Ethics – A Sociological Analysis.”
\textsuperscript{19} Leake, \textit{Percival's Medical Ethics}; Banks, \textit{Medical Etiquette}.
In so far as professionals draw on their social prestige and economic power to promote their own interests at the expense of social welfare, this is a distortion of professional power as opposed to a trust-creating institutional mechanism. The same holds when professionals use their professional legitimacy to justify market allocation mechanisms and anti-competitive behavior such as cartelization. Ultimately, the purpose of the market for professional services is best understood as creating trust through Pareto-improving institutions, and any feature of professions that abuses information asymmetries in conflict with this purpose is a distortion of professionalism.

Finally, ethics codes are another popular trait for defining professions. A large body of literature highlights the altruistic motivations of professionals in promoting the public interest as evidenced by the widespread use of ethics codes. But, in fact, ethics codes are not an intrinsic feature of professions, and they can also be better appreciated through information asymmetries. Since public and non-specialist outsiders are not in a position to properly examine or evaluate professional work or commitment to service, professional associations have historically drawn on an ethos of public service as an institutional mechanism to overcome this information asymmetry and to communicate their positive social value to the public. Additionally, codes of ethics have traditionally drawn on peer-assessment and peer-enforcement as a form of social control. Codes of ethics and norms like service-orientation matter because they govern professionals through formal and informal mechanisms to respond to information asymmetries at (often) transaction-cost-minimizing rates. Critics of professions have pointed out that professionals are “wolves in sheep’s clothing” who take advantage of the appearance of being ethical and promote their own interest at the expense of the public interest. In so far as this true, and in so far as information asymmetries make it difficult to detect such behavior, this would constitute a violation of the implicit morality of professions and a distortion of professionalism.
Widespread disagreement persists in the literature on professions over the professional feature that qualifies as the fundamental defining trait. Power, knowledge, codes of ethics, autonomy, etc. can each explain some set of scenarios well, but fall short of explaining all others at the same time. These characteristics do have one thing in common, however: they are all a response to the possible lack of trust given the prevalence of information asymmetries in the market for professional services. The basic market economy does not readily provide us with adequate trust, and hence we use educational certification, testing, ethics codes, peer-evaluations, etc. as mechanisms for achieving trust and responding to information asymmetries. By creating a reputation for themselves, and communicating their potential for social utility to the public, professionals provide high level, specialized services and enable transactions that might otherwise be lost. In contrast to other traits and features that have been suggested as the “essence” of professions, the need to respond to information asymmetries is a much more robust and flexible explanatory tool.

In sum, information asymmetries in the market for professional services are the key for explaining the professional reliance on trust. A normative model of professionalization describes them as trust-creating, trust-preserving institutional mechanisms that create mutual benefit by responding to information asymmetries and raising Pareto efficiency in the market for professional services. Taking this position has important philosophical implications. Drawing on this model of governance, I can explain why professionalism involves being “ethical,” as in socially beneficial and Pareto-improving, as opposed to altruistic, as many scholars of professions contend. At the same time, I can also use this framework to explain why cartelizing professional institutions are a distortion of professionalism, thereby refuting cartelization as the core feature of professionalization, as many critics of professions have argued. The mystery of professions, then, is that even though they are private economic entities, professional service
providers draw on trust-preserving institutional mechanisms in response to market failures that arise due to information asymmetries, because doing so constitutes a transaction cost minimizing governance structure.

I spent a fair bit of time in the previous chapter discussing economics. This was necessary for making the philosophical point I made here about the nature of professions. Defining professions as trust-creating institutions that promote social welfare allows me to remain agnostic regarding the motivations of professionals for performing their tasks, and also from equating professional morality with altruism. Conceiving of the normative nature of professions as involving the facilitation of trust is somewhat of a halfway point between the dichotomy of cartelization and altruism. The information theoretic characterization of professions also allows me to stay clear of any ambitions to define professions using broad, generalizing platitudes, which would feed into criticisms of definitional uncertainty.
5 Professional Morality via Roles

As I showed in earlier chapters, a core normative assumption in the debate about the professions is that professionalism is either all about private interest or all about altruism. My aim has been to use a reconstructive approach that articulates the implicit norms and ideals that structure professional practices to challenge this dichotomous characterization. Chapters one and three articulated an empirically-informed understanding of what professions do, demonstrated by the kind of work they have historically done, in order to lay the groundwork for a more sophisticated approach. My normative model, introduced in chapter four, described professions as trust-creating, trust-preserving institutional mechanisms that promote efficiency in the market for professional services. This model drew on certain descriptive assumptions about the role of professionals in society, in order to provide an account of the normative purposes already implicit in the practices of professionals.¹

By using trust to promote efficiency, professionals go beyond merely promoting their own private interests or the cartelizing interests of their professional association.² Still, it is not immediately obvious how we might arrive at the norms of professionalism through categorical imperative-style considerations about morality. In other words, professional morality involves going beyond micro-level considerations of self-interest, strictly speaking, but it also serves a narrower cooperative interest than macro-level considerations of morality. The Scylla and

² Note that my focus in this chapter (and in this dissertation in general) is on internal professional obligations. Professional moral obligations are internal, in so far as they pertain to the internal structure of principal-agent relationships between professionals and their clients (as opposed to the impact of one professional’s work on third parties outside the professional-client relationship).
Charybdis of my account may be characterized as follows: Scylla collapses my view into self-interest and criticizes my emphasis on efficiency for being reductionist. Charybdis requires a categorical imperative-style justification for professional morality, which would appeal to all-things-considered moral considerations for justification. I want to show that professional morality is a self-standing moral phenomenon; it is less than an all-things-considered morality, but still, nevertheless, a moral framework. The question is how to articulate these meso-level professional pursuits, and how to evaluate their moral legitimacy.

In this chapter, I begin by arguing that professional obligations can be moral because of the mutually beneficial efficiency that is achieved by professional institutional mechanisms. Specifically, the professional role sometimes requires incumbents to promote efficiency in the market for professional services, whereas promoting efficiency is ordinarily simply permissible in those circumstances outside the role. I then consider instances where the professional role seems to either require or permit us to do what would otherwise be forbidden. As I show, partiality, the moral division of labor and adversarialism are three increasingly controversial variations of this kind of permission. Providing a moral justification for such claims is one of the most challenging problems in the study of professional ethics. This analysis is nevertheless important, since much of professional practice occurs precisely in this moral space. I suggest that roles are a useful framework for deliberating about the moral relevance of such permissions. After discussing the literature on social roles, I propose a conceptual framework that demonstrates that institutional roles can sometimes permit, to a limited extent, what would otherwise be forbidden. In other words, I argue that, sometimes, roles can be constrained instances of deontic weakening, and sui generis sources of moral obligations. This understanding

---

3 By “all-things-considered morality” I mean impartial, broad considerations of morality, such as consequentialism or deontology.
of roles in turn informs my stance on the moral relevance of professional obligations, as less than all-things-considered moral, but nevertheless moral.

5.1 Professional Morality, Efficiency, & Ordinary Morality

As I have shown in earlier chapters, practitioners sometimes take advantage of market allocation mechanisms and professional regulations for rent-seeking, anti-competitive purposes, at the expense of efficiency. The basic structure of the market for professional services does not necessarily preclude this kind of behavior. However, I noted that the right kind of moral commitment – the kind that qualifies an individual to be called a professional – involves constraints that promote efficiency. These constraints respond to (as opposed to take advantage of) information asymmetries, and enable transactions that would not otherwise occur in the market for professional services. My first task in this section is to remind the reader of the grounds for taking such a stance.

In his 1982 article, “No Need for Morality: The Case of the Competitive Market,” David Gauthier proposed a philosophical formulation of skepticism about ethics in the market. Given certain background assumptions about perfect competition, Gauthier argued that the invisible hand theory entails that there is no need for morality in a capitalist economy. But an understanding of the nature of the firm, and of private associations (e.g. professions) more generally, shows that the invisible hand does not work as well as we might have hoped.

---

5 See, for example, Coase, “The Nature of the Firm”.
As I demonstrated in chapter three, market failures, in the form of information asymmetries, pose serious challenges to the efficiency of principal-agent relationships between professionals and their clients. This threat to efficiency is one reason why we need to talk about ethics. While a part of the solution involves adjusting incentives and other market-oriented institutional mechanisms of professional associations, these only go so far. As I showed in chapter four, we need to specify self-imposed professional moral commitments that go beyond market solutions and incentives. Professions, in comparison with other private economic entities, have traditionally excelled at this task by setting up codes of ethics and internalizing an ethos of service-orientation.

An analysis of agency risks is also crucial for understanding the moral codes that structure the activities of principals and agents in professional settings. Recall that in a principal-agent relationship, “one individual (the agent) acts on behalf of another (the principal) and is supposed to advance the principal’s goals.” Agency theory formalizes the risks posed by agents when they act according to interests that conflict with the interests of their principals. Implicit and explicit moral obligations are often put forward as solutions to agency risks in principal-agent relationships. What does “ethics” mean in this context? According to Alan Buchanan, ethics in agency theory involves “commitments” on the part of agents to reduce the risks that their behavior imposes upon principals. When professionals adopt commitments that help them

---


refrain from posing agency risks, they in effect adopt efficiency promoting deontic constraints that go beyond the basic structure of the market.

The normative force of professional moral obligations is derived from the mutually beneficial efficiency gains achieved via the trust-creating professional institutions that respond to market failures. A key criticism of my approach, however, is that it is too normatively bare for a theory of professional ethics, given its focus on efficiency. The idea of formulating deontic constraints (or obligations) based on responses to market failures and agency risks might also seem peculiar to some. Below, I draw on the work of Christopher McMahon\(^8\) in order to elaborate on the normative significance of promoting efficiency in the market for professional services. As McMahon explains, the moral commitment of individuals in the free market involves certain efficiency imperatives by design. McMahon’s discussion of the implicit morality of the market applies nicely to my analysis of professions, where efficiency imperatives guide a normative model of professionalism.

McMahon sets up a vision of moral life that to some extent resembles my discussion of Scylla and Charybdis earlier in this chapter. The monsters are place-holders for the two moral poles in McMahon’s analysis.

---

This is how McMahon sets up his discussion: he views an ordinary moral life as a “mean between two extremes,” where one extreme is made up of people with entirely “self-regarding” life projects who regularly violate moral requirements to serve their own interest. At the other extreme, there are people who have largely “altruistic ends” and respect the requirements of morality. The middle-ground is made up of people who might have a mixture of self-regarding and altruistic projects, or people who respect moral requirements, but who nonetheless have purely self-regarding projects.

---

For example, in this middle-ground stance, one can refrain from actions that would interfere with the projects of others, while promoting one’s own interests at the same time.\(^\text{10}\) This picture of the relation between morality, self-interest, and altruism is somewhat crude, but through it, McMahon explains why a self-regarding end can nevertheless be moral. As he notes,

… while merely heeding the requirements of morality does not involve sacrifices of self-interest as great as those associated with pure altruism, there is still a potential for conflict between morality and self-interest. Morality is compatible with the pursuit of self-regarding ends, but acting so as to maximize self-regarding satisfaction over one's whole life will almost certainly involve violating moral requirements.\(^\text{11}\)

Applying these ideas to his analysis of markets, McMahon holds that although self-interest is to some extent at odds with morality, it is desirable from the point of view of the market. Following Adam Smith’s invisible hand doctrine and the economic rationale of the market, individuals are in fact sometimes required to pursue self-interest. As McMahon explains, however, this does not mean individuals are to intend only their own gain in economic contexts. Instead, self-interest should be understood as a limited condition: “the normative import of the doctrine of the invisible hand is merely that economic agents are not to enter into transactions unless there is some expectation of self-regarding gain, but are otherwise to respect the full set of moral restrictions on the pursuit of self-interest.”\(^\text{12}\) Economic theory can thus be viewed as having certain “normative implications”\(^\text{13}\): beyond mere maximization of utility and profit, “the implicit morality of the market consists primarily of the hypothetical imperatives which are generated by economic theory when the achievement of economic efficiency is taken as an end.”\(^\text{14}\) An efficient allocation and distribution of resources thus requires certain conditions to be

\(^{10}\) Ibid.
\(^{11}\) Ibid., 251-2.
\(^{12}\) Ibid., 253.
\(^{13}\) Ibid., 254.
\(^{14}\) Ibid., 255.
in place, and from these conditions certain *efficiency imperatives* may be derived.\textsuperscript{15} Although McMahon calls these efficiency imperatives “hypothetical,” crucially, they need not be instrumentally motivated. In the professions, for example, efficiency imperatives promote mutual benefit by creating trust in environments where purely self-interested behavior would result in lost transactions and non-optimal outcomes.

McMahon’s analysis of the free market is applicable to professional markets, since professional associations are private economic entities. The implicit morality of the professional market would thus consist of the imperatives that are generated when the goal of economic efficiency is taken as an end. Efficiency in the market for professional services would require a set of conditions to be in place, from which we may derive efficiency imperatives. Avoiding market failures that arise from information asymmetries constitutes one such efficiency promoting imperative in the market for professional services. The efficiency promoting notion of trust has normative significance, according to this approach, because the implicit morality of markets for professional services involves achieving economic efficiency as an end.

I now map various alternatives for how the principles constitutive of professional morality may (or may not) resemble the principles of ordinary morality. I use the term “ordinary morality” to encompass broader, deliberative, all-things-considered conceptions (e.g. consequentialism and deontology), as well as our everyday intuitive understanding of right and wrong.\textsuperscript{16} The following three observations may be made by comparing professional and ordinary morality:

\textsuperscript{15} Ibid.

• Category (1) cases where the two codes require similar actions.
• Category (2) cases where the codes conflict.
• Category (3) cases where an action required by ordinary morality is neither required nor forbidden by professional morality.17

Here, we may identify the components of ordinary morality which 1) must be respected if a professional market is to function optimally, 2) those which must be suspended, and 3) those for which respect is optional from the professional point of view.18 This three-part categorization can be represented in figure 5.2 as follow:

Figure 5.2:

P = Permissible, F = Forbidden O = Obligatory

(Pa means that an action a is permissible, Fa means that an action a is forbidden, and Oa means that an action a is obligatory)

Hence, we have:

<table>
<thead>
<tr>
<th>Ordinary Morality</th>
<th>Professional Morality</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Oa</td>
<td>Oa</td>
</tr>
<tr>
<td>(2) Fa</td>
<td>Pa</td>
</tr>
<tr>
<td>(3) Oa</td>
<td>~ Oa ∧ <del>Fa (= therefore P</del>a ∧ Pa)</td>
</tr>
</tbody>
</table>

As we can see under category (1), ordinary morality and the implicit morality of the market sometimes oblige us to carry out the same actions. Recall that I formulated professional morality in terms of efficiency-promoting deontic constraints in the market for professional services.

distinguish them from “role morality,” which applies specifically in institutional social contexts discussed later in the chapter.

17 This list is modeled on McMahon’s analysis of the implicit morality of the market. See McMahon, “Morality and the Invisible Hand,” 255.

Sometimes, promoting efficiency is not only required by the implicit morality of the market, but also required by ordinary morality. For example, in the medical profession, consider the professional trappings (e.g. education and peer-evaluation) that respond to information asymmetries. These institutional mechanisms promote the implicit morality of the professional market by preventing efficiency losses due to lost transactions. At the same time, however, these professional trappings facilitate an environment of trust where patient’s autonomy and privacy are respected, and considerations of fairness and inclusiveness of care are upheld. A variation of this first category is \( P_a \rightarrow O_a \), where professional morality requires what would otherwise be merely permissible. This is an instance of deontic strengthening since the professional framework turns moral permissibility to obligatoriness. Anyone is permitted to keep an illness a secret, for example, but a patient’s physician is required to keep the secret.

The more difficult question is whether we might be able to justify professional morality under category (2), where the implicit morality of the professional market permits actions that are otherwise forbidden by ordinary morality. Similarly, category (3) involves outlining the kinds of actions that are neither obligatory, nor forbidden, or in other words, actions that are optional for professionals to carry out, even though ordinary morality requires them. Note that from the perspective of deontic logic, transformations in (2) and (3) can be shown to be equivalent. This is because:

\[
Fa = O\sim-a, \text{ and } \sim O\sim-a = Pa. \text{ Thus, we have:}
\]

\[
(2) \ O\sim-a \rightarrow \sim O\sim-a.
\]

We can also derive \( \sim Oa \) from \( \sim Oa \land \sim Fa \). This means that (3) can be represented as:

\[
(3) \ Oa \rightarrow \sim Oa.
\]

Thus, (2) and (3) are equivalent, since in both cases, the consequent negates the antecedent.
My aim in the rest of the chapter is to articulate and evaluate the professional norms that permit or require us to carry out actions that are forbidden by ordinary morality. In these circumstances, professional morality can be said to generate deontic weakening, since it turns what is forbidden into something permissible – and sometimes further, into something obligatory.

Are professionals permitted to carry out moral obligations owed to their clients, if those obligations are forbidden by ordinary morality? Take the legal profession, for example. In classic texts such as *Boswell’s Life of Johnson*, we encounter passages like the following:

BOSWELL. But what do you think of supporting a cause which you know to be bad?

JOHNSON. Sir, you do not know it to be good or bad till the Judge determines it....

As William Edmundson explains, “most laypeople (and many fellow lawyers!) look askance at criminal defense lawyers because they regard them as regularly engaged in helping those they know to be guilty evade punishment – something which would expose anyone else to censure, perhaps even to punishment, for being an accessory ‘after the fact.’” Although some might attribute a certain level of “sleaziness” to the criminal lawyer’s role, it may be argued that professional morality modifies epistemic obligations, such that the lawyer in fact does not know that her client is guilty, *even if her client confesses*. The epistemic obligations of the criminal defense attorney, the jury member, and the judge, are in fact all different, so what might readily appear as right or wrong to a jury member or judge, is not right or wrong to a lawyer. The

21 Edmundson, “Contextualist Answers to Skepticism, and What a Lawyer Cannot Know,” 2 (Italics mine).
epistemic obligations of the lawyer, the jury member, and the judge are in turn different from ordinary requirements of individuals outside these roles.

But how might we substantiate such a claim, that professional roles can modify epistemic obligations? One way to explore the relation between professional morality and ordinary morality is to explain the former as derived from the latter: the moral force of professionalism may be derived from its adherence to ordinary conceptions of morality. Another route is to show that professional conduct is somehow independently moral. According to this latter stance, professional conduct may, under limited conditions, be moral, even though it is forbidden according to ordinary morality. To elaborate on these solutions, I turn to a discussion of institutional roles. As I show, according to one definition, roles can never permit or require us to act in ways that would be forbidden by ordinary morality. Role obligations thus derive their significance solely from all-things-considered morality. Let’s call this the derivative approach. A second approach provides less stringent requirements for justifying roles, so that roles themselves can provide moral guidance – although there is disagreement over whether all-things-considered morality ever supplements this guidance. Let’s call this the generative approach. Siding with the generative approach, the rest of this chapter is devoted to arguing that, depending on the institutional environment within which they operate, roles can be distinct moral categories and constrained instances of deontic weakening.
5.2 Roles & Role Morality

Many contemporary moral philosophers have neglected role obligations, or have regarded them as, at best, marginal. Why should we look into roles as a theoretical area of inquiry? Depending on the definition of roles we adopt, roles may allow us to talk about norms and obligations without having to rely solely on guidance from all-things-considered morality. In daily professional life, our actions are governed by various formal and informal norms and institutional codes of conduct. These norms are important to articulate, especially when we have a difficult time coming to an agreement over substantive conceptions of the good as a society.

The literature on the definition of roles is wrought with disagreement. According to one approach, ordinary moral reasons are excluded from deliberations about professional roles. Let’s call this the standard approach. Instead of making judgments about the overall rightness or wrongness of actions, according to this definition of roles, incumbents restrict their reasons by considering a limited set of interests or facts. For adversarial professional service providers, for

---


example, this standard approach prescribes role-relativity and role-neutrality. “Role-relativity” exempts professionals from broad moral concerns and allows “specialized moral aims across roles.”25 “Role-neutrality” entails that professionals act “uniformly,” and are “precluded from counting the most local [i.e. personal] of deliberative concerns” when acting on behalf of others.26 The standard approach is typically adhered to in the competitive market, representative law, electoral politics, and other adversarial environments. For lawyers, for example, this view of role morality entails that: (i) attorneys should advance the legal rights of their clients (partisanship), (ii) attorneys should not consider their own views of the moral merits of their clients (neutrality), and (iii) attorneys should be shielded from moral scrutiny and criticism by observers (non-accountability).27

Here, the occupant of the role is “permitted or required to ignore or weigh less heavily what would otherwise be morally overriding considerations….28 The phrase “strong differentiation” has been used in this context to describe the relation between roles and ordinary morality.29 The role is “differentiated” in the sense that it posits role-specific obligations (or permissions) that “preempt, supersed, or outweigh what would otherwise be requirements of ordinary morality.”30 When a role is strongly differentiated, it requires its norms and “unique principles” to be weighed more heavily than they would be against other principles in other contexts.31 Another way to put this point is that roles may constitute “a sufficient reason for

25 Ibid.
26 Ibid.
30 Wendel, Lawyers and Fidelity to the Law, 550.
doing or not doing something that would otherwise be… morally wrong.” Roles can thus be a source of authority that changes the reasons we have for action. According to this approach, roles generate deontic weakening, since the requirements and permissions of ordinary morality are weakened by the context and structure of roles. In electoral politics or the competitive market, for example, a deontic constraint against, say, deception, is weakened so that limited instances of deception are permissible given specific institutional structures and safeguards.

The standard approach is an instance of the generative view about roles, since roles are taken to be a source of moral guidance.

Critics have pointed out that the standard account offers professionals a “simplified moral universe,” in the sense that role morality simply requires us to follow the rules of professional conduct along with the formal and informal norms of the professions. The worry is that this simplified moral universe is achieved at the cost of cutting roles off from the resources of ordinary morality. In this context, scholars have often been critical of what roles can make us do. Roles are worrisome since they can lead to moral alienation and “engulfment,” i.e. the loss of an individual within the role. Roles can turn into a “mask” and thereby deny or suppress an individual’s humanity. We may mistakenly believe that occupying one particular role excludes us from occupying other important roles in our lives. And we may worry about losing sight of the fact that we are responsible for the roles we decide to take up, and are also free not to adopt

35 Wendel, Lawyers and Fidelity to the Law, 552.
36 Cohen “Beliefs and Roles.”
37 E.g. Noonan, Persons and Masks of the Law, 21.
38 See Held, Rights and Goods.
certain types of roles.\textsuperscript{39} Finally, and crucially, the moral force of roles is in question, since roles may allow us to treat some people differently because of their special relationship with us. But a fundamental principle of morality is that individuals should count equally from the moral point of view.\textsuperscript{40}

The main alternative view to the standard approach all but eliminates the conflicts between professional roles and ordinary morality. We may call this the \textit{nexus} view, since it involves defining roles as a nexus of obligations and values that have moral weight.\textsuperscript{41} According to this approach, roles may be viewed as “constellations of institutionally specified rights and duties organized around an institutionally specified social function.”\textsuperscript{42} In other words, roles are a shorthand way of describing a “cluster” of obligations and permissions that apply in virtue of being in a particular kind of relationship with others.\textsuperscript{43} Crucially, according to the nexus approach, \textit{roles are not a moral category of their own. They do not involve a distinct evaluative domain.} Rather, “there is only one normative domain, that of ordinary morality.”\textsuperscript{44} Roles thus matter because of the more general moral reasons that are typically applicable to the role-holder.\textsuperscript{45} Here, the essence of particular roles is indeterminate, and the best way to specify them may be reasonably disagreed upon. For example, people have competing reasonable interpretations\textsuperscript{46} of the role of parents, and the duties that parents have towards their children (in

\textsuperscript{39} Sartre, \textit{Being and Nothingness}, 101-3.

\textsuperscript{40} See Jollimore’s and Rachel’s views on “impartial morality” discussed later in the chapter.

\textsuperscript{41} Applbaum, \textit{Ethics for Adversaries}, 45.

\textsuperscript{42} Hardimon, “Role Obligations,” 334. See also Applbaum, \textit{Ethics for Adversaries}, 48-9. Similarly, according to R. S. Downie’s definition, a role is “a cluster of rights and duties with some sort of social function.” R. S. Downie, \textit{Roles and Values: An Introduction to Social Ethics} (London: Methuen and Co., 1971), 128.


\textsuperscript{44} Wendel, \textit{Lawyers and Fidelity to the Law}, 551.


\textsuperscript{46} Dworkin, \textit{Law’s Empire}, 45-86; Hardimon, “Role Obligations,” 337.
comparison to, say, the duties to students, if these parents happen to occupy teaching roles as well).

The source and authority of the morality of role obligations are indeterminate, according to the nexus approach. Just because a role gives rise to obligations, it does not mean that those obligations are moral in nature.\textsuperscript{47} When roles do give rise to obligations, these obligations typically have “moral weight without appeal to the role as a moral category.”\textsuperscript{48} Crucially, proponents of the nexus view tend to speak derivatively about role morality by discussing ordinary morality instead.\textsuperscript{49} This view is thus representative of the \textit{derivative} approach to roles, since moral significance is derived from broader considerations of morality. Here, either roles are not considered to give rise to moral obligations at all, or if they do, these obligations are not fundamental ones. In effect, according to this stance, a role is an application of ordinary morality, grounded by one principle or another. One prominent proposal puts forward “protecting the vulnerable” as the key principle of morality that guides all social roles.\textsuperscript{50} In the case of civil law, some prominent legal scholars have proposed contracts as the core illuminating principle governing the lawyer’s role,\textsuperscript{51} and others argue that role obligations are almost universally trumped by ordinary moral obligations.\textsuperscript{52} Along these lines, one proposal indicates justice to be the core principle embodied in the legal role.\textsuperscript{53}

\begin{flushleft}
\textsuperscript{48} Applbaum, \textit{Ethics for Adversaries}, 45.
\textsuperscript{49} Andre appropriately calls this trend the “simplifying view” of role morality: “Role Morality as a Complex Instance of Ordinary Morality,” 74.
\textsuperscript{51} An array of such positions are put forward in David Luban, ed., \textit{The Good Lawyer: Lawyers' Roles and Lawyers' Ethics} (Totowa: Rowman and Allanheld, 1984).
\textsuperscript{53} Simon, \textit{The Practice of Justice}, 38.
\end{flushleft}
The significance of the role is thus at best “constrained” by background normative considerations, according to the nexus view. According to this approach, roles can require us to act in ways that would otherwise be optional (deontic strengthening). While roles can change permissions into obligations or prohibitions, the permissibility or obligatoriness of an action depends on the legitimacy of the institution within which the role exists. But roles cannot permit us or require us to do what would otherwise be forbidden (deontic weakening), according to this view. As Applbaum famously put it, roles “ordinarily cannot mint moral permissions to do what otherwise would be morally prohibited.”

The nexus view resolves the conflicts between the role and ordinary morality, in effect by stripping roles of independent moral relevance. The worry is that this moral simplification “undervalues” the ends served by roles. The goals and purposes of the social institutions within which roles operate deserve consideration. Another problem is that the nexus approach requires independent, case-by-case analysis of every deliberation and its underlying moral reasons. This is arguably an overly demanding requirement for our daily activities and interactions within social roles. The nexus view thus risks eroding the “morally valuable ends of a role” and the “long-run stability and proper functioning of institutions” at the cost of securing moral accountability of individuals within roles. The result is that the role can rarely justify a

---

54 See Applbaum, *Ethics for Adversaries.*
55 Luban, *Lawyers and Justice.*
57 Wendel, *Lawyers and Fidelity to the Law,* 552.
60 Wendel, *Lawyers and Fidelity to the Law,* 553.
departure from ordinary morality.\textsuperscript{61} This, in turn (arguably) produces too dismal a picture of much of what constitutes everyday, standard conduct by role occupants.

A third approach that merits consideration, in addition to the standard approach and the nexus approach, provides somewhat of a middle-ground perspective.\textsuperscript{62} Here, we can view roles as being designed and constituted for an end or purpose that can be reached by following the directives of the role.\textsuperscript{63} These rules and directives include professional codes of conduct, alongside formal and informal norms of the profession. It is possible to imagine scenarios, however, whereby following the constitutive rules of a role are not the best way to achieve the ends of the role.\textsuperscript{64} In such cases, the role occupant might be able to achieve the aims of the institution best through actions that are not permitted by the rules of the role. We may call this the \textbf{“recourse”} view of the role, since the role occupant is granted a permission to switch to an all-things-considered stance and recourse to the ends of the role, in order to determine whether to follow the directives of a role.\textsuperscript{65}

The recourse view creates “genuine obligations” and often requires “departures from the requirements of ordinary morality,”\textsuperscript{66} with the goal of furthering the “substantive, underlying policies embodied in the role.”\textsuperscript{67} Unlike the nexus view, a case-by-case rebalancing of underlying moral reasons is not permitted by the recourse view. Instead of a broad moral deliberation, the recourse view has built within it a permission for discretion with respect to one

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{61} Dare, \textit{The Counsel of Rogues?}, 39.
\item \textsuperscript{62} See Wendel, \textit{Lawyers and Fidelity to the Law}.
\item \textsuperscript{64} Wendel, \textit{Lawyers and Fidelity to the Law}, 551-2.
\item \textsuperscript{65} See Kadish and Kadish, \textit{Discretion to Disobey}.
\item \textsuperscript{66} Wendel, \textit{Lawyers and Fidelity to the Law}, 553.
\item \textsuperscript{67} Ibid.
\end{itemize}
\end{footnotesize}
set of considerations, namely those pertaining the specific tasks and ends the role is designed to accomplish. The conditions required for justifying discretion in this approach are thus tied to the reasons for which the role was constituted in the first place.

According to this view, roles are defined somewhat broadly, as “acting in a certain capacity, … being constrained by one's position, or … standing in a certain relationship to someone.” There is also some flexibility regarding the formation of the rules that guide the role. The formal and informal norms that surround roles might be the product of “deliberate institutional design,” or they may result from “a gradual evolution of social expectations regarding appropriate action.” The key difference between the standard and the recourse accounts is that roles are considered to be distinct from ordinary morality according to the standard account. Occupants of standard roles thus need to “exit” the role completely in order to account for ordinary moral considerations. The recourse view, however, allows incumbents to evaluate directives “within” the role, since the role grants a “permission to go back to an all-things-considered evaluative standpoint.” Thus, the recourse view is an instance of the generative approach to roles, since roles themselves are taken to provide moral guidance for incumbents. While roles are not a moral category according to the nexus view (which is broadly derivative), both the standard view and the recourse view (which are broadly generative) present roles as sui generis sources of moral obligations.

Our understanding of roles goes a long way to changing our evaluation of whether professionals are granted permissions to do what is otherwise forbidden. Consider Spaulding v.

---

68 Kadish and Kadish, *Discretion to Disobey*, 22 and 29.
69 Wendel, *Lawyers and Fidelity to the Law*, 554.
70 Kadish and Kadish, *Discretion to Disobey*, 15.
71 Wendel, *Lawyers and Fidelity to the Law*, 555; Dare, *The Counsel of Rogues?*, 30.
73 Ibid., 552.
Zimmerman as an example.\textsuperscript{74} The case involves a conflict between the moral obligation to take steps to save another's life (when it is possible and relatively easy), on the one hand, and the legal duty of confidentiality, on the other hand.\textsuperscript{75} Spaulding and Zimmerman were involved in a car accident, and Spaulding filed a post-accident negligence lawsuit, alleging that Zimmerman’s negligence was the cause of his injuries. Zimmerman’s lawyer required Spaulding to submit to an independent medical examination, whereby a doctor hired by Zimmerman examined Spaulding to verify that his injuries were not exaggerated.\textsuperscript{76} At this independent examination, the doctor diagnosed a serious injury: an aortic aneurysm (a dilation and weakening of a major blood vessel) was found, which needed immediate surgical attention. The injury was of traumatic origin, which meant it likely resulted from the car accident. Spaulding’s own physician, however, had missed the aneurysm upon treating him after the accident, so Spaulding’s lawyer was preparing the case for trial with the assumption that Spaulding’s damages were modest. Although Spaulding’s lawyer was entitled (under the relatively new civil discovery rules in Minnesota at the time) to request a copy of the medical examination report from Zimmerman’s lawyer, he inexplicably did not make such a request. Zimmerman’s lawyer could eliminate the risk of death for Spaulding by disclosing the information about the aneurysm, but in addition to violating the duty of confidentiality, this would likely drive up Zimmerman’s damages, perhaps even beyond what his liability insurance policy could cover.\textsuperscript{77} Here, it would seem that the duty of confidentiality on the part of Zimmerman’s lawyer was in conflict with the requirement to take steps to save Spaulding’s life.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{74} Supreme Court of Minnesota, \textit{Spaulding v. Zimmerman}, 116 N.W.2d 704 (Minn. 1962).
\item \textsuperscript{76} \textit{Spaulding v. Zimmerman}, 116 N.W.2d 706-7.
\item \textsuperscript{77} Ibid., 707-9.
\end{itemize}
\end{footnotesize}
The most straightforward solution in this scenario seems to be for Zimmerman’s defense lawyer to counsel his client to disclose Spaulding’s injury, in accordance with the Model Rules of Professional Conduct.\textsuperscript{78} If Zimmerman were to refuse to provide informed consent to disclosure, the lawyer could use his “fundamental disagreement” with the defendant to ask the trial judge for a permission to withdraw from representation.\textsuperscript{79} But this merely “passes the buck” so that another lawyer would have to deal with the dilemma between the duty of confidentiality and the ordinary moral requirement to save Spaulding’s life.\textsuperscript{80}

The nexus view of the role would require Zimmerman’s lawyer to disclose the critical information about Spaulding’s injury. According to this approach, the role does not permit what is forbidden by ordinary morality, and considerations of justice require Zimmerman’s attorney to take steps to save Spaulding’s life, regardless of the obligations of his role. The recourse view would allow Zimmerman’s lawyer to appeal to all-things-considered moral considerations in determining whether to uphold his duty of confidentiality. Thus, remaining aware of the requirements of the role, the attorney can deliberate about the fact that he is failing to save Spaulding’s life by refusing to disclose the injury. If the lawyer takes the ends of her role to ultimately involve serving a system of justice, then the recourse approach may justify his dissent from confidentiality rules.


\textsuperscript{79} On the duty of confidentiality and informed consent permitting disclosure, see “Model Rules of Professional Conduct,” R.1.6 (a) (b) and (c). http://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_1_6_confidentiality_of_information.html (Accessed May 16, 2013)

\textsuperscript{80} Wendel, Lawyers and Fidelity to the Law, 568.
In contrast with the recourse view, the standard view would place more emphasis on the Model Rules of Professional Conduct, and might underscore the blameworthiness of Spaulding’s doctor and Spaulding’s lawyer in fulfilling their respective professional obligations. Further, it is important to note that the rules of confidentiality are one of the most contentious aspects of the law of lawyering. As Wendel points out, “every sector of the bar and the academy, as well as the public, has weighed in on the debate.”\(^\text{81}\) Given the lack of consensus on the controversial nature of the rule, proponents of the standard view might warn against overlooking the normative force of the professional code of confidentiality (as the nexus view does), and against making discretionary judgments about the constitutive goals of the norm of confidentiality (as the recourse view does). One way to interpret the Model Rules of Professional Conduct is as an embodiment of the agreement over rules that are mutually beneficial for parties that otherwise disagree over the nature of the good.\(^\text{82}\) Proponents of the standard view might thus argue that the normative force of mutually beneficial professional rules should not be neglected, as it is by proponents of the nexus view. They also might point out that the recourse view overestimates the capacity of individual lawyers to evaluate whether the Model Rules properly represent the ends embodied in the role, given that these rules are often contentious. The worry here is that the recourse view assigns lawyers a discretionary freedom that may be somewhat arbitrary.

Providing a conclusive argument in support of the right approach to defining roles is beyond the scope of my thesis. Instead, my goal here has been to highlight some prominent frameworks for thinking about roles, and their implications for justifying certain permissions.

\(^{82}\) Wendel, Lawyers and Fidelity to the Law; Wendel, “Three Concepts of Roles”.
sought by professionals. What I will argue for is the limited claim that the proponents of the nexus view (i.e. the derivative approach) make too strong a claim when they strip the category of roles of all independent normative significance. In contrast, the recourse view and the standard view (i.e. the generative approach) can accommodate some such moral relevance for roles. In the next section, I provide a conceptual framework to demonstrate the moral relevance of the generative approach. As I show, roles are sometimes a sui generis source of moral guidance. This is not to say that roles routinely permit what would otherwise be forbidden, or that all roles are unique sources of morality. My point is more minimal. I show that given certain background conditions, roles can be constrained instances of deontic weakening – they permit what would otherwise be forbidden – pace the nexus view. In this context, I discuss partiality, the moral division of labor, and adversarialism as three increasingly controversial instances of permissions granted by professional roles but forbidden by all-things-considered morality. I start by discussing partiality, in 5.2.1 below, and then move to a discussion of the moral division of labor and adversarialism.

5.2.1 Partiality

One way to formulate and evaluate the moral implications of professional roles is through a discussion of moral partiality. As I show below, there is a longstanding debate in moral philosophy over whether the nature of moral obligations is by definition impartial. My first task in this section will be to show that partiality can be insightful for thinking about professional morality, since partial obligations sometimes purport to permit what would otherwise be forbidden according to impartial moral considerations.
An impartial choice is one in which some quality or property of the individual to whom the obligation is owed has no influence. Mainstream contemporary ethicists tend to view impartiality as a fundamental component of morality. Partial moral obligations, in contrast, require us to treat one group of people or one set of relationships differently. We appear to have partial obligations in many areas of our private and public life. Partial obligations arise in friendship, marriage, and parenting, but also in academic supervision, hospital administration, and terrorist group membership. The background environment in which partial obligations are performed might be intimate and personal, or social, political and institutional. The obligations are sometimes contractual (informal and gradational, or formal, discrete, and punctual), and other times non-contractual (tacit, by birth, or by virtue of being a social animal). Of special interest to the topic of professional morality, we have partial obligations to our professional colleagues, clients, and communities. How do we determine whether a particular partial obligation is morally justifiable? Answering this question involves engaging with a long-standing dispute over whether morality is inherently impartial or partial.

84 Jollimore, “Impartiality.”
Impartialists argue that there is no such thing as “morally admirable partiality,” and if there is, that the case is ultimately “reducible” to impartial standards. In other words, a partial obligation is moral when it is derived from some more general principle. According to this stance, morality, by definition, is impartial. As James Rachels puts it in his *Elements of Moral Philosophy*, “morality is, at the very least, the effort to guide one’s conduct by reason – that is, to do what there are the best reasons for doing – while giving equal weight to the interests of each individual affected by one’s decision.” Consider the classic example of choosing between saving one’s mother and saving an archbishop during a fire. The impartial (in this case utilitarian) recommendation notes the equal weight of the interests of the mother and the archbishop, on the grounds that doing so would best promote utility from an objective universal standpoint. Special relations with one’s mother would not be grounds for saving her over the archbishop in this case,

---


86 Jollimore, “Impartiality.”


88 See Godwin, Enquiry Concerning Political Justice and its Influence on General Virtue and Happiness. This example and variations of it have been discussed extensively in the contemporary partiality/impartiality debate.
according to the impartialist stance. Impartiality might be manifested in the form of moral rules that are impartially applied; we might use impartial benevolence as a guide to practical decision-making; or, we could draw on impartiality to explore the content of moral rules. Classic impartialist ethics thus represents a common ground among utilitarians and deontologists, in so far as both approaches identify morality with impartiality, impersonality, objectivity and universality. Moreover, the most common violation of impartiality is to privilege one’s self—and so impartialists generalize the idea that we cannot privilege our own interest to the idea that we cannot privilege anyone’s interest.

Consider the following example as an illustration of the difference between the partialist and impartialist approaches. Most people would agree that a doctor’s expert service to her patient is morally admirable. A doctor fulfills a special set of obligations toward her patient in virtue of the special doctor-patient relationship – this, therefore, is a partial obligation. Impartialists would argue that the doctor’s services do not constitute morally admirable partiality. Instead, they would point out that in so far as the doctor’s deeds are morally worthwhile, the underlying reason is that the services are reducible to impartial standards, whereby the doctor fulfilled a general duty of beneficence and benevolence by curing her patient. Thus, according to impartialists, the source of authority of the doctor’s moral obligations is not the special professional relationship, but rather the broader concerns about general morality that would have made it moral for anyone in the doctor’s place (and with the doctor’s skills) to carry out those same duties.

89 Hooker, “When is Impartiality Morally Appropriate?”
One manifestation of partiality is advocacy. For example, family doctors act as advocates for their patients by securing access to services for them, even when doing so involves bumping back other patients on waiting lists. A doctor’s daily interactions with her patients do not involve respecting the equal interests of *all* individuals in need of treatment. Instead, each doctor commits to advocating for the patients under her care. Similarly, professors write letters of recommendation for their students, advocating for them, even though doing so pits these students against other professors' students on the job market. Even though student placements of each department contribute equally to the school’s broader rankings and reputation, each professor’s partial obligations to her student give her reasons to advocate for that student, but not for other professors’ students, within the same school.

Given the dominance of impartialist moral theories, it is a challenging task to explain the moral legitimacy of partial obligations without reference to general claims about all-things-considered morality. But this is a task we must nevertheless take up. More broadly, beyond professional life, most people would agree that we have partial obligations to our parents, in virtue of our social roles, even though all-things-considered evaluations do not readily identify these obligations as morally admirable. As Christina Hoff Sommers explains, “… not a few moralists dismiss [filial obligation] as an illusion, or give it secondary derivative status.” ⁹¹ But sometimes, impartial concern falls short of explaining certain practical realities in our lives. For many, everyday moral intuitions place more emphasis on friends, parents, students, and colleagues, than all-things-considered morality would recommend. Evidence for this claim is readily available, given that most people fail to do much to assist strangers by alleviating poverty.

---

and disease around the world even when our impartial moral calculations would direct us to do so.

Further, in addition to concerns about the practicality of what impartiality requires of us, another problem is that many decisions are too controversial, and we cannot always reach a consensus over what impartial morality requires of us. Who should be included in distributive considerations (given the scarcity of resources), and what is a just political system of distribution, for example? Partiality, in such circumstances, plays a crucial role (e.g. we take it for granted that we owe distributive duties to our fellow citizens), but the moral relevance of such obligations are no always clear from the standpoint of impartiality.

In an effort to assign partiality greater weight within their impartial moral frameworks, consequentialists and deontologists have attempted to revise their positions. One such effort is the “division of moral labor.” A division of moral labor between individuals and social institutions is a strategy for accommodating diverse values. I focus specifically on Thomas Nagel’s discussion of this concept in Equality and Impartiality. There are other responses to the challenge of moral partiality, of course, but I focus on the division of moral labor since I believe it offers a promising solution.

---


94 Nagel, Equality and Partiality.
5.2.2 Division of Moral Labor

Nagel frames the problem of reconciling the partial and impartial moral stance as a question about reconciling each individual’s relation to herself, on the one hand, and her relation to society, on the other. As he puts it, ethics must be understood as “arising from a division in each individual between two standpoints, the personal and the impersonal.”\(^95\) The impersonal stance represents, in the political realm, the “claims of the collectivity” without which humans would only possess a notion of private morality.\(^96\) The personal stance, in contrast, gives rise to “individualistic motives” and requirements, which typically are an obstacle to the realization of the collective ideal.\(^97\)

In this system, social arrangements that coordinate the relations among individuals depend on a corresponding balance of forces within the self, involving the personal and the impersonal standpoints. Crucially, no social arrangement can claim the support of those living under it without appealing to both individual as well as collective interests. These forces are often in conflict within the individual. For Nagel, the impersonal stance produces “a powerful demand for universal impartiality and equality,” while the personal stance involves “individualism”\(^98\) As he puts it: “When we try to discover reasonable moral standards for the conduct of individuals and then try to integrate them with fair standards for the assessment of social and political institutions, there seems no satisfactory way of fitting the two together. They respond to opposing pressures which cause them to break apart.”\(^99\)

\(^95\) Ibid., 4.
\(^96\) Ibid.
\(^97\) Ibid., 5.
\(^98\) Ibid., 5.
\(^99\) Ibid., 5-6.
How can we deal with impartiality given that we cannot escape our individual personal stance? Nagel explains that according to the utilitarian tradition, the solution to the conflict between the personal and impersonal stance is to give dominance to the impersonal and impartial values, and to live “under the direction of an impartial benevolent spectator of the world.” But Nagel believes that the personal standpoint has to also be considered in justifying a system of ethics. Infusing general impartiality with the personal stance is a Kantian move, because it requires us to think about what impartiality would be like when viewed from each person’s point of view. This involves not just asking “what can we all agree would be best, impersonally considered?” but also “what, if anything, can we all agree that we should do, given that our motives are not merely impartial?” We thus see things from each person’s view in order to arrive at a kind of motivation everyone can share.

Integrating the personal and the impersonal has to be approached through both individual conduct and institutional design. What we need is a combination of persons changing their behavior and institutions providing the right incentives. As a psychological fact, our concern with our personal lives cannot (and perhaps should not) be minimized significantly. Although Nagel does not provide a complete solution, he describes the form of the solution for reconciling the personal and impartial standpoints through institutional design. Institutions can “penetrate and in part reconstruct their individual members, by producing differentiation within the self between public and private roles.” Nagel’s strategy for integrating the personal and the impersonal is thus to “externalize through social institutions the most impartial demands of the
impersonal standpoint.” In certain domains, drawing on institutions can help specify the moral relations among individuals. As Jon Garthof has put this point, without social institutions, the abstract nature of impartiality entails that “moral values underdetermine the obligations and entitlements of individual persons.” Of course, not all impartial moral concerns get externalized through institutions. And it may be a challenging task to find individuals to operate and support institutions impersonally, while at the same time making room for these individuals’ personality.

Even though we externalize some of the most impartial requirements of the impersonal standpoint, our support for these institutions depends on the fact that they respond to core individual demands at the same time. In this manner, “if the most serious impersonal claims can be externalized and met through occupation of a social role, the individual can pursue his remaining personal aims within that framework with a good conscience.” In sum, by dividing our moral labor, we can separate our personal moral commitments from our impartial moral considerations, some of which we may confront through our roles in social institutions. See Figure 5.3 below for an illustration.

---

106 Ibid., 5 and 53 (Italics mine).
109 Ibid.
110 Ibid., 54.
I now provide an example to illustrate how the division of moral labor works. Managers, directors, assistant deputy ministers, and deputy ministers in the Canadian civil service are permanent members of staff, regardless of which government wins elections. The highest ranked individual to whom all civil servants owe obligations is their respective minister—an elected representative assigned to the post after each election. Staff members at a given ministry naturally have political leanings of all stripes. Following figure 5.3, the division of moral labor entails that personal political beliefs are separated, within each individual, from impartial institutional commitments. Thus, staff members provide neutral professional advice and exhibit a dedicated commitment to serving the minister, regardless of their alignment with the minister’s political beliefs. As Max Weber put the point, office holding is a “vocation,” guided by specific
ethos and regulations, such that the “proper vocation of the genuine official” involves “impartial administration … *Sine ira et studio*, ‘without scorn and bias’.”\(^\text{111}\)

In addition to the duty to serve the political hierarchy, the special institutional circumstances surrounding civil servants also define a second set of obligations – Crown obligations to serve the public interest. Taking the oath to the Queen on their first day of work commences this special relationship between civil servants and the public. In this realm, the rule of law oversees the civil service, and ensures that civil servants enjoy institutional autonomy and independence from the political office.\(^\text{112}\) Civil servants draw on this institutional independence while formulating non-partisan, professional opinions, with the goal of serving the public interest. This professional opinion is distinct from the conceptions of the good that civil servants set aside in the personal realm, and also free from undue political influence from the government of the day. The division of moral labor thus allows us to formulate and theorize about the partial obligations of professionals to serve the chain of command and fulfill Crown obligations.

Now consider a civil servant who is ordered to carry out a policy that poses some small-scale harm to the public. Let’s suppose this policy involves replacing a number of parks in lower-income neighborhoods in the process of building a shopping mall – note that we are not considering nuclear war or a holocaust here. Assume that the civil service institution in question is generally just, and that the shopping mall policy poses minimal harm in comparison to alternative policy options that lead to comparable beneficial outcomes. Notice, however, that unless we assign moral relevance to the civil servant’s partial obligations (to serve the political


hierarchy and fulfill Crown obligations), the impartial stance may not permit the harm arising from reducing low-income neighborhood access to free recreational green space. Without a division of moral labor, an environmentalist on staff might put up a fight against such a policy because eliminating green space is in conflict with what that individual considers personally morally permissible. Thus, we can see that sometimes, given certain background assumptions, (i) we need to go beyond impartiality and consider the partial obligations of the civil service to the minister, and (ii) we need to draw on the division of moral labor to ensure that bureaucratic obligations stay non-partisan and professionally neutral. Impartial morality alone seems insufficient in this context.

The paradigm case that highlights the challenge of explaining professional obligations via impartiality is adversarialism. Here, the challenge of formulating the morality of partial obligations arises routinely, not just in cases like the shopping mall policy.

5.2.3 Adversarialism

Adversarial roles abound in professional settings such as law, business, and electoral politics. Individuals in adversarial roles typically take their partial obligations through the role to permit or require what would otherwise be forbidden. The question is whether our commitment to the adversarial role can give us permission to act in ways that would be wrong if not for the role. Can a role render an act (e.g. violence, deception, coercion) morally permissible, when that act would ordinarily be morally forbidden? According to Arthur Applbaum, individuals in adversarial roles sometimes claim such a moral permission, citing “the rules of

their profession” for justification.\textsuperscript{114} Applbaum offers the example of the lawyer in the adversarial legal system, who skillfully makes a case for what she knows to be unjust. Is her action morally permissible just because she has made a promise to her client to win the case? How might a professional be obliged to do something on behalf of a client, which would be wrong if done on her own behalf? Is the action permissible because it is the lawyer’s “professional responsibility” and it realizes “social values served by the division of moral labor?” Or is the professional role to blame, as opposed to the person occupying the professional position?\textsuperscript{115}

The most straightforward justification for the morality of adversarial obligations is a consequentialist one. According to the argument of “aim redescription,” the aggregate goal of adversarial institutions is to bring about good ends, so that the violations that adversaries inflict on one another are considered “unlucky accidents” or else are outweighed by other good consequences.\textsuperscript{116} Each adversary can be taken to ultimately aim at the overall social good of the institution, as opposed to the violation that particular adversarial tactics may inflict. Thus, the argument from aim redescription implies that the institution’s aim “distributes” over its practitioners. Every lawyer, then, can “appropriately be described as aiming at justice,” even if “her local efforts are directed at advocating an unjust cause.”\textsuperscript{117}

To elaborate on the argument for aim redescription, Applbaum draws on an insightful sports analogy. In the sport of fencing, for example, fencers wear protective gear, in order to protect themselves from injury, and to promote the practice of the sport without worrying about

\textsuperscript{114} Ibid., 8.
\textsuperscript{115} Ibid.
\textsuperscript{116} Ibid., 179.
\textsuperscript{117} Ibid.
injuries. As he puts it: “with the gear, they can attack without fear.”\textsuperscript{118} Becoming good at the sport, requires thrusting the sword \textit{as if} aiming to seriously injure or kill an opponent. Injuries might occur on rare occasions, despite the fact that fencers do not aim to kill each other. A well-designed adversarial institution, like the sport of fencing, can generally “anticipate and counteract the harms that could result from adversary action.”\textsuperscript{119} The fencer thrusts on the assumption that her opponent is protected, and so does the lawyer advocating for her client, or the manager pushing for profit among competitors.

In equilibrium, the good results of justice, efficient markets, and safe sport are all usually realized.\textsuperscript{120} The problem with drawing the comparison between sports and general adversarial professional settings, however, is that the good ends of adversarial institutions in equilibrium are “notoriously underdemonstrated for most of the institutions about which it is invoked.”\textsuperscript{121} General equilibrium theory in economics is an idealization, after all, with conditions of perfect market competition never being actualized. To apply this analysis to social institutions, we need to specify the necessary conditions that must be in place, and to offer an account of how robust the good results are based on those conditions.\textsuperscript{122} Similarly, according to McMahon, the morality of an efficiency-promoting free-market system is ultimately a matter of “ideal-theory,” since the background assumptions of the doctrine of the invisible hand and the perfectly competitive market almost never hold.\textsuperscript{123} The same insights apply to the realm of professions.\textsuperscript{124} Higher-level, all-things-considered consequentialist justifications for the professional role are hard to come by, in the absence of idealized conditions. The harm that is posed by coercion, deception,

\textsuperscript{118} Ibid.
\textsuperscript{119} Ibid.
\textsuperscript{120} Ibid.
\textsuperscript{121} Ibid., 180.
\textsuperscript{122} Ibid., 181. See, for example equilibrium conditions in the scholarship on democratic deliberation.
\textsuperscript{123} McMahon, “Morality and the Invisible Hand,” 256.
\textsuperscript{124} Appelbaum, \textit{Ethics for Adversaries}, 15, 257.
etc., caused by adversarial roles, also makes professional adversarial roles difficult to justify under a categorical-imperative style of analysis.

Does this mean that adversarial professional roles have no moral significance whatsoever? Not necessarily. Below, I provide a conceptual solution, building on Nagel’s division of morality framework, in an effort to draw attention to the moral significance of adversarial professional roles. As I show, at least some of the time, the professional role is a unique moral category that is not derivative from all-things-considered morality. I then provide some examples that show the moral significance of adversarial role obligations.

5.3 A Conceptual Solution

As we saw in figure 5.3, for each individual practicing within a profession, the personal standpoint, on the one hand, and the impersonal/impartial standpoint, on the other hand, coexist through a division of moral labor. This division has descriptive, historical accuracy: as I showed in chapter 1, the scientific, technical nature of professional knowledge was typically celebrated by early proponents of the professions in part because of its universal nature and its requirement of impartiality and person-neutrality. The division of moral labor involves externalizing a part of our impartial standpoint onto institutional structures such as professions. Private individual commitments, meanwhile, remain in the personal realm. We need to explain why professional norms typically involve partial obligations towards clients, and how such partial obligations are fulfilled within a broadly impartial realm of institutional morality. The answer is as follows.

Once we externalize our impartial moral concern, we may acknowledge that the way social institutions conspire to achieve outcomes involves differentiating responsibilities into distinct
*roles*. For example, in the legal profession, individuals take up different institutional roles as judges, defense attorneys, prosecutors, etc. These various institutional commitments require each individual to act partially through his or her role, since, by definition, job descriptions require professionals to treat one group of people or one set of relationships differently. The partial obligations of the defense attorney and those of the prosecutor naturally differ. Still, for every attorney, the personal moral concern remains outside the institution, and individuals are not personally responsible for their institutional roles. These partial institutional obligations are not all-things-considered morality, since they represent the obligations attached to specific social roles. Nevertheless, even though the partial moral concerns are less than the cumulative considerations of morality, they remain a subset of the externalized and institutionalized impartial moral concern. See figure 5.4 below:
There is an interesting conceptual gain in dividing up partial and impartial obligations in this manner. If we think of ourselves as legislators in the kingdom of ends, we would have a difficult time arriving at a complex division of individual roles within social institutions by asking ourselves “what if everyone followed that rule?” In other words, categorical imperative-style thinking would lead to a set of rules guiding the actions of, say, a defense attorney, but not to two different sets of rules guiding the defense attorney and the prosecutor. Offloading some of our categorical imperative-style considerations onto institutions allows institutions to take care of certain “what if everyone did that?” moral concerns. This externalization process in turn allows for the creation of complex divisions within our social obligations, according to some of which we might have partial obligations to fulfill.

As I showed earlier in this chapter, many scholars deny that there are any pre-packaged obligations attached to roles. Among those who recognize role obligations, there is considerable disagreement over the moral nature, and the source of authority or legitimacy, of these obligations. Figure 5.4 demonstrates roles as a moral category of their own, within the externalized realm of institutional morality. As proponents of the standard account of roles might point out, roles are the manifestation of the norms that guide social institutions. Alternatively, as proponents of the recourse view might argue, we recourse to all-things-considered morality with respect to the ends for which the role is constituted, in order to evaluate the moral force of roles—all the while remaining in role. According to both these accounts, roles can be a sui-generis source of morality. The specific institutional circumstances within which roles take place play a crucial role in determining when this happens, and in providing justification for the circumscribed deontic weakening posed by roles.
Through this proposal, social roles can be drawn upon to explain instances of moral relevance that are less than all-things-considered morality, but nevertheless have moral significance. Partiality, moral division of labor, and adversarialism are three incrementally controversial instances of permissions that have this kind of meso-level moral relevance. I have already given examples that drew attention to the moral relevance of professional roles: recall the discussion of advocacy (partiality), and the shopping mall policy example (division of moral labor). I now apply my findings about roles to a series of more challenging cases. As I will show, roles can modify the epistemic obligations of incumbents. Specifically, roles may raise epistemic standards to a higher level than those applicable to ordinary persons outside the role; they can also lower standards of epistemic obligation and relieve incumbents of requirements that apply outside the role. The weakening of epistemic standards is not of any intrinsic importance, of course. Rather, it is an example of the kind of deontic weakening that roles may permit.

A prominent contextualized epistemic standard in the Anglo-American system of law is proof beyond a reasonable doubt, and jurors are required to apply this standard to evidence when determining guilt. Often the evidence at trial is sufficient to convey knowledge of the defendant’s guilt to a layperson (if one happened to be present in the room). But jurors might nonetheless vote to acquit. How might this be justifiable? Why might the epistemic standards applicable to the lawyer, the judge, and a layperson differ? As Edmundson explains, roles import increasingly exacting epistemic standards upon jurors and lawyers, higher than those applicable to a layperson. This is why jurors might vote to acquit someone that the public views as clearly guilty. In comparison to the juror, the epistemic standards attached to the lawyer’s role are even more exacting – the lawyer is often exposed to evidence that would prove guilt beyond

---

a reasonable doubt, if she were occupying a juror’s role. Edmundson provides a contextualist argument in favor of such role-relative epistemic standards: “contexts alter, and with them, epistemic standards alter. … Among the things that can alter the context is the role of the knower and the relationships the knower has to others as determined by that role. Roles alter contexts and contexts alter the epistemic standard.”

Consider a similar example. In comparison to a civilian who might confirm the presence of gravity by observing a bag of stones being dropped on the ground, a scientist studying gravity who drops that same bag of stones is required to meet higher epistemic standards before confirming the observation of gravity. This is because the scientist’s role provides her with more stringent epistemic standards in comparison to non-experts whose expertise and daily interactions do not involve the study of gravity. While the scientist’s role involves finding enough evidence for her own experiments and research goals, the attorney’s role is typically to provide evidence for others (namely prosecutors, jury members, and the judge) so that they may meet the epistemic standards that are applicable to them. This is why the attorney may continue her efforts to prove that reasonable doubt exists, even when she herself has enough evidence to rule out reasonable doubt.

It might be objected that the attorney is taking part in deception, since she is persuading others to believe what she herself does not believe. Arguably, however, at least in certain instances, this harm may be minimal. Attorneys focus on proposing one essential proposition to their audience: “the evidence before you, the jury, does not exclude every reasonable doubt of

---

126 Ibid., 7.
127 Ibid., 4.
128 Ibid., 10.
129 Ibid.
130 Ibid., 18.
This proposition does not refer to the whole body of evidence, but rather to a subset of evidence that is properly before the jury, subject to limiting instructions given by the judge. Thus, a limited deontic weakening may be justifiable by the role. Ultimately, for all the lawyer knows, by causing the jury to have serious doubts about the evidence at hand, she might have helped an innocent person to be acquitted. The jury gets to decide factual guilt, not the lawyer – this is why it is sometimes justifiable to be an adversarial, zealous advocate. As Edmundson puts it “it would be absurdly harsh to say that those who have been properly convicted, having been found guilty beyond a reasonable doubt, deserve punishment whether or not they are factually guilty.”

The lawyer’s role modifies the epistemic obligations of the attorney in other ways as well. For example, the lawyer’s role can be taken to permit him to focus on providing evidence beyond a doubt in support of, but not against his client. In ordinary scenarios outside the role, the relevant epistemic standards would be more broad and comprehensive. For example, during a routine disagreement between his children at the dinner table, he would inquire about the facts supporting each of the opposing perspectives. To make a fair judgment about the dispute, he is expected to consider both sides of the story. Outside his role, the lawyer would typically look at evidence on both sides of a dispute. At work, however, his role permits him to lower his epistemic standards when it comes to evidence supporting her opponent’s client. Lawyers are sometimes accused at cocktail parties of being liars. This is a typical instance of the failure of ordinary moral consideration to recognize the normatively salient features of adversarial professional roles.

---

131 Ibid.
132 Ibid., 20.
Let’s consider a different example now, this time in a division of moral labor scenario. Lower-level civil servants (with limited access to discretionary decision-making capacities) typically have access to fragmented knowledge, because the structure of bureaucratic hierarchies limits clearance for information. Luban et. al argue that moral responsibility in such bureaucratic settings involves breaking down hierarchical barriers to information and adopting “preemptive” duties to gain knowledge. Specifically, they call for “obligations of investigation, communication, protection, and precaution,” as a framework of bureaucratic moral responsibility. As the authors put it, the kind of “virtue” that civil servants should aspire towards involves appealing to higher epistemic standards, similar to those applicable to “scientists” and “trial lawyers.” This is in effect a derivative moral evaluation of the bureaucratic role, where knowledgeable decision-making is proposed as the appropriate higher-level guiding principle. There are two problems with this proposal.

(i) Institutional roles are sometimes too complicated to accommodate the requirements of such knowledgeable decision-making. We may call this the argument from complexity. In the case of the civil servant, whenever possible, accessing information that allows for more knowledgeable moral evaluation is of course desirable. The problem is that demanding such a requirement may not be practically feasible given the complex bureaucratic structure and the institutional settings of the civil servant’s role. Incumbents may not be able to gain access to knowledge and fulfill higher-level epistemic obligations, despite making preemptive attempts to do so. A low-level front-line civil servant receives orders from her manager, who in turn receives orders from multiple layers of hierarchy, each of whom has access to information that is above

133 Luban, Strudler and Wasserman “Moral Responsibility in the Age of Bureaucracy,” 2385.
134 Ibid.
135 Ibid., 2391-2.
the clearance threshold of lower ranks. Multiply many of these levels of hierarchy, and it becomes that much more difficult for a lower-level agent to gain enough information to routinely evaluate directives.

(ii) Beyond the practical infeasibility of gaining access to information due to complex institutional structures, at least some of the time, front-line role occupants have access to fragmented knowledge *for good reason*. The formal and informal norms guiding different institutional roles can thus render derivative attempts for moral evaluation inappropriate or counterproductive. We may call this the argument from **local norms**. For example, consider lower-level admin staff and factory workers in a competitive market institution such as a tech firm, where market knowledge is a competitive advantage. Secrecy is crucial to almost all aspects of work in this context. Apple has allegedly shipped products in unmarked Tomato boxes at least once, in order to block industry spying on their supply chain. Electoral politics provides another such example. Lower ranked bureaucratic and political staff members have access to limited knowledge about policies in forthcoming political platforms, but arguably this is for good reason. The success of the political platform depends to a large extent on the response from opposition parties and the media. The slightest lack of caution in social media or a simple careless dinner conversation prior to the official release date can jeopardize months of planning and policy making. Notice that this is not an all-out argument against access to information or transparency. Instead, my aim is to underscore that we should study roles more closely and attend to their independent moral guidance, since in certain limited circumstances, the local norms guiding roles successfully generate deontic weakening.

Consider one more example, in the realm of non-political officials and administrators in the Canadian civil service. These officials possess sufficient authority to practice discretion –
potentially in an arbitrary manner – in the process of fulfilling their duties to the public. Managers, directors, assistant deputy ministers, and deputy ministers meet this criterion formally. In contrast to lower level, front-line bureaucrats, these civil servants not only have access to information, but they often have access to *more specialized knowledge* than their superiors. Unlike litigators whose epistemic obligations require them to provide evidence for others, these civil servants’ epistemic obligations are internal to the profession, aimed at producing top of the line, independent, social scientific pronouncement of the impact of policy on the public interest. The norms of confidentiality and bureaucratic independence are two informal standards which ensure that the knowledge of civil servants is not compromised either by political meddling or by the media and opposition parties prior to planned and timed release.

The structure of institutions and the partial obligations arising from roles alter the epistemic obligations applicable to incumbents. As we have seen, these epistemic obligations are not uniform within or across institutions, and crucially, they may differ from the epistemic obligations that are taken to apply to individuals outside the role. In division of moral labor (e.g. civil service) and adversarial scenarios (e.g. litigation), I have shown that roles can be a limited source of deontic weakening since they modify epistemic standards applicable to partial obligations (e.g. to serve the public interest & providing adversarial criminal defense).

Applbaum argues that most adversarial role obligations are not moral, since they cannot mint moral permission to do what would be wrong if not for the role. McMahon similarly holds that a general relaxation of the standards of general morality is unjustified in the

---

136 For an in-depth discussion of the definition of the civil service and the various uses of the term in different regions in Canada, see Sossin “Speaking Truth To Power? The Search for Bureaucratic Independence in Canada.”

137 Applbaum, *Ethics for Adversaries*. 
competitive market (at least within advanced developed societies). Accordingly, a large section – perhaps even a majority – of our contemporary social roles claim permissions that are not morally legitimate. But as I have argued, the way social institutions conspire to achieve outcomes involves differentiating responsibilities into distinct roles. These roles can constitute circumscribed cases of deontic weakening, and merit attention because of their importance in shaping our daily moral experience. Not all professional roles are moral, and not all roles successfully generate deontic weakening, but the derivative view makes too strong a claim by stripping roles of all independent normative significance. In contrast, the generative approach (recourse or standard) allows us to explain why roles can be sui generis moral categories. Whether through distinct institutional settings or through recourse to the ends which roles are designed to achieve, it is possible for some roles to have constrained, yet independent, moral weight, at least some of the time. This is why professional morality, as a realm of institutional morality constituting a set of social roles, can be said to have a meso-level moral significance – not quite all-things-considered moral, but still moral.

In the next chapter, I apply my findings thus far in the thesis to the field of management. After arguing that managers are professionals, I explore some of the implications of the discussions about professional role morality for determining managerial moral obligations.

---

6 Management Professionalism

Professions and professional associations are an important (though underappreciated) topic in business ethics. Even though they are private economic entities, professions have historically been portrayed by some as ethically motivated, and there is a long-standing literature that documents professional moral norms and codes. Since early in the twentieth century, scholars have argued over whether management is a profession, and whether we may theorize about the managerial\(^1\) role using the normative and theoretical trappings of the professions. In this chapter, I argue that managers are professionals, properly understood, since they draw on trust-creating, trust-preserving institutional mechanisms to respond to market failures caused by information asymmetries in corporate principal-agent relationships.

Viewing managers as professionals has attractive implications for a reconstructive normative model of managerial responsibility. Promoting intra-firm efficiency involves adopting imperatives that are generated when the goal of reducing efficiency losses is taken as an end by managers. Professional managers are thus required to promote efficiency, so long as doing so is permitted outside the role. Can the professional role ever legitimately permit managers to act in ways that would otherwise be forbidden? As I show, partiality, division of moral labor, and adversarialism are three controversial instances where such permissions may be granted by the professional managerial role. Thus, the professional managerial role, properly understood and occupied, can be a sui-generis source of moral guidance.

\(^1\) By managers I have in mind a group of salaried employees, distinct from suppliers of capital and labor, who work within the hierarchical administrative structure of private organizations (bureaucratic, entrepreneurial, or other). This definition is admittedly somewhat vague, but it allows me to discuss both formal and informal managerial roles within the firm. My focus here is on upper-level managers. See Alfred D. Chandler, Jr., \textit{The Visible Hand: The Managerial Revolution in American Business}, (Cambridge: Belknap Press, 1977).
6.1 Is Management a Profession? A Literature Review

Management has traditionally not been considered a profession. For early generations of merchants in colonial United States, for example, commerce did not qualify as a profession due to the “mediocre social credentials” of merchants.\(^2\) Although merchants slowly began to gain social recognition,\(^3\) still, commerce was not thought fit to acquire the title “profession” through to the end of the 19\(^{th}\) century, since “the professional man was expected to avoid the pursuit of wealth.”\(^4\) The question of whether the language of professionalization is appropriate for analyzing business and the role of managers in society remains a persistent problem for scholars of the professions and of management. Some scholars have argued that business should become professionalized. As Emile Durkheim asked, in *Professional Ethics and Civic Morals*, “[t]here are professional ethics for the priest, the lawyer, the magistrate...Why should there not be one for trade and industry?”\(^5\) Others, such as Louis Brandeis, argued that this transformation was already underway by early in the 20\(^{th}\) century, suggesting that business “to some extent already is, one of the professions.”\(^6\) Parsons, who produced some of the most important early writings on the


\(^6\) Brandeis, *Business – A Profession.*
professions, originally noted in 1937 that professions and business were distinct due to their respective motivations of altruistic service and acquisitiveness. But by 1939, he shifted his position, and argued that business was functionally and rationally organized, and was thus indistinguishable from the professions. Still, business was regarded, at best, as an “emerging or marginal profession” as late as the 1960s.

Management scholars first discussed this issue widely early in the 20th century when university-based business schools were being established for the first time, but the discussion continued throughout the 20th century, with dynamic exchanges appearing in the *Harvard Business Review*. In the very first issue of the journal in 1922, for example, John Callan noted that "[b]usiness… may be thought of as a profession [and] we may profitably spend a good deal of time in considering what is the best professional training for [those] who are to take important executive positions in the coming generation.” Similarly, Lippmann argued that business

---

7 Parsons, “Remarks on Education and the Professions.”
8 Parsons, “The Professions and Social Structure.”
should become “a profession with university standing equal to that of law, medicine, or engineering.”\(^\text{14}\) There has been a revival of this literature over the past decade, as a response to the growing concerns about corporate scandals and the role of business in society.\(^\text{15}\)

The question at hand is whether the manager’s role, situated within the corporate organizational structure, can accurately be described as a professional role. Consulting management scholars’ contributions to this question over the past century turns up an inconclusive answer at best. I argue that the confusions regarding this question arise from two broad misunderstandings in the literature. First, scholars on both sides of the debate often identify particular professional characteristics and traits to argue for (and against) the professionalization of management. But the same definitional uncertainty that surrounds the literature on professions carries over to discussions of professional management. To this day, widespread disagreement continues over the trait(s) that constitutes a necessary and sufficient condition for the existence of a profession of management.

Second, and more problematically, there is a curious lack of appreciation among management scholars for the structural critiques that have been made against professionalization.

---


Supporters of professional management tend to dichotomize individualistic profit-orientation and professional service-orientation. But this approach is suspect, until it is substantiated by an account of professions that can block criticisms that accuse professional practice of being necessarily self-serving and cartelizing.

6.1.1 Professional Traits

Professionalization is embraced by some scholars as an institutional and normative shift that is bound to make managers more socially responsible. Others reject the claim that professionalization is suitable for managers, and question whether professional roles are necessarily more socially responsive than other roles. There is an assumption here, on both sides of the argument, that there are particular traits and features that can be drawn upon in defining a profession. But this assumption betrays a lack of familiarity with the definitional uncertainty and the unresolved disagreements regarding the fundamental features of professions that played out over the course of the 20th century. As I mentioned in my first chapter, the literature on professions is fraught with definitional uncertainty, despite a century long quest by theorists to establish specific traits as the defining features of professions.

In the most extensive recent study on this subject, Khurana provides a socio-historical analysis to demonstrate that management scholars legitimized the field of management in the late 19th century by drawing on elements of professionalization and science, and through the establishment of the university-based business school. Khurana argues for a return to this era of “higher aims” for professional management. In order to promote professionalization in

---

16 Khurana, *From Higher Aims to Hired Hands*. 
management, he draws on a certain understanding of the definition of professions. For example, he identifies “expertise, autonomy, and an ethos of service to society” as “distinctive features” of occupations that qualify as professions.\(^{17}\) Elsewhere he notes that “the ideal of professionalism had always rested on combining mastery of specific knowledge with adherence to certain formal or informal codes of conduct and, even more fundamental, to an ideal of service.”\(^{18}\)

In another statement, Khurana declares that the fundamental traits of professions include a common body of theoretical knowledge, a system of certification and licensure, a code of ethics, along with “a commitment to promote the public good….”\(^{19}\) Thus, Khurana highlights a variety of traits as definitionally imperative, without spending any time explaining his reasons for choosing one set of traits over another. But service orientation, specialized knowledge, the promotion of the public interest, and formal professional associations are distinct elements, with sometimes conflicting implications, in the definition of professions. Scholars of professions have so far failed to put forward a winning proposal for a definition.

Interestingly, on the other side of the argument, scholars draw upon a similar hodgepodge of contrasting traits to argue against the case for characterizing business as a profession. For example, Pfeffer argues that management is not, and cannot be, a profession because it does not meet the necessary standards for a “specialized body of knowledge,” which involves the use of the ultimate professional traits of “scientific data” and “data mining.”\(^{20}\) Barker provides a sophisticated analysis of professions based on permanent knowledge asymmetries, but he insists that two particular traits—clear boundaries for a unique and specialized body of knowledge, and a

\(17\) Khurana quotes Freidson’s *Professionalism: The Third Logic* (18-32), but these definitional features are in fact contested, and disagreements continue regarding the fundamental traits that define professions. See Evetts, “The Sociology of Professional Groups,” 133.
\(18\) Khurana, *From Higher Aims to Hired Hands*, 291.
\(20\) Pfeffer, “Management a Profession? Where’s the Proof?”.
consensus over those boundaries— are necessary features of professions.\textsuperscript{21} Based on these criteria, he concludes that managers cannot be professionals. Mintzberg has also argued that managers are not professionals, on the grounds that management is in fact a “soft practice,” and a combination of art, craft, and science.\textsuperscript{22}

Thus, in arguments for and against viewing business as a profession, management scholars have drawn on a variety of definitional traits, without due attention to the long-standing disagreements among scholars of professions over the fundamental attributes of professionalization.\textsuperscript{23} It is evident that the definitional uncertainty surrounding the concept of professionalization has caused confusion among management scholars eager to apply this institutional structure to business.

\section*{6.1.2 Lack of Critical Attention}

In addition to confusions in the literature about the quintessential traits of professions and about their role in resolving the definitional uncertainty surrounding professions, both sides of the argument appear to rely on a problematic normative stance with respect to the nature of professions. Whether they are for or against viewing managers as professionals, the characterization of professions by management scholars is to a large extent positive, although conclusive justification for this positive portrayal is often lacking. Management scholars have almost uniformly overlooked the problematic features of professionalization - such as

\begin{thebibliography}{99}
\bibitem{21} Barker, “The Big Idea: No, Management Is Not a Profession.”
\bibitem{22} Henry Mintzberg, \textit{Managers Not MBAs: A Hard Look at the Soft Practice of Management} (San Francisco: Berrett Koehler, 2004).
\bibitem{23} I have discussed these problems in detail in the first chapter of this dissertation.
\end{thebibliography}
monopolistic, self-serving attempts to create social and economic closure in society.\textsuperscript{24} There is thus a failure to attend to the potentially problematic legitimizing role of professionalization, and its various implications for the social role of professionals. It is not unreasonable to worry that this legitimacy cloak may in fact be what some business schools are quietly hoping to acquire by drawing on the language of professionalization.\textsuperscript{25}

The tendency to ignore the criticisms of professionalization, and to draw on exclusively laudable traits to define professions, has a long history in writings about business. The first scholars to do so in fact argued that the commendable features of professions disqualify business from ever becoming a profession. Abraham Flexner, a pioneer scholar of professions, argued vehemently against the case for seeing business as a profession in 1930. As he put it, “modern business does not satisfy the criteria of a profession; it is shrewd, energetic, and clever, rather than intellectual in character; it aims … at its own advantage, rather than noble purpose within itself.”\textsuperscript{26} He thus declared that Harvard Business School, which claimed to draw on a model of professionalization to define business education at the time, failed in its attempts, because its curriculum “raises neither ethical nor social questions; it does not put business on the defensive; it does not even take a broad view of business as business.”\textsuperscript{27} Thorstein Veblen, another early opponent of the expansion of business schools on the pretext of professionalization,


\textsuperscript{25} As Pfeffer and Fong have shown in “The End of Business School 'Business',' the MBA has been the fastest-growing graduate degree in American universities for the last twenty years: “There is little doubt that business education is big business and for many, including business schools and their professors, a lucrative business at that.”


\textsuperscript{27} Flexner, \textit{Universities}, 166-7.
argued that “no gain comes to the community at large from increasing the business proficiency of any number of its young men.”

This tendency to glorify professional altruism was shared also by proponents of professionalizing managers. According to Joel Podolny, a former dean of the Yale School of Management, professional managers, unlike regular business people, do not care about profits, which is a sign of their concern for the public interest. As he puts it: "[a]n occupation earns the right to be a profession only when some ideals, such as being an impartial counsel, doing no harm, or serving the greater good, are infused into the conduct of people in that occupation.”

This view is also evident in Walter Lippmann’s 1914 writings about the new generation of salaried managers in American corporations. These administrators have “no tradition to work with, and the old commercial morality of the exploiter and profiteer still surrounds these new rulers of industry.” But as Lippman puts it, this old “commercial morality” is now being put behind. He explains:

The real news about business, it seems to me, is that it is being administered by men who are not profiteers. The managers are on salary, divorced from ownership and from bargaining. They represent the revolution in business incentives at its very heart. For they conduct gigantic enterprises and they stand outside the higgling of the market, outside the shrewdness and strategy of competition. The motive of profit is not their personal motive. That is an astounding change. The administration of the great industries is passing into the hands of men who cannot halt before each transaction and ask themselves: what is my duty as the Economic Man looking for immediate gain? They have to live on their salaries, and hope for promotion, but their day’s work is not measured in profit. There are thousands of these men, each with responsibilities vaster than the patriarchs of industry they have supplanted.

---

29 Podolny, “Are Business Schools to Blame?”.
31 Ibid.
Surveying the contributions of management scholars to the literature on business as a profession, Khurana emerges as the sole theorist who seems familiar with the critical literature regarding professions. But Khurana’s approach has certain shortcomings.

According to Khurana, professionals live up to a “standard of ‘unselfish, reasonable, ascetic, scientific, and impersonal judgment’;” they “place the interests of those whom they advised above their own;” they “actively avoid situations in which their disinterestedness could be questioned;” they also “refrain from activities that yielded them personal advantage.” This is arguably why managers went from being an obscure entity in early twentieth century American corporations to acquiring an image as a “trustworthy steward of the economic resources represented by the large, publicly held corporation.” This transition is said to have occurred only “after a sustained quest for social and moral legitimacy—finally achieved through the linkage of management … to the common good.”

Khurana then implies that the social value of the professional’s work somehow justifies questionable structural features of professional work. This point is not really developed, although in a footnote he notes that professions are a “social contract” that give professionals certain privileges in exchange for their contribution to “social order” — a position espoused in the works of Hughes, Parsons, Merton, and (the later work of) Freidson. In other words, Khurana believes that the negative aspects of professionalization somehow get cancelled out by its

---

32 Khurana, From Higher Aims to Hired Hands, 69. Khurana quotes Freidson, Larson, and Parsons in this context, but a careful treatment of these authors’ contributions would have led to a more critical discussion of professions. In other words, in order to portray the ends of professions as synonymous with “higher aims,” Khurana needs to do further conceptual work to respond to the extensive critical literature surrounding professions.
33 Khurana, From Higher Aims to Hired Hands, 3.
34 Ibid.
positive social contributions. As he puts it, professionals are committed to a renunciation of the goal of profit-maximization in return for professional autonomy and monopoly power.36

But although this line of argument is promising, it needs further elaboration. Specifically, Khurana’s treatment of the definition of professions does not adequately deal with the definitional uncertainty in the literature. The assumption that professionals necessarily strive to bring about the public interest in an unselfish manner is also unwarranted. As I showed in my discussion of partial morality in the previous chapter, professionals sometimes claim permission to fulfill partial obligations that are forbidden by general morality. Other times, professionals take their role to give them permission to treat the demands of general morality as optional. We thus need a more careful analysis of the normative nature of professional work. An appropriately sophisticated analysis of professionalization stays clear of glorified altruistic characterizations, while remaining sensitive to the socially beneficial nature of professional work.

My definition of professions and my approach to the normative role of professionals in society thus provides a core missing step in the kind of argument that management scholars like Khurana have been trying to make. Based on my normative model of professionalization, I can outline the features of management professionalism that characterize managers as concerned with more than merely self-serving – yet less than altruistic – ends. I turn to this task below.

36 Khurana, Nohria, and Penrice, “Management as a Profession,” 5. By “profit-maximization,” the authors mean personal profit – as opposed to corporate profit. As they explain, “professionals are specifically enjoined from using the laws of the market to reap economic gain at the expense of their professional obligations.”(5)
6.2 Trust and Management Professionalism

In this section, I argue that managers are professionals, properly understood, since they draw on trust-creating institutional mechanisms to promote efficiency in response to information asymmetries. To lay the grounds for my discussion of management professionalism, I begin by reminding the reader about the features of my model that distinguish professions from other private economic entities.

As I showed in chapter four, a deficit of trust is a typical structural feature of private economic markets. Well-designed market mechanisms create social welfare despite the absence of the relationships of trust that govern personal interactions between friends and family members. Recall the example of the used car market: buyers assume that sellers will try to get the highest price possible regardless of the actual value of the car. In the absence of mandatory disclosure regulations, sellers rarely volunteer information that might reveal the true quality of the car. This structural feature is not typically present, however, in interactions with professionals such as lawyers and engineers, even though these services are also offered in a private economic market. So why might we have an expectation that professionals like lawyers or engineers can be trusted, unlike used car salespeople?

My answer to this “mystery of the professions” was that the presence of information asymmetries in the market for professional services explains the need for professionals to rely on mechanisms that create trust. Here, trust is the expectation of “regular, honest, and cooperative behavior, based on commonly shared norms,” on the part of members of a given community. Professional mechanisms are typically governed through market solutions (e.g. competition and

---

37 Fukuyama, Trust: The Social Virtues and the Creation of Prosperity, 26.
reputation-building) as well as non-market solutions (e.g. trust created through codes of ethics). Professions, according to my approach, create and preserve trust through institutional mechanisms that respond to information asymmetries and raise Pareto efficiency in the market for professional services.

But why should we rely on trust? Why not just focus on improved contract design and more effective rewards and punishments? After all, the dominant economic framework in the business and economic literature focuses on “external” incentives, such as performance pay and promotional systems, as strategies for responding to information asymmetries. In effect, this is a system of incentives where agents are motivated through punishments and rewards to carry out the expectations of their principals, within principal-agent relationships. As some agency theorists have pointed out, however, the over-reliance on external incentives has had a negative impact on the explanatory power of agency theory and on the management solutions that rely on it. Consider, for example, the professional identity of agents and the professional ethos that may prompt them to work diligently and loyally on behalf of a principal. These agents may happily work hard without any real promise of reward or threat of punishment. Such a motivation serves as an “internal” incentive for the agents to obey principals.

Internal incentives, in contrast to external incentives, may involve the monitoring of agents’ preference changes and ensuring that employees are inspired by their jobs. External

---

40 Prominent theorists often ignore the role of internal incentives. For some of the negative consequences of this mistake, see Milgrom and Roberts, Economics, Organization, and Management. For further discussion see Boatright, Ethics in Finance, 48-49, and Dees, “Principals, Agents and Ethics,” 35. As Heath explains in “Uses and Abuses of Agency Theory,” this negligence may in part be explained through modeling, since economists have typically had more success modeling external incentives in comparison to internal incentives.
incentives are often formal and written; in comparison, internal incentives are informal and unwritten, yet tacit and operational. While informal incentives do not appear in job descriptions and codes of conduct, they play a pivotal role, alongside formal external incentives, in reducing agency risks within principal-agent relationships. This, then, is one reason to also focus on trust as opposed to just performance pay. Trust typically represents internal incentive structures: it can involve loyalty, mutual benefit, shared identity, and a culture of promise-keeping, and it supplements external incentives in motivating agents to fulfill obligations in principal-agent relationships.

In addition to incentive structures in agency theory, trust is relevant to my project because of its impact on professional governance structures. We can choose between different governance mechanisms by comparing their transaction costs. The governance system with the lowest transaction costs is the superior choice, since it allows resources to be reallocated for more efficient purposes. In professions such as law or medicine, for example, professional ethics codes are a governance mechanism that creates trust and guides professional behavior through enforcing an ethos of service among service providers. Of course, I don’t claim that trust is the only relevant governance mechanism. Ordinary market mechanisms that impact professional conduct through competition and market reputation, and legislative interventions that impacts


professional conduct through regulations are also relevant. Often, we might draw on a combination of governance mechanisms.

With these preliminary considerations in mind, I now argue that based on my normative model of professionalization, managers can be viewed as professionals. I begin by showing that the corporate institutional structure allows managers to make efficiency-improvements within the firm by responding to information asymmetries. I then propose that the separation of ownership and control in corporations is precisely the kind of transaction that might have been lost, if not for the environment of trust surrounding managers.

6.2.1 Managers as Professionals

Information asymmetries are pervasive in corporations, and fundamentally impact the relationship between managers and members of the firm’s chain of command. Specifically, moral hazard makes it difficult for managers’ work to be adequately evaluated and for their effort levels to be accurately measured.

Inside the firm, the division of labor along the chain of command entails that information is often localized. A technician, for example, may know how to solve a complex problem quickly, but she may not share this knowledge with her supervisor, who is not in a position to evaluate how much time should be allotted for such a task. Similarly, managers have a wealth of information about how corporations actually work (no matter what the formal manuals say), and what corporations actually accomplish (regardless of how things are presented on financial reports). Shareholders, owners, board members, and employees are not always privy to this information, because their organizational ranking assigns them varied informational access.
While these members might have access to unique information that is in turn important for the firm’s general success, they do not have access to information about the procedural details that managers oversee. Managers control how much information is shared, how it is shared, and when it is shared. The complex nature and structure of corporate management is often not observable by individuals outside the firm’s organizational hierarchy either. This is at least in part a structural feature of corporations by design, which helps them gain an advantage in markets where information about products, strategy, and governance translates to gains in competitiveness.

This feature of the firm – information asymmetries surrounding the work of managers– can create moral hazard. Here, managerial work quality and effort levels cannot be readily evaluated. Adverse selection, in comparison to moral hazard, typically occurs prior to when an employment contract is signed, i.e. prior to the manager being hired. The manager can try to communicate information about her quality (her “type”) to potential employers (a signaling game). At this stage, my focus will be on moral hazard problems between managers and various individuals along the hierarchy of the firm, since those are some of the most prominent and problematic sources of efficiency loss. Below, I discuss two sets of relationships: those between managers and superiors in the chain of command, and those between managers and employees.

**Managers vs. superiors**

For the purpose of my discussion, by “superior,” I have in mind any individual or group in the organizational hierarchy to whom the manager might report. Members of the hierarchical corporate structure ranked above the manager may include the board of directors, executives, and company owners. Consider figure 3.2 (a) – moral hazard (hidden action).
In this standard information theoretic model, the manager (agent) and her superior (principal) have symmetric information when the manager is hired. But Nature makes a move to create a state of the world after the game begins. This move is observed by the agent, but goes unobserved by the principal. Since managers may fail to act in the best interests of their superiors, incentives and monitoring solutions may result in efficiency improvements within the firm. Robust standards of practice, peer assessment, and disciplinary repercussions are also professional institutional solutions for informing/improving incentive and monitoring structures. In addition to pecuniary incentives, an ethos of service and loyalty and codes of ethics can also be effective governance mechanisms.

What moral obligations do managers have in virtue of occupying a professional role? Beyond mere maximization of utility and profit, economic theory requires economic agents to behave in certain specific ways. As I showed in the previous chapter, an efficient allocation and distribution of resources requires certain conditions to be in place, and from these conditions certain “efficiency imperatives” may be derived. Applied to this discussion, management professionalism involves following efficiency imperatives despite the prevalent moral hazard problems that make evaluation difficult.

Here are some instances of moral hazard:

(i) The manager could make overly risky financial decisions, thereby posing a moral hazard risk to her superiors who may experience financial and reputational loss;

(ii) Managers could make insufficiently risky decisions, thereby lowering profit margins for their superiors;

---

(iii) Superiors often have difficulty accessing information that allows them to evaluate managerial performance and decision-making. Managers could do nothing to change this status-quo, or could act so as to reinforce it;

(iv) Not only do shareholders not know what the managers are up to, they also don’t have the legal authority to control them, because of the “business judgment rule.” The judicially-created business judgment rule allows courts to defer to the business judgments of managers, on the condition that the managers’ decision fulfills the fiduciary obligation of being reasonably well-informed and being made in good faith. According to the Supreme Court of Canada, the court’s enforcement of the duties of managers has to “[respect] the fact that [managers] and officers often have business expertise that courts do not.” Managers could elect to use the business judgment rule to their personal benefit;

(v) Executive compensation is also a problem, since while there is transparency with respect to what the compensation should be, managers could manipulate how their performance is evaluated. Managers may also manipulate stock prices to increase the value of options, etc.

Managers are in an agency relationship with their superiors in the corporate hierarchy, but this is not the only internal agency relationship within the corporate structure. Managers are also in principal-agent relationships with individuals below them in the chain of command.

Manager vs. employees

The typical manager-employee analysis views the manager as the principal who hires the employee as an agent. But the roles of principals and agents can shift, such that an individual is an agent to someone in one relationship, and a principal to that same individual in a different relationship. I consider the role of managers as agents of employees, since I want to underscore the professional nature of the manager’s role. Here, a manager is an agent, and an employee is a principal, in so far as the agent (manager) promises to deliver on certain objectives (e.g. wages, safe working environment, etc.) in return for the employee’s work.

---

44 See Peoples Department Stores Inc. (Trustee of) v. Wise, 2004 SCC 68 at para 64.
Information asymmetry in the form of moral hazard may entail that an employee does not know how to evaluate her manager’s efforts to deliver on promises made to her at the time of signing the contract. For example, the manager might make unsound financial decisions which lead to the bankruptcy of the firm, thereby posing a moral hazard risk to the employee who may lose her salary and her job. Consider the 12,500 Lehmann Brothers employees who were left jobless and blindsided by the bankruptcy news of their company in 2008.\textsuperscript{45} Employees also develop employer-specific human capital, which is similar to making an investment in the firm.\textsuperscript{46} When employees are laid off, those investments are in effect lost since industry-specific human capital is not easily transferrable.

Here are some relevant examples of moral hazard:

(i) Managers could decide not to provide adequate information to employees about the financial health of the company, so that employees do not expect to be laid off;

(ii) Managers could hide information from, or actively mislead employees, regarding the competitiveness of employee benefit packages and salaries;

(iii) Managers could neglect to provide support, or actively stall opportunities for employee retraining and recertification.

Based on this approach, the failure to fulfill efficiency-promoting hypothetical imperatives is a distortion of managerial professionalism. Inappropriately risky behavior that eventually leads to laying off employees and losses to shareholders and owners is thus a deviation of the professional role for managers.

\textsuperscript{45} The actual number of employees left jobless following the company’s bankruptcy was 25,000, but about half that number went to Barclays and Nomura Holdings after the two companies purchased divisions of Lehman Brothers. (Accessed February 22, 2012) http://www.thedailybeast.com/newsweek/2009/09/13/left-behind-by-lehman.html

\textsuperscript{46} See Margaret M. Blair, “Firm-Specific Human Capital and Theories of the Firm” In Employees and Corporate Governance, Margaret M. Blair and Mark J. Roe, eds. (Washington: Brookings, 1999).
One question that needs to be addressed at this stage is whether there is ultimately a
difference between considerations about the morality of the market and the morality of the
professions, when it comes to thinking about corporations. If all we need is a set of efficiency
imperatives, then what work are the professional trappings doing here? The answer lies in the
professions’ increased relative reliance on non-market governance mechanisms. Ethics is an
attractive governance mechanism for the professions from a transaction cost perspective. As we
have already seen, professional institutional structures draw on a combination of market,
government, and ethics governance mechanisms, with the goal of minimizing transaction costs.
But while market and government solutions have garnered plenty of attention from management
scholars, ethics continues to receive less attention as a viable governance mechanism in the
corporation. This is an area that needs attention, and my model of management professionalism
aims at this task.

Ethics is an attractive governance mechanism from a transaction cost perspective, because
it creates trust in managers, the same way that the presence of ethics governance mechanisms
like the Hippocratic oath creates trust in the services of physicians. Norms like loyalty,
cooperativeness, honesty, and mutual benefit, and trust all play important roles as internal
incentives in corporations. Of course, there are more fundamental normatively substantive
conversations to be had about ethics in the market: notions of fairness, autonomy, and harm, for
example, deserve special attention in a complete analysis of professional and non-professional
markets. I have purposefully chosen to focus my attention on the more limited task of identifying
the role of ethics in promoting efficiency. This is to a large extent due to the reconstructive
attraction of this approach, since existing corporate practices already draw on ethics as an


47 See Arrow “Uncertainty and the Welfare Economics of Medical Care.”
implicit efficiency-promoting norm that structures business interactions. Even at this minimal level, however, many decision-makers and policy-designers have difficulty justifying a discussion of ethics in the corporate setting because of controversies surrounding the nature of the good.48

In fact, despite ethics being a descriptively accurate feature of how the firm actually operates, ethics research has had much less uptake than it deserves in the study of corporations and management practices. The relative transaction cost improving attractiveness of ethics has been for the most part off the radar of management scholars. One reason for this is that ethics is difficult to model for game theorists given the background assumptions of their approach. Another reason is that scholars in the management field have to a large extent been critical of applying the logic of transaction cost economics to firms.49

To better appreciate the lack of adequate attention to ethics, we may note that the core governance mechanism under study in management school “corporate social responsibility” classes is the market, with a focus on the marketing and the reputation-oriented costs of ethics scandals. Front-runner “progressive” schools focus also on the role of the government and teach “corporate citizenship.” In many of those classes, market solutions as well as government solutions are discussed, along with an overview of business practice in competitive, regulatory environments. Thus, the efficiency-promoting benefits of ethics governance mechanisms has been deemed too normatively demanding to take on under core business school curricula. The

48 The deputy minister in charge of my unit at the Ontario Ministry of Finance joked collegially about the prospects of leaving an economic policy job to do doctoral research in “business ethics”.
Society for Business Ethics, the professional body that underscores the role of ethics in business scholarship, represents only a small fraction of general management researchers.\textsuperscript{50}

So far, I have shown that my reconstructive normative model of management professionalism views managers as professionals, because emphasizing ethics in management would in turn prove efficiency-improving. In the next section, I show that we can find descriptive evidence for the role that ethics already plays in the practice of management by looking at the history of salaried managerial positions.

6.2.2 The Separation of Ownership & Control

Thinking about corporations over the past century more broadly, we may note that the moral hazard problems that I warned against in the previous section, have not, on balance, been debilitating. We have financial crises from time to time, and these crises are sometimes directly related to the failure of managerial responsibility. But as economic historians have pointed out, markets are also responsible for an incredible array of remarkable achievements, impacting growth and industrialization around the world.\textsuperscript{51} Corporations have been at the forefront of these achievements (in terms of efficiency, as opposed to, say, fairness, human rights, or environmental sustainability).

One way to explain this success is to note the role of managerial incentive structures and performance pay. But external incentives can only go so far in explaining the dominance of

\textsuperscript{50} In comparison, the Academy of Management (the professional body for the study of management) has continued to grow, alongside the expansion of business schools in North America over the past couple of decades.

corporations. The widespread growth of corporations may be taken to suggest that managers were also trusted to occupy their position despite the information asymmetries that made it difficult to evaluate them. Recall from chapter four, that a core descriptive assumption underlying my approach is that the proper fulfillment of the professional role involves carrying out a generally positive social function. Even if practitioners currently have a tendency to ignore the social implications of their roles, my assumption is that professional institutional structures were initially configured to serve social aims, and these design features may continue to impact how given profession is practiced today. With this assumption in mind, I claim that the increased reliance on trust that is traditionally typical of professionals (as opposed to say, used car sales people) may in fact be observed through the phenomenon of the separation of ownership and control in corporations.

How might the separation of ownership and control be related to trust? We know that firms are subject to Pareto-improvements when managers impose higher standards of care and practice upon their work. This form of utility improvement is rather obvious. But there is a further crucial gain in utility that is somewhat less obvious, and which explains why the exercise of robust professional management norms can be extensively socially beneficial. As I showed in chapter three, professional structures can save lost transactions and make Pareto-improvements by preventing dead-weight losses.\(^52\) Information asymmetries make it difficult for clients to evaluate the quality of the services they receive from professionals. Moreover, since the provision of such services does not occur on a regular basis, clients have a limited chance to learn through repeat

purchases, and the professionals’ reputation will also have a minimal impact.\textsuperscript{53} Thus, we need to draw on non-market oriented structures, for example the norms of service orientation and promise-keeping, in order to counter information asymmetries.

Applying this analysis to corporations, arguably, a core transaction saved by the managerial professional structure is the separation of ownership and control.\textsuperscript{54} Ownership is a position shared by members of the hierarchical organizational structure superior to the manager. Control, in turn, is the job of the manager. The separation of ownership and control allows a division between decision and risk-bearing functions in the corporation, which in turn produces the benefits of specialization in management.\textsuperscript{55} This separation facilitates an effective approach for controlling agency problems that arise as a result of the separation of the ratification and monitoring of decisions, on the one hand, from the initiation and implementation of decisions, on the other.\textsuperscript{56}

A core problem in separating the ownership and management structures is that the corporation’s total value is less than it would be if the managers themselves were the sole owner. As Adam Smith originally put the problem:

\begin{itemize}
  \item Fama and Jensen, “Separation of Ownership and Control,” 302.
\end{itemize}
The directors of such [joint-stock] companies, however, being the managers rather of other people’s money than of their own, it cannot well be expected, that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own. Like the stewards of a rich man, they are apt to consider attention to small matters as not for their master’s honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company.57

But if managers generally fail to maximize value efficiently, then why were salaried managers able to gain traction and grow exponentially as a group in the 20th century? As Jenson and Meckling have argued, the managers’ failure to maximize the value of the corporation is in fact consistent with efficiency.58 The reduced value of the corporation due to the managers’ failure is inefficient, “only in comparison to a world in which we could obtain compliance of the agent to the principal’s wishes at zero cost or in comparison to a hypothetical world in which the agency costs were lower.”59 But as it turns out, agency costs are unavoidable, and the value of the manager’s work is Pareto-improving.

A virtue of my approach to professional management is that it allows me to provide a complementary explanation for the growth of the salaried managerial position—an explanation that does not rely on pecuniary incentives, and one that has so far received less attention in the literature. Aside from the optimal agency costs of separating ownership and control, it is important to attend to the corporate organizational norms that allowed and facilitated the separation of ownership and control. The norms surrounding the managerial role arguably allow owners and managers to trust each other—enough to enter into transactions with each other that they would not enter into if they did not expect to reap mutual gains from the cooperation. In effect, in contrast with standard market contracts fraught with a deficit of trust (recall the used

59 Ibid., 336.
car sales experience), managers, somehow, have convinced owners that they are not likely to cheat their contracts.

One way to explain this phenomenon is through traditional agency costs: it made sense for owners to hire managers even if managers did not end up working as hard as they would if they were owners themselves. But while managers may not have worked as hard as owners, arguably they still worked hard enough for the phenomenon of the separation of ownership and control to grow exponentially. The question is why managers worked “hard enough,” and did not cheat whenever they could. Over time, expectations have been formed about a culture of management. And somehow, at least in its early days of development, the culture of management acquired a different form than, say, the culture of the used car salespersons. How might we explain this? One (reconstructive) solution is to attend to the cultural norms that surrounded managers at the time as a source of motivation against cheating. Norms such as loyalty, honesty, and promise-keeping could arguably have influenced managerial work. The presence of these norms in turn could have made managers to some extent trustworthy in the eyes of owners. In this manner, managers could facilitate many times more profits and efficiency gains for companies, as compared to what the owners would be able to create on their own. This transaction between managers and owners could have been “lost,” if owners were not able to see past the adverse selection problems that surrounded managers. Managers in effect signaled successfully about their efficiency-promoting potentials, at least in part via norms that made them trustworthy.

Organizational structures are arguably strongly influenced by the social environment within which the economy operates.60 Consider the following example as an illustration of this point. A 1995 study of the industrial structure of countries like the People’s Republic of China, 60 Francis Fukuyama, Trust: The Social Virtues and the Creation of Prosperity (New York: Free Press, 1995).
Taiwan, Hong Kong, and Singapore revealed the overwhelmingly small scale of economic enterprises in contrast with Japan, Korea, and Western countries.\(^{61}\) One reason for the difference in the size of companies in Chinese regions is arguably that they have had access to a different kind of social capital than companies in the West.\(^{62}\) Historically, families owned and managed almost all private businesses in Chinese societies.\(^{63}\) In contrast with the Western corporate experience, family ownership in Chinese corporations continued to be prevalent late into the 20\(^{th}\) century, and Chinese tycoons did not begin to hire salaried managers until a couple of decades ago.\(^{64}\) As Francis Fukuyama explains, in Chinese societies, “*family businesses seem to have difficulty in making the transition from family to professional management...*”\(^{65}\) The reason for this difficulty might have been that the Chinese tend to distrust people outside family and kinship groups, which blocks unrelated people from managing enterprises.\(^{66}\)

To be sure, Fukuyama’s observation about Chinese family-owned corporations does not prove that the separation of ownership and control requires trustworthy managers. Rather, the point is that the social environment and the norms that guide everyday interactions, e.g. promise-keeping, honesty, and trust, function as a sort of social capital, and affect the operation of corporate governance structures. Consider an imaginary society where no one ever lies. Here, the separation of ownership and control might supplement market-solution-style incentives with a heavy reliance on honesty, as a transaction-cost minimizing governance strategy. In Chinese

\(^{61}\) Ibid., 71.
\(^{62}\) Ibid., 74.
\(^{65}\) Fukuyama, *Trust*, 74.
societies where the social capital of trust used to apply only to family members (as opposed to strangers), the ownership and control structures may have been accordingly restricted to the family ranks rather than being open to strangers.

As I have shown, external incentive structures and norms can be viewed as combatting the adverse selection problems that surrounded the birth of management. If norms like trust and loyalty in fact contributed to the successful separation of ownership and control, this would constitute descriptive efficacy for my normative model of management professionalism. In this manner, managers are professionals since they typically draw on trust-creating, trust-preserving mechanisms to respond to information asymmetries and promote efficiency. Now that I have made a case for viewing managers as professionals, I turn to exploring and evaluating their moral obligations.

6.3 Managerial Role Morality

In this section, I apply my findings about professional morality from the previous chapter to the role of professional managers. Recall my support for the *generative* nature of social roles. As I argued, roles can be constrained instances of deontic weakening and sui generis source of moral guidance. As I demonstrate in this section, partiality, division of moral labor, and adversarialism, are three controversial instances where managers sometimes claim a permission to do what would otherwise be forbidden.

I begin by showing that following the formula I set up in figure 5.3 in the previous chapter, dividing the personal and impartial realms is a simple but important step in the process of determining managerial responsibility. Consider a manager’s commitment to the auditing
institution, for example. This involves a commitment to the neutral, impartial, and universal principles according to which accurate and fair audits are guided. The manager can thus offload part of her impartial stance onto the auditing institution and commit to the institution as a universally valuable entity. This commitment, however, is separate from the manager’s personal interests, friendships, and religious convictions. The individual may in fact have personal interests in not getting audited. The division of moral labor allows us to explain these two different levels of personal and impartial interest. Note that although this distinction seems simple enough, large-scale ethics scandals often come about from a failure to uphold this separation of the personal and impartial interests. The ex-Goldman Sachs Group Inc. director, Rajat Gupta, for example, was convicted in 2012 in the largest hedge fund insider trading scheme in U.S. history, for leaking stock tips to the co-founder of Galleon Group LLC, Raj Rajaratnam.\footnote{See http://www.bloomberg.com/news/2012-10-15/gupta-s-admirers-urge-mercy-as-insider-sentence-nears.html (Accessed March 7, 2013)} The Wharton Business School classmates had been family friends for decades, served on boards of various foundations, and started a number of companies together over the years.

Let’s now turn to discussing the notion of partiality in management. Within the realm of institutional morality, partiality entails that managers are obliged to promote the interests of the principals with whom they are in principal-agent relationships, along the internal hierarchical structure of the corporation. At minimum, this involves acting so as to not take advantage of moral hazard. Efficiency is thus taken as an end that generates deontic constraints upon professional managers. Consider the following example. A risk manager’s partial obligations might involve committing her company to accountability standards (e.g. the Global Compact or the Equator Principles). These standards require extensive social and environmental
commitments that use up company time and resources, but which purport to promote long-term efficiency through proactive stakeholder dialogues and beyond-compliance commitments to international standards. Promoting efficiency in this manner is permissible for individuals outside the managerial role, but not required of them. The risk manager’s role obligations generate deontic strengthening by requiring him to do what is otherwise merely permissible.

Does the role ever generate deontic weakening? In other words, can the managerial role ever permit what would otherwise be forbidden? I provide an answer to this question through the following example. Consider a fictional corporation x. For each professional manager that works at x, the personal standpoint, on the one hand, and the impartial/professional standpoint, on the other hand, coexist. Following discussion of Figure 5.4 in the previous chapter, part of the manager’s impartial moral concern is externalized onto managerial institutional structures, such as those that guide the efficient and profitable operation of x. Hence, we have a division of moral labor. We may then acknowledge that the way social institutions conspire to achieve outcomes, involves differentiating responsibilities into distinct roles. The institutional role of one manager differs from another, with titles such as human resources (HR) manager, research and development (R&D) manager, accounting manager, marketing manager, risk manager, etc.

Within this realm, managers have externalized part of their impersonal commitment onto their roles via the institutional structure of the corporation. All-things-considered moral considerations require all managers to follow certain maxims (e.g. no insider trading, no shirking, no stealing, no harmful activity, etc. in the process of promoting Pareto-efficiency). However, we can readily accept that each managerial role (HR, R&D, etc.) requires individuals to fulfill specific partial obligations. For example, the human resources manager’s duty of partiality might require her to fire an employee, based on recommendations about the financial
state of the firm by the accounting manager. The risk manager’s and the research and development manager’s partial concern might find that same employee’s dismissal inadvisable (e.g. because she might sue the company due to the nature of her contract, or because she has useful research skills). From the perspective of ordinary morality, managers should uniformly refrain from posing harm to employees. Individuals outside the managerial role, e.g. friends, family members, and fellow employees of someone who is about to get fired would agree. Yet, each managerial role in this scenario may produce a different prescription of permissible action, distinct from ordinary morality, because the managerial role can supply certain (limited but nevertheless) genuine reasons for action.

Let’s assume the managers decide to proceed to fire the employee to save on labor costs following the financial crisis. The human resource manager in charge of delivering the bad news may personally be sad or disappointed about having to let this employee go. After being fired, the employee will be left jobless, the psychological impact of being laid off will be painful, and the material and emotional impact on the employee’s family will be certainly harmful. If it were not for the manager’s role, bringing on this kind of harm onto another human being would be morally unacceptable. In other words, the manager’s role seems to generate deontic weakening, by permitting what would otherwise be forbidden. But what about the harm that comes about from firing the employee? Arguably, this harm can be mitigated to some extent. Codetermination policies in German corporations may be a good governance model for mitigating the harmful effects of lay-offs. Employment contracts, employee unions, etc. are other institutional

68 “Codetermination” (Mitbestimmung) legislation requires substantial labor participation in the governance of large German corporations. German corporations typically have a dual board structure: A “management” board is charged with carrying out the day-to-day business, while a “supervisory” board guides the company’s general agenda and broader policy decisions. These supervisory boards have representatives not just from the company’s shareholders, but also from the company’s employees. See Frank Dornseifer, ed. Corporate Business Forms in Europe: A Compendium of Public and Private Limited Companies in Europe (Munich: Sellier European Law Publishers,
responses that seek to alleviate such harm. Well-designed institutional structures that reduce information asymmetries surrounding employees (e.g. the codetermination corporate board system which facilitates substantial labor participation in the governance of corporations) are also pivotal counterparts of corporation structures in this context.

A similar analysis holds in adversarial scenarios where one company competes with another, eventually pushing the losing party out of the market and out of business. Limited instances of harm are arguably permitted here by the managerial role, since the role modifies obligations to some extent while focusing on the goal of profit-maximization. Similarly, the (enforcer) hockey-player’s role modifies the norms governing hockey games so that fist-fights – of the variety that would send players to jail off-ice – are routine and expected during game time. This is not to say that all adversarial roles can generate deontic weakening, or that all impermissible activity can become required by all roles. Rather, my point is that within specific institutional settings, the generative nature of the professional role entails that it is possible for roles to sometimes modify obligations, so that we may be permitted to do what is otherwise forbidden. The presence of alternative institutional safeguards (that guard against harm to employees, and more broadly, guide corporations to refrain from taking advantage of market failures in the process of profit- maximization) is a background requirement for the generation of deontic weakening by the managerial role.

Thus, an argument can be provided for the constrained moral relevance of controversial obligations arising from partiality, division of moral labor, and adversarialism within corporation institutional structures. Such an appreciation of the limited force of the professional managerial role goes a long way towards responding to the dreaded cocktail party criticism: Lawyers get accused of being liars and business ethics gets called an oxymoron.

### 6.4 Implications for Business Ethics

Developing a model of management professionalism has important implications for the field of business ethics. Shareholder theory has been a dominant conceptualization of the firm, whereby managers’ obligations are owed to shareholders. Within this paradigm, research in transaction cost economics and agency theory, and an adherence to neoclassical economic methodologies, have shaped both the theoretical and the practical understanding of the management role and managerial social obligations. As critics have pointed out, however, the shareholder perspective cannot account for much of the social and institutional intricacies of business organizations. This is by now a familiar stance advanced by proponents of the stakeholder theory of business ethics. These proponents have traditionally worried that the influence of shareholders may force managers to override the responsibilities they owe to other entities. The solution, stakeholder theorists suggest, is to adopt a framework for responsibility that takes into consideration not just shareholders and members of the corporate hierarchy, but

---

69 See, for example, Williamson, *Markets and Hierarchies*; Jensen and Meckling, “Theory of the Firm”.

also other individuals and entities impacted by the work of managers. The problem is that stakeholder theory does not quite qualify as a “theory.” Proponents of this view see themselves as providing an alternative to shareholder primacy, but their proposals are arguably best characterized as perspectives or paradigms, rather than theories.\(^71\) Still, one in every five articles published in the two leading business ethics journals thus far in the 2000s discusses stakeholder theory.\(^72\)

Leading scholars in the field have been pointing out for some time that we now know enough about the shortcomings of stakeholder theory to “put a stake through stakeholder theory” (to use language that resonates in the current age of vampire fandom) and move on to the next big theoretical debate.\(^73\) One alternative framework of responsibility is embodied in the ethos of professionalization. This line of argument is still in the very early stages of development, but it is gaining momentum with the introduction of professional oaths in certain sectors (such as banking and economic policy-making).\(^74\) According to Khurana, professionalism in management entails that the corporate interest be “subordinate” to the “public interest.”\(^75\) Thus, he argues for a return to “higher aims,” and has proposed (along with Nitin Nohria, Dean of Harvard Business School) an “oath” for MBA students to operate as responsible professionals upon graduation.\(^76\)


\(^72\) Ibid., 2. The relevant publications are *Journal of Business Ethics* and *Business Ethics Quarterly*.


\(^74\) In the wake of the recent financial crisis, a banking oath was uniformly adopted in the Netherlands in 2010. See George F. DeMartino, *The Economist's Oath: On the Need for and Content of Professional Economic Ethics* (Oxford: Oxford University Press, 2011). A professional oath has also been proposed for MBA students. See Max Anderson and Peter Escher, *The MBA Oath: Setting a Higher Standard for Business Leaders* (New York: Portfolio, 2010).

\(^75\) Khurana, *From Higher Aims to Hired Hands*, 304.

\(^76\) Khurana and Nohria, “It's Time to Make Management a True Profession.” See also Anderson and Escher, *The MBA Oath*. 

Although Khurana’s position is headed in the right direction, as I argued earlier in this chapter, it is inadequately substantiated. Familiarity with the nature of professions and with the myriad scholarly studies of them reveals many instances of what I have called distortions of professionalization, where professionals have used their title and institutional authority for cartelizing purposes. Far from fostering the public interest, the title “professional” has been used to promote anti-competitive, rent-seeking behavior and to facilitate exclusion and discrimination against potential new members. Scathing critiques abound of over-paid professionals who, under the auspices of contributing to social order and doing “gentlemanly work,” have exerted self-serving authority and reaped benefits at the expense of those they supposedly serve.  

Critics of professionalization have argued that the professional rhetoric is often taken advantage of by profit-seekers, who are in reality “wolves in sheep clothing, monopolists who live by the rule of caveat emptor, but lack the integrity to admit it.”

We need a theoretically robust definition of professions, and an understanding of the normative essence of professionalization, in order to properly identify deviations from professionalism. Otherwise, we risk losing meaningful traction from the classification of professions. We also need a nuanced approach to professional ethics that explains the extent to which professional obligations are moral, even when partial obligations of managers differ from impartial general morality. With this theoretical ammunition, which this dissertation does some work to provide, management professionalism starts to look like a useful theoretical framework to replace the shareholder-stakeholder dichotomy.

---

77 See, for example, Greenstein, “Against Professionalism”; Johnson, Professions and Power; Larson, The Rise of Professionalism; Lierbermann, The Tyranny of the Experts; Melosh, The Physician’s Hand; Robinson, Pritchard, and Ellin, eds., Profits and Professions.

78 Thomas Haskell, “Professionalism versus Capitalism”, 181.
More fundamentally, the feature of professional management that makes it theoretically fruitful is that it is a market-centered theory of business. Broadly speaking, the market-centered approach to business ethics is an approach of grounding the norms that structure the relation between business and society.\(^79\) According to this approach, the standards of conduct in business ought to be explained in terms of the normative presuppositions of the market. \(^?\) These presuppositions entail that responsible business conduct preserves a) the conditions necessary for markets to be socially beneficial, and b) the behaviors needed for market transactions to be free and fair. By socially beneficial, here, I mean, broadly speaking, Pareto-improving. A key dimension to this theory, which is receiving more attention recently, is the “market failures approach” to business ethics.\(^80\) According to this approach, firms have a basic responsibility not to exploit market failures in the process of making profit. Profit-seeking facilitates the Pareto-efficient allocation of goods and services in the free market, but managers should observe constraints that limit or prohibit practices that lead to information asymmetries, collusion, externalizing costs, and other market failures.

My definition of professions and my approach to management professionalism contribute to the market failures approach to business ethics. As I have outlined, the obligations of professionals involve responding to market failures that arise from information asymmetries. To


some extent, these obligations parallel the obligations of managers in the standard market failures approach in the market for goods and services. By defining managers and their organizational realm as a profession, in effect, I propose that we imagine these two market realms (the market for goods and services, and the professional market) as merged into one, under the broad banner of the private economic realm. To be sure, these realms remain different in important ways, and there are multiple distinct frameworks and institutional structures within each market that deserve separate analysis. But visualizing the two realms as one stresses my point that institutional structures in both realms respond to market failures by drawing on market as well as non-market institutional mechanisms. See figure 6.1 below.

Figure 6.1

---

Maximizing profit is Pareto-improving in the market for goods and services, given certain efficiency-promoting deontic constraints against taking advantage of market failures. Similarly, professional services are Pareto-improving in the professional service market, given certain efficiency-promoting deontic constraints against taking advantage of information asymmetries. Thus, fundamentally, managers of private economic entities operate as professional service-providers would with the goal of profit-maximization.

Private economic firms aim at profit-maximization, but as I have shown, the obligations of managers to the chain of command within the firm do not involve only profit-maximization. More broadly, the managers’ responsibility, parallel to the professionals’ responsibility, is to fulfill principal-agent obligations. When the principal is an employee, the obligation might be to provide a safe working environment. When the principal is the owner of a non-profit environmental organization, the obligation might be to clean up a lake or work to preserve an endangered species. The benefit of theorizing about managers as professional service providers within the market failures approach is the extra normative guidance we can tap into within the market for professional services. Beyond incentives and institutional regulations, the trust-creating, trust-preserving professional institutional mechanisms are fruitful in a normative model of management professionalism.

6.4.1 Conclusion

In sum, as I have shown, the question of whether the concept of professionalization is appropriate for analyzing business and the role of managers in society has remained a persistent problem for management scholars over the past century. A survey of the management literature
reveals inadequate attention to the definitional uncertainty surrounding the professions. There is also a questionable underlying assumption among management scholars (for and against professionalization), that professionals necessarily strive to bring about the public interest in an unselfish, altruistic manner. An adequately nuanced analysis of the definition of professions and the normative nature of professionalization is thus missing in the literature.

To fill this gap, I have provided a normative model of professionalization that stays clear of glorified altruistic characterizations, while remaining sensitive to the socially beneficial nature of professional work. I argued that managers are professionals, since in addition to external market-oriented incentives, they typically appeal to trust-creating norms in the process of making Pareto-improvements in the firm. Management professionalism involves following a set of hypothetical efficiency imperatives that guide managers to do their job, despite the prevalent moral hazard problems in the corporation. A failure to fulfill efficiency-promoting imperatives is thus a distortion of managerial professionalism.

One virtue of my approach is that it provides an alternative explanation to the mainstream understanding of the phenomenon of the separation of ownership and control, and the growth of the salaried managerial position in the twentieth century. Instead of merely explaining the phenomenon via optimal agency costs and external pecuniary incentives, I drew attention to the presence of norms (e.g. loyalty) as a prevalent social capital. The separation of ownership and control is a transaction that might have been lost amidst the adverse selection information asymmetries that affected the relationship between owners and managers. Managers in effect signaled successfully about their efficiency-promoting potentials, at least in part via norms that made them trustworthy. Thus, the increased reliance on trust that is traditionally typical of professionals (as opposed to say, used car salespersons) may in fact be observed in the history of
the development of management practice. Drawing on my definition of professions, I then outlined the normative features that characterize managers as concerned with more than merely self-serving – yet less than altruistic – ends. The managerial role generates a set of deontic constraints towards the end of efficiency-promotion within the internal hierarchical chain of command of the corporation. Given certain background conditions, and due to their generative moral nature, managerial roles sometimes qualify as a sui generis source of moral guidance, as demonstrated through the examples of partiality, division of moral labor, and adversarialism.

Two outstanding points require attention here. First, my focus in this chapter has been on the moral obligations of management professionals within the corporate hierarchy of the firm. Although I have not discussed extra-firm efficiency promotion, a complete theory of managerial professionalism must address such concerns. Clients and customers that might purchase a firm’s products, other corporations with which the firm competes, and third parties impacted by corporate activity (e.g. people in a neighborhood where a factory is built). Beyond the “internal” principal-agent relationships with members of the hierarchical chain of command (shareholders, owners, employees, etc.) management practice has extra-firm moral relevance. The subprime mortgage crisis and resulting international financial woes are a reminder of this fact. These “external” managerial obligations need special attention.

This observation extends beyond management practice to professional practice in general. For example, as a professional civil servant, individuals owe obligations not just to their chain of command (minister), but also, more broadly, to the electorate and to promoting the public interest. The external impact of professional work may be positive or negative. Enhanced services to the electorate in this province, for example, may have an adverse effect on the electorate in other regions. The city of Toronto boasts a reputation for being clean among
similarly sized urban centers, but residents of Michigan have long complained about Toronto trash deposited in their landfills. Civil servants in charge of garbage disposal are fulfilling their internal obligations well by providing timely and efficient garbage removal, but they are claiming a permission to ignore the negative effects on external parties who bear the undue costs of that service. Michigan citizens might conceivably be disgruntled about the garbage dump operations in their neighborhoods, despite the fact that Toronto pays to use those facilities. The impact of waste on environmental degradation in general and on future generations fit under this same category of external concerns. A doctor might similarly prescribe a well-priced pain killer that effectively cures her patient of an ailment, but in so far as that drug causes anti-biotic resistance among the public at large, the drug adversely affects patients in the long run who may not even benefit from the painkilling remedies. Ordinary morality typically specifies (and professionals sometimes claim permission to ignore) requirements pertaining to the third-party impact of professional work, as demonstrated in the examples of garbage disposal and drug prescription. These third-party concerns require attention in a complete theoretical treatment of the obligations of professionals.

Finally, in addition to internal obligations within principal-agent relationships, and external obligations beyond, professionals often owe “mutual” quid pro quo duties of reciprocity to their colleagues. These obligations are guided by norms of reciprocity and trust, but principal-agent relationships do not exist among professionals who fulfill these obligations. The agency risk approach to thinking about professional interactions does not provide much guidance regarding the morality of these quid-pro-quo obligations among professional agents. Most professional associations draw on peer-evaluation and peer-assessment mechanisms of some kind or other to provide a measure of quality control, especially given the information asymmetric nature of professional work, which makes it difficult for non-specialists to evaluate professional
performance. Taking part in such institutional mechanisms is a core professional responsibility, and takes up considerable time and commitment on the part of professionals.

For example, university professors spend a considerable amount of time providing services to their professional colleagues. Professors have various mutually beneficial obligations (e.g. reviewing papers, organizing conferences, serving on admissions, tenure review, and departmental administrative committees, etc.). We need a framework to explain why they owe these obligations. The standard agency theory approach captures relations with students, departments, schools, and universities, but leaves out collegial, reciprocal, quid-pro-quo obligations. A potential strategy worthy of further attention is to determine whether mutual obligations are a transaction cost reducing norm in professional associations. Mutual obligations could thus make relative transaction cost-minimizing contributions to Pareto-improvements, in conjunction with market, government, and alternative non-market solutions. Applying this discussion to the management profession and introducing mechanisms for enhanced mutual obligations may also be fruitful.

Thinking about professionalism in terms of efficiency allows me to explain why we trace internal moral obligations along principal-agent relationships (since agency risks are efficiency-depleting). Applying this framework to corporations and the role of managers is instructive for the field of business ethics. Recall that the main task of the stakeholder approach to the firm was to provide a paradigm that articulates managerial moral obligations that go beyond shareholders. A market-failure-style approach, like the one I’ve presented in this dissertation, does that job well, and has more robust theoretical qualifications for doing so. Other than shareholders, managers have professional moral obligations to others along the chain of command including boards of directors, and crucially, employees.
By framing a normative model of management professionalism, my approach grounds prescriptive claims upon certain descriptive facts about norms and practices and their historical development. This allows me to provide a complementary explanation for the separation of ownership and control and the dominance of management professionalism over the past century. Visualizing the market for goods and services and the market for professional services as one merged realm, under the auspices of private economic entities, has the added explanatory benefit of letting management scholars tap into the professional norms of trust, loyalty, promise-keeping, etc., in formulating managerial moral obligations. This dissertation thus pushes the boundaries of traditional business ethics research by introducing a normative model of management professionalism and by proposing a dynamic limited justification of managerial moral obligations.
Bibliography


Baldus, J. *Disputatio an medicina sit legibus politicis preferenda vel econtra*, 1415.


Bernstein, M. A. “American Economists and the ‘Marginalist Revolution’: Notes on the
Intellectual and Social Contexts of Professionalization.” *Journal of Historical Sociology*
16, 1: 135-180.

Biddle, B. J. *Role Theory: Expectations, Identities, and Behaviors.* New York: Academic Press,
1979.

Black, A. *Guilds and Civil Society in European Political Thought from the Twelfth Century to


Blair, M. M. “Firm-Specific Human Capital and Theories of the Firm.” In *Employees and

1980.

1985.

Bledstein, B. J. *The Culture of Professionalism: The Middle Class and the Development of


217-38.


Boatright, J. “Ethics and Corporate Governance: Justifying the Role of the Shareholder.” In *The

Boswell, J. *Boswell’s Life of Johnson.* Edited by George Birkbeck Hill. Oxford: Clarendon Press,
1934.

Bowen, H. “Business Management: A Profession?” *Annals of the American Academy of Political


http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/vwapj/Professions%20study%20final%20E.pdf/$FILE/Professions%20study%20final%20E.pdf


DeLancy, F. P. *The Licensing of Professions in West Virginia.* Chicago, 1938.


Hughes, E. C. “Professions.” *Daedalus* 92 (1963), 655–68.


John of Arezzo, *De medicina et legum presentatia*. 1450-70.


Munk, W. *Roll of the Royal College of Physicians*. 1861.


Peoples Department Stores Inc. (Trustee of) v. Wise, 2004 SCC 68.


Petro Paolo Vergerio. De ingenius morbus et liberalibus adolescentiae studiis, c. 1400.


Webster, N. *An American Dictionary of the English Language*. New Haven, CT, 1828.

Webster, N. *An American Dictionary of the English Language: Revised and Enlarged by Chauncey A. Goodrich*. Springfield, MA, 1850.


