Corporate Governance: A Critical Assessment
Benjamin E. Hermalin

Contents
1 Introduction 1
2 What Is Corporate Governance? 4
3 Threats and Responses 6
4 The Determination of Responses: Corporate Governance in an Equilibrium Framework 10
5 The Many Dimensions of Governance 20
6 Dispersed Shareholders, Managerial Power, and Governance 26
7 More Thoughts on Shareholder Control 31
8 Where Does the Field Go Next? 33
References 35
1 Introduction

People invest in organizations. They invest their human capital and they invest their physical capital. In exchange, they expect some return and, all else equal, they would like the organizations in which they invest to maximize that return. A problem arises, however, when the people who invest in an organization are not the same people who control the organization. Lack of control makes the investors vulnerable to mismanagement of the organization, misallocation of its resources, or even misappropriation of their returns. In short, investors fear that their returns will not be maximized because their desires are not necessarily aligned with those of the controlling parties. In response to these fears, investors will wish to put in place various institutions to constrain the controlling parties to better act in their interests. In short, they will put in place some sort of governance structure for the organization.

Although this problem is endemic to all organizations (consider, e.g., the concerns of a philanthropist worried that a charitable organization will make the best use of her money), it has proved to be of critical importance with respect to corporations (especially, for-profit corporations). In the corporate setting, the parties typically deemed to have control are management and, if they exist, dominant shareholders. In some cases, the controlling parties are both (e.g., companies dominated by a single family, as the Ford Motor Company was historically by the Ford family or as many companies around the world are today).

The potential misalignments of incentives in corporations are many. For instance, managers are more inclined to tolerate their own incompetence than shareholders. They are also unlikely to see eye-to-eye with shareholders about managerial amenities, such as lavish offices, corporate jets, etc. The terms of an asset transfer from one firm to another could be disadvantageous to the first firm (and hence the majority of its investors), but advantageous to its dominant shareholder if the other company is one in which she owns even a greater proportion.\(^\text{2}\)

Concerns about corporate governance,\(^\text{3}\) in particular the issues that arise when investors and controlling

---

\(^{1}\) There is a limited literature on the governance of non-profit corporations. See, e.g., Bowen (1994) and Ehrenberg (2004).

\(^{2}\) Johnson et al. (2000) refer to such transfers as “tunneling.”

\(^{3}\) Becht et al. (2003, fn. 6) suggest that the term “corporate governance” was first used in 1960. The field is obviously considerably older.
parties are not one and the same, are long standing. Smith (1776) is an early, if not the earliest, example in economics:

The directors of such companies, however, being the managers rather of other peoples money than of their own, it cannot well be expected that they should watch over it with . . . anxious vigilance . . . Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such compan[ies] — Book v, Part iii, Article 1, “Of the Publick Works and Institutions which are necessary for facilitating particular Branches of Commerce,” paragraph 18.

The “modern” era of economic interest in corporate governance appears to have started in the 1920s, where initially the focus was on financial reporting. Berle and Means (1932) was perhaps the most important early modern work and it directly influenced a number of reforms undertaken by the New York Stock Exchange (NYSE). Berle and Means were among the first to note that the stark separation of ownership and control, common to the modern corporation, created a number of potentially serious problems. The 1950s and ’60s saw economists looking at issues of executive compensation and starting to model the divergence of interests between owners and managers (see Lewellen, 1968, as an example and survey of the former; Williamson, 1963, is a noteworthy example of the latter). An article of great importance in the evolution of the field was Jensen and Meckling (1976), which, in addition to surveying some of the literature to that point, identified many of the issues that have continued to occupy economists working on corporate governance to this day. As with much of life, good timing is essential and the renewed interest in corporate governance generated by Jensen and Meckling occurred just as the tools of information economics, game theory, and contract theory were coming to the fore in economics. These tools led to an explosive growth in the literature on organizations generally (as this handbook attests) and on corporate governance specifically. For instance, Becht et al. (2003) cite over 550 works in their survey of corporate governance, almost all of which are 1980 or later; surveying just the literature on boards of directors, Adams et al. (2010) cite approximately 170

---

4 An important work in this regard was Ripley (1927), which was initially a series of articles in the Atlantic Monthly. Baskin and Miranti (1997, p. 197) describe Ripley’s work as a “widely read polemic.” Scholz (1928) discusses “the commotion which was aroused in financial and official circles” by its publication.

5 See Baskin and Miranti (1997, p. 200) for a discussion of the influence of Berle and Means on policy.
This vast literature has generated a number of survey pieces. In addition to the two just mentioned, a sample of such surveys includes Shleifer and Vishny (1997), Tirole (2001), Farinha (2003), and Zingales (2008) on the literature overall; John and Senbet (1998) and Hermalin and Weisbach (2003) on boards of directors; Bertrand (2009) on CEOs; Devers et al. (2007) and Edmans and Gabaix (2009) on executive compensation; Bhagat and Jefferis (2002) on econometric issues; Denis and McConnell (2003) on international governance; and Durisin and Puzone (2009) on the intellectual structure of governance research. One can rightly ask whether another survey is needed; in particular, given the existence of the aforementioned surveys, the social benefit—even accounting for my egotistical exaggeration of my writing’s value—is modest and, given the huge literature, the cost—to me at least—immense. Consequently, I propose a different course of action: This chapter will seek to present an overarching framework from which to assess the issues of corporate governance and a means of framing the existing literature, both empirical and theoretical. Such an approach will, I believe, also be more conducive to suggesting future lines of research than a more traditional survey chapter.

Briefly, the chapter is organized as follows. The next section considers the scope of corporate governance, arguing that it is primarily the study of what happens when investors seek to protect themselves against mismanagement, misallocation, and misappropriation of their investments by those who control the corporations in which they wish to invest. The section thereafter seeks to tie these three-Ms to the various economic literatures that have explored these contract and agency problems. Section 4 deals with the fact that governance institutions are arrived at endogenously. This has important implications for how empirical work needs to be conducted. Perhaps more importantly, it also has important implications for how that work is interpreted. This second point is explored further in Section 5, where the consequences of the fact that different governance structures govern different firms are discussed. Section 6 focuses on the dispersion of ownership and its consequences. The last two sections are essentially conclusions. The first of these suggests that both the literature and policy have placed too great an emphasis on improving shareholder control.

Adams et al. report that over 200 papers on boards—not all of which they cite—were written in just the period 2003–08 alone; see their footnote 6.
The second sketches directions for future research.

2 What Is Corporate Governance?

I suggested above that corporate governance is concerned with protecting the interests of investors when they are not the ones wholly in control of the organization. Such a view echoes Shleifer and Vishny (1997, p. 737):

Corporate governance deals with the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment. How do the suppliers of finance get managers to return some of the profits to them? How do they make sure that managers do not steal the capital they supply or invest it in bad projects? How do suppliers of finance control managers?

For companies in which a few dominant shareholders have control, one could replace “managers” in this quote with “dominant shareholders.” One could, and perhaps should, also extend this to cover all suppliers of capital, including human capital. That said, the economic literature has been almost exclusively focused on the suppliers of finance (Roberts and van den Steen, 2003, and Tirole, 2001, are notable exceptions).

One could, following Tirole, take an even more encompassing view: Corporate governance is the means by which the externalities controlling parties generate are regulated.\(^7\) With respect to the very largest corporations, we are all affected to some degree by the actions of those in control. Typically, however, the set of individuals affected is more compact: its employees, customers, suppliers, local community, and, of course, investors—a set typically referred to as the firm’s stakeholders.

While one might properly be concerned with stakeholders’ welfare, such a broad view is not, in my opinion, the best way to look at corporate governance. First, the potential externalities of the controlling parties’ actions are many and include matters such as environmental degradation, the health and safety features of the products and services provided, even contributing to rush-hour congestion through transportation decisions. Further, these externalities may generate subsequent effects so, like the proverbial butterfly that

\[^{7}\text{Tirole, p. 4, writes \textit{I will \ldots define corporate governance as the design of institutions that induce or force management to internalize the welfare of stakeholders.}}\]
triggers a hurricane, everyone is potentially affected by each corporate decision. This is simply too broad a perspective to be workable in terms of analysis or policy setting.

Second, even if one limits the set of externalities at issue, the set of those affected is not necessarily clear. If a firm closes a plant, who would we say has been affected? The workers certainly and, perhaps, their local community as well. But what about the workers of companies that supplied the plant and who now face reduced hours or layoffs? What about their communities? Or what about a distant community whose economy is, in part, reliant on remittances from those who work at the plant? In this regard, a plus for focusing on investors is that they are readily identifiable.

Third, if the focus is on externalities, broadly defined, then what distinguishes corporate governance from other forms of environmental or health and safety regulation? Even a cursory examination suggests that there is some value in studying corporate governance apart from the general study of state regulation of economic activity.

One way in which corporate governance stands apart from state regulation generally is that corporate governance involves, to a great extent, the way private parties choose to regulate their own dealings. In contrast, the only way in which all stakeholders, broadly defined, can have a say is if the state intervenes on their behalf. Hence, the focus in the chapter will be on the institutions that the private parties—investors and managers—wish to put in place to govern their relations. Although corporate governance is highly regulated by the state, at least in most of the world and in all OECD countries, it is possible to consider corporate governance independently of any state action. Indeed, there is much to be gained by studying corporate governance assuming away state regulation. Only by knowing what private parties would otherwise put in place can we properly assess the wisdom—or lack thereof—in state regulation of firms’ governance structures. To be sure, one cannot wholly ignore the state—or at least the law—with respect to corporate governance; in particular, some institutions could not be created privately absent certain features of corporation law or political institutions.\(^8\) Nevertheless, the relevant parties still possess considerable freedom in designing the

---

\(^8\)See, for instance, Hansmann and Kraakman (2000a,b) and Kornhauser and MacLeod (in this volume) on this topic. Hansmann and Kraakman argue that private contracting could not replicate the corporation’s legal personhood. The ability of the corporation to own property in its own name—as opposed to having said property owned collectively by the shareholders—is essential for a corporation to function. Beyond simplifying the sale of shares, it also simplifies the corporation’s ability to sell and acquire assets. Perhaps most importantly, legal personhood prevents a shareholder
institutions that govern their relation.

3 Threats and Responses

Investors fear losing returns due to mismanagement of the organization, misallocation of its resources, or even outright misappropriation. Of these, the last is arguably the biggest threat, at least measured in terms of the response. A moment’s reflection plus some quick back-of-the-envelope calculations will show that expenditures on procedures and methods to detect and discourage misappropriation are the lion’s share of resources devoted to the control of the corporation. Every large organization has numerous procedures and methods to detect theft. Examples include auditing of transactions, requiring multiple levels of approval for transactions, and direct monitoring of employee activity. A non-trivial portion of various governance regulations, such as Sarbanes-Oxley, are concerned with these processes.9

Misappropriation can occur at any level of the organization. Within economics, the study of misappropriation mitigation at the lower levels has been essentially indirect, with the focus on matters such as management of hierarchies, budgeting, and collusion.10 With respect to misappropriation at the top, the literature has focused on how securities should be designed to mitigate the threat of misappropriation—of non-disgorgement of funds more specifically. For example, one topic has been how debt can induce those in control to disgorge the returns due investors. Townsend (1979), Jensen (1986), and Hart and Moore (1998) are some examples of this literature.11

A related concern is that a dominant shareholder may engage in self-dealing in one form or another. As Johnson et al. (2000), among others, observe, there is a danger that a controlling shareholder could engage in transfer pricing between companies she controls in a manner that is self-enriching (e.g., she has a firm from selling his or her portion of the assets or pledging them as collateral and, similarly, it prevents any creditor of a shareholder from seizing the firm’s assets as a form of recovery should the shareholder default on his or her debts. This is what Hansmann (this volume) refers to as “entity shielding.”

Fligstein (2008) discusses some of the political and societal structures necessary for a corporate governance regime aimed at maximizing firm value to be acceptable.


10See, for instance, Williamson (1975), Tirole (1986), and Mookherjee and Reichelstein (1997).

11See also Tirole (2006, §3.6–3.8) for a survey and discussion of the relevant literature.
of which she owns 50% buy inputs from a firm of which she owns 75% at an inflated price). If she is an executive, she could also approve excessive compensation for herself. There are other abuses of minority shareholders that could arise, such as freeze-outs (i.e., effectively compelling them to sell at less than true market value). Here, the principal “response” to this threat has been legal: It is laws on the protection of minority-shareholder rights that mainly serve to limit such behavior.

Although the distinction between misappropriation and misallocation is fuzzy—has the CEO stolen from shareholders when he has the company purchase a corporate jet?—we can view misallocation as a catch-all for various agency problems that arise because management and investors have differing preferences concerning the use of corporate resources, including managerial time. If we treat management’s allocation choices as a hidden action, then basic agency theory (e.g., Holmstrom, 1979, Shavell, 1979, and Grossman and Hart, 1983) provides a sense of the responses. In particular, we would postulate that companies would provide executives with various forms of incentive compensation. (Although it is important to note that incentive compensation is not the only way to respond to the threat of moral hazard; a point to which I return later.)

Executive compensation has been much studied (see, e.g., Bertrand, 2009, and Edmans and Gabaix, 2009, for two recent surveys). Although I will return to the topic later for a more critical discussion, it is worth mentioning some of the more important work in this area. The popular press and others (e.g., Bebchuk and Fried, 2004) have bemoaned the seemingly endless rise in executive compensation. For historical perspectives on this phenomenon, see Hadlock and Lumer (1997) and Frydman and Saks (in press). The literature has also sought to empirically test whether compensation practices are consistent with basic agency theory, with mixed results. For example, agency theory predicts that the optimal compensation scheme should seek to eliminate noise from the performance measure. Bertrand and Mullainathan (2001) find that CEO compensation schemes fail to do so. On the other hand, Gibbons and Murphy (1990) do find evidence of relative-performance compensation, a means of reducing noise. As Bertrand (2009) observes, one reaction to the rapid run-up in CEO compensation and the poor record of basic agency theory has been the development

---

12 For more on freeze-outs see Hermalin and Schwartz (1996) and the citations therein. There is also considerable cross-country variation in the law governing minority-shareholder protections, see La Porta et al. (1997).

13 If allocation decisions are observable to investors, then presumably management and investors will bargain to an efficient agreement concerning these decisions; that is, there won’t be misallocation.
of new models of compensation based on changes in the executive labor market or other trends. A partial list of these include Murphy and Zábojník (2003, 2004), Hermelin (2005), Gabaix and Landier (2008), and Terviö (2008).\footnote{Bertrand (2009, §3.2.3) offers a critical assessment of the Gabaix and Landier (2008) model with respect to its fit with the empirical evidence.}

The third investor concern, mismanagement, deals with the competence of the managers. Managers or would-be managers vary in ability. If talent were known \textit{ex ante}, then some market mechanism would allocate managers to firms (see, e.g., Terviö, 2008, and Gabaix and Landier, 2008, for such models). In equilibrium, investors are happy with their manager’s ability in the sense that they are unwilling to affect an alternative outcome.

Alternatively, managerial ability could be unknown, including, possibly, by the manager. This is the approach taken in Holmstrom’s seminal article on career concerns (1999).\footnote{The article first appeared in 1982 in a now out-of-print \textit{festschrift} for Lars Wahlbeck.} How well a newly appointed CEO will do is uncertain, but if his previous record is publicly known, then everyone should hold the same beliefs about his ability at the start of his tenure as CEO.\footnote{The underlying logic of career-concern models does not rely on the CEO’s being equally ignorant of his ability as other players. Provided there is no scope for him to clearly signal his type, the CEO will still have incentives to engage in the signal-jamming activities discussed here. The technical difficulties that would, however, arise are daunting, which is why the equal-ignorance assumption prevails in these models. Another variation—unexplored to the best of my knowledge—would be for the CEO to hold a biased estimate of his ability, consistent with some of the excessive optimism and overconfidence models in the behavioral economics literature. See, for example, Malmendier and Tate (2005). Bertrand (2009, §4.2) contains a nice survey of the literature on “cognitively challenged CEOs.”} In addition to Holmstrom (1999), other articles adopting a career-concern approach include Holmstrom and Ricart i Costa (1986), Stein (1989), Hermelin and Weisbach (1998), and Hermelin (2005).\footnote{Gibbons and Murphy (1992) is an empirical study of career concerns, which finds evidence to support the idea that firms increase explicit compensation to compensate for the reduced career concerns CEOs have toward the end of their careers.} In these models, players learn information that improves their estimate of the CEO’s ability. Because the CEO is rewarded if his estimated ability increases (he gets, e.g., higher salary offers) or punished if his estimated ability falls (he, e.g., gets fired), the CEO has an incentive to manipulate what information is learned by the other players. This manipulation can be beneficial to the firm (the CEO works harder, as hypothesized by Fama, 1980), but it can also create agency problems (see,
in particular, Stein, 1989). For instance, the CEO may boost signals of his ability in the short term in ways that have long-term adverse effects that outweigh, in NPV terms, any benefit the boost might create. An example of this is a CEO who behaves in a myopic fashion to boost performance in the short term at the expense of long-term value (see Stein). Alternatively, the CEO may choose actions based on how much or little they reveal of his ability rather than based on the investors’ best interests (see Holmstrom and Ricart i Costa). In other words, investors’ desires to limit mismanagement can create costly agency problems.

Another feature of these models is that improving the estimation of CEO ability exposes the CEO to risk. Under certain circumstances, he will require compensation for bearing that risk, which in turn could motivate investors to limit that risk even at the cost of reducing the probability that a bad CEO will be fired. Hermalin and Weisbach (1998) and Hermalin (2005) consider, in different contexts, what the consequences of this tension are for the design of governance structures. In particular, these last two articles examine, among other issues, how the choice of corporate directors—the people who actually dismiss the CEO—balances the desire to be able to get rid of low-ability CEOs against the potential agency problems and compensation demands that monitoring induces.

Monitoring is another response to agency problems. Generally, the job of monitoring top management has been seen as a responsibility of a company’s board of directors (although the takeover market and proxy contents can also take on this role). Directors are also supposed to guard against misappropriation and misallocation. As noted earlier, the literature on directors is voluminous (see Hermalin and Weisbach, 2003, and Adams et al., 2010, for surveys).

This section has briefly considered the various threats investors face and the range of responses they may take (more precisely, the economics literature that considers how to respond). What it has ignored is what determines the choice and degree of response. In particular, because those in control naturally do not wish to be constrained, they can be expected to push back against the various responses, seeking to modify or limit them. This is the topic of the next section.

18 This can be seen intuitively by imagining the estimate of ability were based on a signal that is complete noise. Consequently, the posterior estimate would equal the prior estimate. Given there is no change in the estimate, there is no uncertainty and, hence, no risk. In contrast, if the signal is informative (not noisy), then the posterior estimate will put considerable weight on it, meaning, from an *ex ante* perspective, that the posterior estimate is quite variable, which translates into risk for the CEO.
4 The Determination of Responses: Corporate Governance in an Equilibrium Framework

As noted, investors wish to limit the ability of controlling parties to behave contrary to the investors’ interests. Such limits are not in the controlling parties’ private interests and the controlling parties can, thus, be expected to resist their imposition. Or, if they cannot block their imposition, then controlling parties will seek compensation for them. This suggests that the corporate governance institutions we observe are the consequence of some—perhaps implicit—bargaining game between investors and controlling parties.

In economics, the standard presumption is that if the private parties are symmetrically informed at the time they bargain, they will typically reach an efficient outcome. Because the optimal solutions to agency problems are almost always second best, the efficiency standard is second-best efficiency; a standard often referred to as constrained Pareto optimality. Specifically, given the constraints inherent in dealing with agency issues, the outcome is optimal only in the sense that money has not been left “on the table.” That is, there is no alternative contract or institution the parties could adopt that at least one party favors more than the original and the other no less than the original. Moreover, if we assume—as seems reasonable in this context—an ability to make transfers between the parties, then there is no alternative contract or institution that would generate greater total welfare in equilibrium than the one the parties put in place.

Admittedly, the view that governance institutions could be constrained Pareto optimal is not without controversy. It is a perspective that is certainly at odds with how many commentators on corporate governance see corporate governance. How can one use the adjective “optimal” with regard to practices—such as apparently impotent boards and apparently undeservedly huge executive compensation—that seem so suboptimal? Where is the optimality in spectacular governance failures such as Enron, Worldcom, and Parmalat? These commentators are, I would argue, overlooking one or more the following:

1. Even if there were no constraints, this does not imply that investors would get everything they want.

As noted, there is bargaining between investors and management, which presumably reflects relative

\[19\] See Hermalin et al. (2007), especially §§2.2 and 2.3, for a survey of the relevant literature.

\[20\] As examples of work suggesting current governance is highly suboptimal see Bebchuk and Fried (2003, 2004), Lorsch and MacIver (1989), and MacAvoy and Millstein (2003).
market position. If capital is plentiful, but managerial talent scarce, then the returns generated by combining capital and talent will tend to flow disproportionately to the talent.\(^{21}\)

2. Even if there were no constraints, there are costs to governance and some amount of failure should occur even in the first-best solution. By way of analogy, auto-related fatalities can be reduced by requiring automobiles to travel more slowly. But there is an opportunity cost to the extra time spent traveling. Hence, society accepts some failures (auto-related fatalities) because eliminating or reducing them would be too costly.

3. There are constraints and they are not trivial. As an analogy, we would not say that Henry Aaron, Babe Ruth, and other Hall-of-Fame hitters were failures because over 65% of their at-bats ended in failure.\(^{22}\) Major-league pitchers, backed up by major-league fielders, represent a serious constraint with respect to hitting safely.

4. Finally, the corporation is not the only way to organize an enterprise. Alternatives include sole proprietorships, partnerships, cooperatives, and state-owned enterprises.\(^{23}\) For over 400 years, at least, the corporate form has successfully competed with these other forms and has, economically, come to dominate all alternatives.\(^{24}\) Hence, there is a limit on just how bad corporate governance can be.\(^{25}\)

\(^{21}\)In this regard, it is worth speculating about the extent to which the run-up in executive salaries at the beginning of this century were due to the availability of cheap capital. Did the relative bargaining positions of talent and capital tip towards talent?

\(^{22}\)Henry Aaron’s career average was .305. Babe Ruth’s .342.

\(^{23}\)See Hansmann (in this volume) for a survey of the literature on why different organizational forms are chosen.

\(^{24}\)The East India Company is often taken to be the first corporation, chartered on December 31, 1600 by Queen Elizabeth I (Baskin and Miranti, 1997). Another early corporation was the Vereenigde Oost-Indische Compagnie (usually referred to in English as the Dutch East India Company), which was established in 1602. Neither is as old as Stora Kopparberg, a Swedish mining company that sold its first share in 1288 and was chartered by King Magnus IV of Sweden in 1347. Malmendier (in press, p. 1) argues “the earliest predecessor of the modern business corporation was ... the Roman societas publicanorum.” Societes publicanorum rose to prominence in the last two centuries BCE.

\(^{25}\)This is not to claim that corporate governance is better than the governance of other forms of enterprise. Rather it is the observation that it cannot be so bad as to outweigh whatever benefits the corporate form offers over these other forms. In this regard, it is worth noting that one of the benefits of the modern corporate form, limited liability, cannot be what has made the corporate form survive: Limited liability did not come into being until the British
I am not, however, suggesting one adopt a Panglossian approach to governance. It is doubtful that we live in the best of all possible worlds; that is, there could be scope for improvement via regulatory action (more on this later). My point is rather that the efficacy of regulatory action is more limited than many commentators would suggest.

Investors and management *choose* the institutions by which their relation is governed. Hence, governance institutions are *endogenous*. As such, any regression analysis that attempts to determine the effect of one or more of these institutions on firm performance—for instance, a regression of return on assets ($\text{ROA}$) on the proportion of the board of directors who are outside directors—needs to control for the joint endogeneity of the dependent and independent variables. Although researchers have not always been careful in this regard, the better work does attempt to deal with this endogeneity (see Bhagat and Jefferis, 2002, for a more complete treatment of this issue).

A more fundamental issue though is the following: How should one interpret a regression of financial performance on governance attributes? A common presumption in the literature, for example, is that outside directors are better monitors of management than inside directors. Naïvely, one might seek to test this hypothesis by regressing a measure of financial performance on a measure of the relative number of outside directors. The “rationale” being that, if the hypothesis is correct, firms that score higher on the measure of outside directors will tend to score better in terms of financial performance. That is, if the

26The outside directors of a corporation are those directors who are not current or former officers of the corporation (i.e., part of its management team). Inside directors are directors who are current or former officers. In the literature the two groups are often referred to as outsiders and insiders, respectively. It is generally assumed that outsiders’ interests are more aligned with investors than are insiders; in particular, the presumption is that outsiders exercise some control over management. See, e.g., Hermalin and Weisbach (2003) and Adams et al. (2010) for a discussion of the relevant literature.

27Baysinger and Butler (1985), Brickley et al. (1994), Rosenstein and Wyatt (1990), and Durschin et al. (2009), among others, find evidence that measures of outside directors are positively correlated with firm performance. Hermalin and Weisbach (1991) and Yermack (1996) also study the issue, but find little evidence that firm performance is related to board composition. Other studies have investigated the relation between other attributes of the governance structure and firm performance: Yermack (1996) and Eisenberg et al. (1998) find a negative relation between board size and firm performance (Coles et al., 2008, however, find that firm performance is increasing in board size for certain types of firms); Stulz (1988) and Morck et al. (1988) find a positive relation between managerial stock ownership and firm performance (up to a certain threshold of ownership, at which point the relation reverses; see §6 infr and Becht et al., 2003, p. 63, for a more detailed appraisal of the managerial stock ownership literature); Jarrell and
data resemble those in Figure 1, then there is support for the hypothesis.\footnote{To be more precise, imagine in Figure 1 that the dependent variable is the residual of financial performance regressed on the relevant controls other than the measure of outsiders.}

A problem with such an interpretation is that if more outside directors really mean better financial performance, why does firm A in Figure 1 choose to have a low measure of outside directors? That is, if the hypothesis is true, then it would appear firm A is leaving money on the table. Surely the investors and management of firm A would have had a larger “pie” to divide had they emulated firm B in terms of outside directors. In short, one cannot accept both the hypothesis that governance structures are constrained Pareto optima and the hypothesis that the regression in Figure 1 is a valid test of whether more outside directors improves financial performance.

Furthermore, as Demsetz and Lehn (1985) and others have noted, if one accepts that the observed governance structures are constrained Pareto optima, then it is no longer clear that any relation should

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{A hypothetical regression of financial performance against a measure of outside director presence on the board.}
\end{figure}
exist in the data between governance structures and performance. If the relevant parties are solving an optimization program unique to their firm, then, as Demsetz and Lehn observe, we could readily expect to see data like that shown in Figure 2. In this second figure, in addition to the scatter plot of the “data,” I have also drawn parabolas to indicate, heuristically, the different optimization programs faced by the investors and controlling interests of these three firms.29

Consistent with the idea that the parties achieve constrained Pareto optima, each firm in Figure 2 is shown resting on the top of its respective hill; that is, no set of parties could do better. Clearly, a regression should reveal no relationship between the variables in question. Yet, many carefully executed empirical studies find evidence of correlation; moreover, this correlation tends to be positive, measures that

---

29See Adams et al. (2010) for a survey of models that seek to explain the optimal mix of outside and inside directors. Within this literature, there are three basic explanations: (i) bargaining between management and investors (directors) results in a compromise (see, e.g., Hermalin and Weisbach, 1998, and Cyert et al., 2002); (ii) getting management (the CEO) to reveal information or otherwise take actions risky to their careers (see, e.g., Dominguez-Martinez et al., in press, and Almazan and Suarez, 2003); or (iii) as a means of inducing efficient information sharing among directors (see, e.g., Adams and Ferreira, 2007, and Harris and Raviv, in press). In all these models, restricting the proportion of outside directors is essentially a means of committing the board to future actions that cannot be contracted on formally.
arguably indicate stronger governance (e.g., smaller boards, higher governance indices, etc.) go hand-in-hand with better firm performance and vice versa. Why, in a world of constrained Pareto optimal governance structures, do we observe these positive correlations?

Could the view that Figure 1 is showing something causal be resuscitated by assuming that firms sometimes get out of equilibrium or that there are large adjustment costs? The problem with an out-of-equilibrium explanation is the following: If firms are truly seeking to be at the peaks of their respective parabolas, but get knocked off the peak for various reasons, why aren’t they equally likely to land to the right of the peak as to the left? That is, an out-of-equilibrium explanation of the data in Figures 1 must posit that firms are more likely to slip to the left—have less than the optimal measure of outside directors—than to the right—have more than the optimal measure. To be sure, such a bias is possible. Indeed, the Hermalin and Weisbach (1998) model predicts that corporations will “drift to the left” over the tenure of a CEO. The Hermalin and Weisbach model does not, however, also predict that financial performance will fall as the corporation drifts to the left.

Adjustment costs would seem to offer another explanation for why firm A does not emulate firm B in Figure 1. But then one is left with the puzzle of how firm A came to be so far from the optimum board structure in the first place. Perhaps L was, at some point in the past, the optimal measure of outside directors. But if so, then H was suboptimal and we are left wondering how firm B came to be so far from the optimum. An evolutionary twist might help: A greater percentage of outsiders on the board is a “mutation” largely enjoyed by new firms, such as B. Older firms, such as A, see that this is a positive mutation, but find it hard to adjust. Evolution and organizational adjustment are important phenomena (see, e.g., my chapter on corporate culture in this volume), but it seems unlikely this could be the explanation given the long time

---

30 See, e.g., Bhagat and Jefferis (2002), Becht et al. (2003), and Adams et al. (2010), among others, for surveys covering this empirical literature. Examples of articles finding a negative correlation between board size and performance (i.e., that small boards are associated with better performance) include Yermack (1996) and Eisenberg et al. (1998) (but see, also, Coles et al., 2008 for partially contradictory evidence). Examples of articles finding a positive correlation between indices of good governance and performance include Gompers et al. (2003) and Bebchuk et al. (2009).

31 There is also empirical evidence for a leftward drift; that is, for the proportion of outside directors to tend to decrease with the length of the CEO’s tenure. See Hermalin and Weisbach (1988), Baker and Gompers (2003), Ryan and Wiggins (2004), and Boone et al. (2007).
such heterogeneity has existed (e.g., Hermalin and Weisbach, 1988, find considerable variation as far back as the 1970s and such variation continues to exist to this day).

Where, then, does this leave us? The fact that many studies do find a significant and positive relation between governance and performance suggests there could be something in the data that cannot be dismissed. But what is in the data and how do we reconcile it with the previous discussion? One answer is that we’re thinking about the data backwards: Rather than good governance causing good performance, perhaps it is good performance—more specifically the potential for or likelihood of good performance—that is causing good governance.

To have a tangible framework within which to explore the idea that good governance is the consequence of the potential for good performance, consider the following model. Investors deal with a single manager. The manager’s utility is

\[ u = D + v(R - D, g), \]

where \( R \) denotes the firm’s resources, \( D \) denotes the amount the manager diverts to uses he desires, but which are unproductive from the firm’s perspective, \( v : \mathbb{R}^2 \rightarrow \mathbb{R} \), and \( g \) is a measure of the strength or effectiveness of governance.\(^{32}\) The variable \( g \) could represent the percentage of independent directors on the board or on key board committees, a measure of the directors’ diligence, a measure of the effectiveness of the monitoring and auditing systems in place, some measure of the strength of the incentives given the manager, or perhaps some index of governance strength such as the proposed by Gompers et al. (2003). One interpretation, in particular, is worth considering: given the many dimensions of governance, think of \( g \) as the firm’s total expenditure on governance. Provided the owners set the dimensions of governance optimally, spending more on governance must correspond to stronger governance.

\(^{32}\)A comment made about this model is that \( v \) should vary with firm type: “Independence from firm type means that a manager who delivers $9 million of profit will receive the same incentive pay (or threat of being fired) regardless of whether the firm had the potential to deliver $10 million or $100 million.” This objection overlooks the fact that \( g \) is endogenous. As will be shown, the manager of the firm with the greater potential will operate under stricter governance in equilibrium, so his reward for delivering $9 million would be far less than if he were managing the firm with lower potential.

The conclusions of the model would not change if the manager’s utility were \( b(D) + v(R - D, g) \), where \( b(\cdot) \) is an increasing and, at least weakly, concave function. If there is no risk, then the model is robust to any positive monotonic transformation of \( D + v(R - D, g) \).
As with all exogenous functions in this chapter, \( v(\cdot, \cdot) \) is assumed to be twice continuously differentiable. The function represents the benefit the manager derives from behaving in a manner desired by the investors (i.e., not diverting funds or assets for private use). This benefit is, in part, a function of the level of governance.\(^{33}\) All this simply reflects the idea that governance structures operate to reward the manager for good behavior. Consistent with the view that better performance is better rewarded, assume \( v_1(\cdot, g) > 0 \) for all \( g > 0 \).\(^{34}\)

To keep the analysis straightforward, assume that, for any \( g \), there is a unique value of \( D \) that maximizes the manager’s utility. That is, that solves

\[
\max_D D + v(R - D, g).
\]

Observe this maximization has unique solution if the manager’s utility exhibits diminishing marginal return to diversion (i.e., if \( v_{11}(\cdot, g) < 0 \) for all \( g \)). Let \( D(g) \) denote that solution. Again to keep the analysis straightforward, corner solutions are ruled out.\(^{35}\)

By stronger governance, one means an increase in the governance parameter that results in a reduction in agency behavior; that is, leads the manager to choose a smaller \( D \). Formally, assume:

**Assumption 1** Let \( g \) and \( g' \) be two different levels of governance, \( g > g' \) and let \( D \) and \( D' \), respectively, be the levels of private use chosen by the manager in response to those governance levels. Then \( D < D' \).

A sufficient condition for Assumption 1 to hold is that the marginal utility of not diverting resources is increasing in the level of governance; that is, \( v_{12}(\cdot, \cdot) > 0 \). That more governance then leads to better behavior follows from well-known concepts in comparative statics (see, e.g., Topkis, 1978, or Milgrom and Roberts, 1990), which in this context can be summarized by

\(^{33}\)Of course, because a cost is just a negative benefit, this formulation also incorporates specifications in which the governance structure punishes the manager for behaving at odds with the shareholders’ preferences.

\(^{34}\)I am employing the convention that footnotes on functions denote partial derivatives; that is, \( f_n \) is used to denote the derivative with respect to the \( n \)th argument of function \( f \) and \( f_{nm} \) to denote the second derivative with respect to the \( n \)th and \( m \)th arguments.

\(^{35}\)Specifically, if \( g > 0 \) we want \( D(g) < R \). Hence, assume \( v_1(0, g) > 1 \) for all positive \( g \). However, to avoid dealing with corner solutions in the level of governance, assume \( v_1(0, 0) = 1 \); hence, the absence of governance implies \( D(0) = R \). To assure \( D(g) > 0 \) for \( R \) great enough assume \( \lim_{R \to \infty} v_1(R, g) < 1 \) for all \( g \). An example of a \( v \) function satisfying all the assumptions given so far is \( v(R - D, g) = 2g\sqrt{R - D} \).
Lemma 1 Let \( f : \mathbb{R}^2 \to \mathbb{R} \) be at least twice differentiable in its arguments. Suppose that \( f_{12} (\cdot, \cdot) > 0 \). Let \( \hat{x} \) maximize \( f(x, z) \) and let \( \hat{x}' \) maximize \( f(x, z') \), where \( z > z' \). Then \( \hat{x} \geq \hat{x}' \). Moreover, if \( \hat{x}' \) is an interior maximum, then \( \hat{x} > \hat{x}' \).

Suppose that the corporation’s returns, \( r \), are distributed on the interval \( [r, \infty), r > -\infty \), according to the conditional distribution function \( F(\cdot| R - D, \tau) \), where \( \tau \in \mathbb{R} \) is the corporation’s type (more on \( \tau \) below). Assume \( \mathbb{E}\{r| R - D, \tau\} \) exists for all possible values of \( R - D \) and \( \tau \). Via integration by parts, observe
\[
\mathbb{E}\{r| R - D, \tau\} = r + \int_r^{\infty} (1 - F(r| R - D, \tau)) dr = r + \int_r^{\infty} S(r| R - D, \tau) dr ,
\]
where \( S(r| X, \tau) \equiv 1 - F(r| X, \tau) \) is the survival function. It is natural to assume that the more net resources utilized, the greater the expected return. In fact, let’s make the somewhat stronger assumption that an increase in net resources improves the distribution of returns in the sense of strict first-order stochastic dominance; that is, \( \partial S(r| X, \tau) / \partial X > 0 \) for all \( X, r \in (r, \infty) \), and \( \tau \). Finally, as our definition of corporate type, assume that this improvement is stronger the greater is \( \tau \). In other words, the marginal expected return from an increase in net resources utilized is greater for higher-type corporations than lower-type corporations. In terms of calculus, this definition of type can be written
\[
\frac{\partial^2 S(r| X, \tau)}{\partial \tau \partial X} > 0 \tag{1}
\]
for all \( X, r \in (r, \infty) \), and \( \tau \).

The realized profit of the corporation is return less the cost of governance, \( C(g) \). Writing \( X(g) \) for \( R - D(g) \), the investors’ choice of governance will be the solution to the program
\[
\max_g \left( r + \int_r^{\infty} S(r| X(g), \tau) dr - C(g) \right) . \tag{2}
\]
The cross-partial derivative of (2) with respect to \( g \) and \( \tau \) is
\[
\int_r^{\infty} \frac{\partial^2 S(r| X(g), \tau)}{\partial X \partial \tau} X'(g) dr > 0 , \tag{3}
\]
where the inequality follows from (1) and because \( D(\cdot) \) is a decreasing function and, thus, \( X(\cdot) \) an increasing function. Expression (3) and Lemma 1 imply
Proposition 1 *The level of governance a corporation has is non-decreasing in its type (i.e., in its marginal expected return from net resources).*

By imposing more structure on the model, it is possible to ensure that the solution to (2) is always an interior solution. This would allow us to conclude:

**Corollary 1** *If the level of governance that maximizes the investors’ expected net returns is an interior solution within the space of feasible governance levels, then the level of governance a corporation has is increasing in its type.*

Proposition 1 and its corollary explain the difference between firms A and B in the figures above: An additional dollar of net resources is more valuable to firm B than firm A; that is, B is a higher-type firm than A. What hasn’t yet been explained is why the data resemble Figure 1 and not Figure 2. Explaining that requires an additional assumption: One presumes that a corporation that employs no net resources will enjoy a zero return (one typically doesn’t get something for nothing). Hence, the distribution of returns when no net resources are employed is independent of type. In terms of calculus, this means \( \frac{\partial S(r|0, \tau)}{\partial \tau} \equiv 0 \) for all \( r \) and \( \tau \). This and expression (1) imply

\[
\frac{\partial S(r|X, \tau)}{\partial \tau} = \frac{\partial S(r|X, \tau)}{\partial \tau} - \frac{\partial S(r|0, \tau)}{\partial \tau} = \int_0^X \frac{\partial^2 S(r|x, \tau)}{\partial \tau \partial X} \, dx > 0
\]

for all \( r \) and \( X > 0 \). Observe that (4) demonstrates that an increase in type, holding net resources constant, improves the distribution of returns in the sense of first-order stochastic dominance.

Let \( g(\tau) \) be the solution to program (2). Utilizing the envelope theorem, it follows that

\[
\frac{d}{d\tau} \left( \mathbb{E} + \int_0^\infty S(r|X(g(\tau)), \tau) \, dr - C(g(\tau)) \right) = \int_0^\infty \frac{\partial S(r|X(g(\tau)), \tau)}{\partial \tau} \, dr > 0,
\]

where the inequality follows from (4). In words, higher-type firms have greater expected profits in equilibrium than do lower-type firms. This explains why the performance of firm B, the higher-type firm, is better than that of firm A, the lower-type firm. To summarize, I’ve shown that higher-type firms will have both greater levels of governance and greater profit (in expectation) than lower-type firms. This means that governance level and profits should be positively correlated; the data should, in fact, resemble those shown in Figure 1.
Proposition 2  In this model, in which all corporations are making optimal decisions, there will be a positive correlation between level of governance and corporate profits.

Note the path of causation: A corporation with a high marginal return to net resources—which will therefore be, ceteris paribus, a corporation with greater profits on average—is a corporation with a higher marginal cost of agency. It therefore puts in place a higher level of governance than a corporation with a low marginal return to net resources (low marginal cost of agency).

5  The Many Dimensions of Governance

Governance is the product of decisions on many dimensions: board structure, security design, incentive schemes, etc. This observation raises another issue: If the different dimensions are substitutes and different firms face different factor-price vectors, then there could be considerable heterogeneity in the governance practices within any sample of firms. To see this readily illustrated, consider the following extreme “model.” Let \( g = \max\{\gamma_1, \gamma_2\} \), where \( \gamma_i \) is the amount of governance on dimension \( i \). A corporation will, therefore, set the more expensive dimension to zero. If the more expensive dimension is the first for some firms, but the second for others, then the sample will contain a subset of firms that use only the first method of governance and another subset that use only the second. This can create complications for empirical work because empirical work often focuses on the effect of governance on a single dimension only. Because the firms that elect to have zero of the dimension under study are not necessarily poorly governed firms (they could have invested heavily in the other, unstudied dimension), the coefficient on the dimension being studied will be biased downward. Figure 3 illustrates the potential problem. In that figure, there are four data points. Firms A and B achieve their governance objectives by investing along the first dimension (e.g., proportion of outside directors), while Firms C and D achieve their governance objectives by investing along the second dimension (e.g., incentive compensation). Firms B and D are high-type firms, so perform well, while firms A and C are low-type firms, so perform less well. But because firms C and D achieve their governance objectives differently than do A and B, a regression of performance on \( \gamma_1 \) will suggest that the relation between these variables is less strong than it truly is for those firms that invest in governance along the first dimension.
Figure 3: Empirical Consequence of Governance Structures Having Many Dimensions.

Observe that controlling for governance on both dimensions—that is, running the regression

$$\text{Performance} = \beta_0 + \beta_1 \gamma_1 + \beta_2 \gamma_2 + \varepsilon$$

—does not correct the problem identified in Figure 3. Relative to the true relation for those firms utilizing a given dimension, the coefficients $\beta_1$ and $\beta_2$ will be biased downward.

Heterogeneity is a particularly important issue in the study of pay-for-performance incentive contracting. Agency theory (e.g., Holmstrom, 1979; Shavell, 1979; and Grossman and Hart, 1983) teaches us that one way to better align the incentives of investors (the principal) and managers (the agent) is to make the agent’s compensation contingent on signals that are informative about the agent’s action in a way that induces the agent to take actions the principal desires. Because investors desire that management take actions that boost their returns, a natural set of signals upon which to base compensation are various performance metrics (e.g., profit, ROA, etc.). The extent to which different metrics are utilized depends on their informativeness (see, e.g., Holmstrom, 1982, and Kim, 1995), which could differ across different firms or industries. Consequently, one should expect variation in the incentive contracts employed across firms and industries. One should also expect variation because high-type firms will wish to provide their managers...
stronger incentives than low-type firms, as discussed above. Even with otherwise identical firms, variation in incentives can arise in equilibrium as a consequence of product-market competition (Hermalin, 1994). More generally, the discipline of product-market competition can be a substitute for incentive compensation, thus serving as another source of heterogeneity. These theoretical predictions of heterogeneity in incentive compensation are given empirically validity by Kole (1997), who examined executive compensation contracts and documents considerable variation across them.

What is the implication for this variety for empirical work? Consider a possible regression specification:

\[ \text{Comp}_n = \beta_0 + \beta_1 \times \text{Perf}_{1n} + \cdots + \beta_J \times \text{Perf}_{Jn} + \epsilon_n, \]  

(5)

where \( \text{Comp}_n \) is the compensation of the \( n \)th executive (typically firm \( n \)'s CEO) and \( \text{Perf}_{jn} \) is the \( j \)th performance metric (e.g., ROA) for the \( n \)th executive’s firm. What is the interpretation of (5)? Presumably, the coefficients (the \( \beta \)s) are the terms of the compensation contract. But which firm’s? If, as theory and evidence indicate, different firms will utilize different contracts, there is a fundamental misspecification in (5), because (5) rests on the erroneous assumption that all firms are using the same compensation contract.

The consequences of this misspecification are similar to the problem identified in Figure 3. If firms A and B tend to base compensation on \( \text{Perf}_1 \), while C and D tend to base it on \( \text{Perf}_2 \), then the situation will resemble that shown in Figure 4. The true sensitivity of pay to performance for those firms that base

\[ \text{Comp}_n = \beta_0 + \beta_1 \times \text{Perf}_{1n} + \cdots + \beta_J \times \text{Perf}_{Jn} + \epsilon_n, \]  

(5)

where \( \text{Comp}_n \) is the compensation of the \( n \)th executive (typically firm \( n \)'s CEO) and \( \text{Perf}_{jn} \) is the \( j \)th performance metric (e.g., ROA) for the \( n \)th executive’s firm. What is the interpretation of (5)? Presumably, the coefficients (the \( \beta \)s) are the terms of the compensation contract. But which firm’s? If, as theory and evidence indicate, different firms will utilize different contracts, there is a fundamental misspecification in (5), because (5) rests on the erroneous assumption that all firms are using the same compensation contract.

The consequences of this misspecification are similar to the problem identified in Figure 3. If firms A and B tend to base compensation on \( \text{Perf}_1 \), while C and D tend to base it on \( \text{Perf}_2 \), then the situation will resemble that shown in Figure 4. The true sensitivity of pay to performance for those firms that base

36 The idea that market competition serves to discipline managers is sometimes attributed to Hicks (1935) as the “quiet life” hypothesis. Formalizing it has not proved easy (see Allen and Gale, 2000, for a discussion of the intellectual history of the quiet-life hypothesis). Part of the problem, as noted by Hermalin (1992), is that product-market competition introduces a number of contradictory effects. Important articles in this area include Hart (1983), Scharfstein (1988), Schmidt (1997), and Allen and Gale. Empirical studies include Bertrand and Mullainathan (2003) and Giroud and Mueller (2008). That latter, in particular, finds evidence that product-market competition mitigates managerial slack. A related literature (see, e.g., Aghion et al., 2005, and cites therein) considers the complementary question of how competition affects incentives to innovate and become more productive.

37 An actual regression of this sort could also have numerous non-performance controls (e.g., for industry, firm size, etc.). They’ve been omitted from expression (5) to avoid unnecessary clutter. In addition, the dependent variable is often the log of compensation; taking logs would not change the point being made here. Finally, what is compensation and how it is best constructed is a tricky empirical issue (see, e.g., Jensen and Murphy, 1990; Hall and Lieberman, 1998; or Hermalin and Wallace, 2001). How compensation is measured is, however, not relevant for the point being made here.

38 Except under strong conditions (see Holmstrom and Milgrom, 1987), a linear compensation scheme is not guaranteed to be the optimal scheme. Indeed, it is straightforward to construct examples in which linear compensation would be sub-optimal. Given that real-life compensation is not linear (bonuses, e.g., are rarely—if ever—negative and stock-options pay off nonlinearly), the use of a linear specification is, in itself, questionable.
compensation on Perf is the slope of the “true” line. The econometrician will, however, measure it as the slope of the estimated line. Consequently, the estimates of all coefficients are biased downward insofar as they understate the true sensitivity of pay to performance measures for those firms that base compensation on those performance measures.

Hermalin and Wallace (2001) attempt to deal with this issue by allowing the coefficients in (5) to vary across firms (such an analysis requires, obviously, panel data). Consistent with the concerns raised above and in Figure 4, they show, using a two-step estimator of a hierarchical linear model based on Wong and Mason (1991), that the estimated coefficients are greater under such a specification than using OLS. In other words, as predicted, failure to treat the coefficients—that is, the parameters of the compensation contract—as varying across firms leads to a downward bias in their estimation. In addition, consistent with the idea that the different dimensions of governance could be substitutes and that firms vary with respect to

---

39 True is in quotes because there are multiple true relations—in this case one for firms A and B and another for firms C and D.

40 Because they estimate a distribution of coefficients, a more precise summary of their results is that the distribution estimated via their specification lies to the right (has a greater mean) than under OLS.
the performance measures to which they make compensation sensitive, Hermalin and Wallace find a negative correlation between the coefficients on the two performance measures they investigate: Firms that make CEO pay highly sensitive to one measure tend, all else equal, not to make it particularly sensitive to the other and vice versa.

Even if the different dimensions of governance are complements—so a high-type firm will tend to have more governance on all dimensions than a low-type firm—expression (5) is still misspecified. To see this, suppose there is only one performance measure, \( r \), which is distributed normally with mean \( a \tau \) and variance \( \sigma^2 \), where \( a \in \mathbb{R}_+ \) is the CEO's action and \( \tau \) again denotes the firm's type. Suppose the CEO's utility is \(-\exp\left(\frac{1}{2}a^2 - y\right)\), where \( y \) is his income. Limit attention to affine incentive contracts: \( y = \beta_0 + \beta_1 r \). Hence, the CEO's equilibrium expected utility is a monotonic transformation of

\[
\beta_0 + \frac{1}{2}\beta_1^2 (\tau^2 - \sigma^2).
\]

Suppose, perhaps unrealistically, that investors have all the bargaining power and are, thus, able to hold the CEO to his reservation utility. If \( u \) is his, suitably transformed, reservation utility, then, from (6), it must be that \( \beta_0 = u - \beta_1^2 (\tau^2 - \sigma^2)/2 \). Equilibrium expected compensation and firm expected profit are thus, respectively,

\[
\frac{1}{2}\beta_0^2 (\tau^2 + \sigma^2) \quad \text{and} \quad \beta_1 \tau^2 - \frac{1}{2}\beta_1^2 (\tau^2 + \sigma^2) - u.
\]

Investors will set \( \beta_1 \) to maximize the latter expression; hence, \( \beta_1 = \tau^2 (\tau^2 + \sigma^2)^{-1} \). Observe \( \beta_1 \) is increasing in \( \tau \); that is, consistent with the discussion above, higher-type firms employ stronger governance than do lower-type firms. Moreover, as claimed, different firms are using different contracts—the implicit assumption in (5) of a common compensation contract is invalid.

The adverse consequences of this misspecification in (5) can be illustrated via the following simulation. Set \( u = 0 \) and \( \sigma^2 = 1 \). Construct a “sample” of 10,001 firms where \( \tau_1 = 1 \) and \( \tau_n = \tau_{n-1} + 1/10,000 \), \( 1 < n \leq 10,001 \). For each firm randomly generate \( r \) as a normal random variable with mean \( a_n \tau_n \) or,

\[\text{The CEO's expected utility is readily seen to be a monotonic transformation of} \]

\[
\beta_0 + \beta_1 a \tau - \frac{1}{2}\beta_1^2 a^2 - \frac{1}{2}a^2.
\]

Hence, the CEO maximizes his expected utility by choosing \( a = \beta_1 \tau \).
substituting, mean $\tau_n^4 (\tau_n^2 + \sigma^2)^{-1}$. Calculate the corresponding compensation for each firm and then estimate (5) via OLS from these data.\textsuperscript{42} Multiple runs of this simulation all came back with estimates of $\beta_1$ between .575 and .58. The true $\beta_1$s range from .5 to .8, with 82% of them being greater than .58; that is, the estimated sensitivity of pay to performance from (5) dramatically understates the actual sensitivity in the sample.\textsuperscript{43}

This insight is more than just nit-picking. The question one is investigating when estimating the sensitivity of pay to performance is not whether there is a positive relation—it is virtually inconceivable that they are negatively correlated in reality—but rather how sensitive? Do executives operate under high-powered incentives or not? Jensen and Murphy (1990), who estimate an equation similar to (5), argue that CEOs operate under fairly low-powered incentives; they estimate that CEO wealth changes by $3.25 per $1000 increase in shareholder wealth. They cite earlier studies that found a pay-to-performance elasticity of approximately 0.1, a number consistent with their own estimates. They argue that such estimates of sensitivity are “inconsistent with the implications of formal agency models of optimal contracting” (p. 227). They go on to suggest that outside political forces “implicitly regulate executive compensation by constraining the type of contracts that can be written between management and shareholders” (p. 227).\textsuperscript{44} Haubrich (1994) disputes this conclusion. He conducts numerical simulations using parameterized versions of the Grossman and Hart (1983) model, calibrated using real-life data, to show that the Jensen and Murphy estimates need not be at odds with optimal contracting.\textsuperscript{45} Hall and Liebman (1998) also estimate an equation similar to (5) and find a greater sensitivity of pay to performance than did Jensen and Murphy. They ascribe the difference in findings between the two studies to their focus on a later period of time, during which CEOs held more stock and stock options than before. Bebchuk and Fried (2003, 2004) also argue that the pay-to-performance sensitivity is too low, but do so from the perspective that managers have too much power and are being unduly rewarded, almost independently of their performance. Hall and Murphy (2003) take issue with that view, but they too suggest that firms may fail to be setting the sensitivity of pay to performance

\textsuperscript{42}The Mathematica program used for this is available from the author upon request.

\textsuperscript{43}The true average of the $\beta_1$s is .678.

\textsuperscript{44}Bebchuk and Fried (2004) refer to these as “outrage” costs; although, in contrast to Jensen and Murphy, they believe these are not sufficiently deterring high levels of compensation.

\textsuperscript{45}See also Haubrich and Popova (1998) for another calibration exercise that reaches a similar conclusion.
optimally (they suggest that the costs of stock options are misperceived as being “free”—see, in particular, pp. 66–69). What the discussion here suggests, however, is that the debate over executive compensation is being conducted under a possibly misleading understanding of what the data actually say. If there are systematic biases of the sort identified above, then the sensitivity of pay to performance could be higher than heretofore estimated (at least for those corporations that rely on that means of governance).

This section has sought to make the following points. The inherent heterogeneity in governance solutions—both in structure and intensity—is a first-order problem for the empirical study of governance (as well as a source of puzzles for theoreticians). Cross-sectional regressions, in particular, must be viewed with great caution as they are subject to misinterpretation.

6 Dispersed Shareholders, Managerial Power, and Governance

Many of the complaints about governance are that managers have too much power vis-à-vis the shareholders (see, e.g., Lorsch and MacIver, 1989, or Bebchuk and Fried, 2004). Since at least Berle and Means (1932), this disparity in power has been attributed to the dispersion of ownership. In particular, any one shareholder holds too few shares to have any power him or herself; that is, alone, he or she does not possess the votes to control the board of directors and, through it, management. A lone shareholder could, in theory, seek to build a coalition of shareholders with sufficient power to control the board of directors, but—traditionally at least—such action is costly. Because our lone shareholder bears the cost, but shares in the benefit, the usual theory of positive externalities applies: any one shareholder will undersupply efforts to rein in management. Not surprisingly, attempts to gain control of the board of directors by existing shareholders are relatively rare events.

46For more complete surveys of the debate over executive pay, especially whether it is appropriately sensitive to performance, see Devers et al. (2007) and Edmans and Gabaix (2009).

47For large corporations (e.g., Fortune 500), reliance on incentive compensation (e.g., use of bonuses, stock options, etc.) has tended to be the norm. For smaller corporations, however, this has been—historically at least—less universally true. See, e.g., Hermalin and Wallace (2001) for data on the savings-and-loan industry. See also Hall and Murphy (2003) for a discussion on the growing use of stock-option compensation during the 1980s and ’90s.

48Over a 16-year period, Mulherin and Poulsen (1998) find there were only 270 proxy contests (attempts to change the board of directors) at NYSE and Nasdaq-listed firms (an average of 16.875 contests per year). Of these, roughly half led to changes in management. This rate of “success” clearly overstates the ex ante probability of success, since
As Grossman and Hart (1980) observed, a similar problem confronts an outside actor who attempts to take over a corporation. If the motivation for such a takeover is the outside actor’s belief that it can generate greater returns from the assets of the corporation, then it will behoove the outside actor to launch a takeover bid provided it will get a sufficient share of the increase in returns it generates. The problem is that an existing shareholder will reason as follows: “Because I own such a small proportion of shares, the odds that my tendering my shares will be pivotal to the success of the takeover bid are so small as to be negligible. Hence, if I retain my shares, I will enjoy the benefits of the improved management. I, therefore, should not sell unless the bid price equals the expected present value of the returns I will enjoy under this improved management.” Given this reasoning, existing shareholders will sell only if given all the gains created by the outside actor. This, however, eliminates the gains from launching a takeover bid in the first place. Of course, takeovers do occur, so there must be ways around this free-riding problem. Grossman and Hart (1980) suggest some means of reducing the free-riding problem by limiting the returns to those shareholders who hold out (the ex post minority).

Such free-riding problems are less pronounced if there is a shareholder whose holdings are relatively large. The governance literature has, thus, devoted considerable attention to large shareholders (blockholders). The classic article in this regard is Shleifer and Vishny (1986), which considers how the presence of a large blockholder can serve to overcome the free-riding problems identified above. The presence of a large blockholder is not, however, unambiguously good for governance. A large blockholder can use its power for evil as well as good: In a reverse of the free-riding problem, if the blockholder can use its power to engage in self-dealing with the corporation (e.g., induce the corporation to buy inputs at inflated prices from a supplier it fully owns), then the blockholder gets all the benefit, but incurs only a portion of the cost. This problem is potentially more severe when there is a pyramid-structure of ownership, which permits a shareholder to control the votes, but have a more limited financial exposure. Such ownership structures are more common presumably such contents are not undertaken unless there is a reasonable chance of success.

49 See Hermalin and Schwartz (1996) for a discussion of the relevant law concerning freeze-outs of minority shareholders and some of the economic implications of the law in this regard.

50 A related literature looks at the roles of main banks in German and Japanese governance. See, e.g., Hoshi et al. (1990), Hellwig (1991), and Kaplan (1994a,b). Hellwig (2000) contains a partial survey of this literature.

51 As an example, suppose an individual owns 51% of company A, which in turn owns 51% of company B, which in
in Europe and Asia than in the US. See Becht et al. (2003, §5.2) for more on blockholders.

A remark is in order at this juncture. Although dispersed ownership would seem to limit shareholders’ ability to control management effectively, it is nevertheless an equilibrium phenomenon. In other words, as Demsetz and Lehn (1985) note, ownership structure is endogenous. Consequently, if we observe dispersed ownership, then we must at least entertain the idea that the costs of dispersion are not necessarily that great (or that the net benefits of concentration are not that great). Otherwise, it is hard to see how the system could have persisted—$100 bills do not linger long on the sidewalk. This is not—I hasten to add—to claim that dispersed ownership isn’t a binding constraint with respect to achieving the ideal in governance, but rather that its cost cannot be excessive.

This last point can be readily illustrated. Let return, $r$, be distributed normally with mean $\mu + g$ and variance $\sigma^2$. Suppose, initially, the firm has a large shareholder. She oversees management. Suppose she is undiversified and, so, concerned about risk. In particular, assume her utility is $-\exp\left(C(g) - \alpha r\right)$, where $C(g)$ is the cost she incurs to monitor—provide level of governance $g$—and $\alpha \in (0, 1]$ is her share of the company. A monotone transformation of her expected utility is

$$\alpha (\mu + g) - \frac{1}{2} \alpha^2 \sigma^2 - C(g).$$

Suppose that $C : \mathbb{R} \to \mathbb{R}$ is strictly convex. Application of Lemma 1 reveals that the $g$ that maximizes her expected utility cannot increase as $\alpha$ decreases (and is falling in $\alpha$ whenever it is positive).

Imagine, initially, that this shareholder owns 100% of the company (she is its founder, perhaps). Assume she can sell shares to diversified and atomistic shareholders, who can therefore be considered to be risk neutral. They will pay the expected return, $\mu + g$, times the portion of the firm they purchase; hence, the large shareholder gets $(1 - \alpha)(\mu + g)$ with certainty. Her expected utility is, then,

$$\mu + g - \frac{1}{2} \alpha^2 \sigma^2 - C(g).$$

If the large shareholder could commit to a level of governance, then clearly she does best to choose the $g$ that maximizes (8). That could be time inconsistent, however: when she actually chooses $g$, her relevant utility turn owns 51% of company C. Given that the individual controls A, which controls B, which controls C, the individual controls C. But he or she has a claim on just $0.51^3 \approx 13\%$ of company C’s returns.
is given by (7) and her incentive would be to maximize it. Let us suppose that is what will happen. Let $g(\alpha)$ maximize (7). New investors anticipate this and take it into account when bidding for shares; hence, they pay $(1 - \alpha)(\mu + g(\alpha))$ for the portion of the firm they buy. The proportion of the corporation to which the large shareholder hangs on maximizes

$$\mu + g(\alpha) - \frac{1}{2} \alpha^2 \sigma^2 - C(g(\alpha)).$$

Utilizing the envelope theorem, the first-order condition for maximizing that expression with respect to $\alpha$ is equivalent to

$$(1 - \alpha)g'(\alpha) - \alpha \sigma^2 \leq 0,$$

with equality holding only if the optimal $\alpha$ is positive. Observe that $\alpha = 1$ cannot be a solution to (9). The large shareholder will sell some portion of the company even though that guarantees a worsening of governance. The intuition is clear: Starting from $\alpha = 1$, a small reduction in $\alpha$ has a second-order effect on the level of governance (and, thus, share value), but a first-order effect on the risk she must bear. Not surprisingly, therefore, she will sell some of her shares.

Observe that the large shareholder would be strictly better off if she could commit to a governance level before selling her shares (i.e., commit to the $g$ that maximizes (8)). Let $g^*$ be the level of governance that maximizes (8). If such commitment were possible, then the large shareholder would sell 100% of her stake. If we assume that the large shareholder’s commitment to $g^*$ is part of a contract with the purchasers, then it is readily shown to be renegotiation proof: If the large shareholder sought to reduce governance by $\Delta$, then she could get the purchasers to agree only if they receive $\Delta$ in compensation. Given $g^*$ maximizes (8), $\Delta = 0$ must maximize $-\Delta - C(g^* - \Delta)$.

Here, as shown, the large shareholder’s commitment to a governance level is renegotiation proof. This is unlikely to be a general result nor one robust to the introduction of more real-world issues. First, as has been shown in a variety of important settings, many ideal contracts are not renegotiation proof (a particularly pertinent example with respect to governance would be Fudenberg and Tirole, 1990). Second, new information could arise that renders a previously optimal contract suboptimal. Admittedly, in theory, the original contract could be designed so as to allow the incorporation of any new information. In reality,
however, bounded rationality, legal constraints, and difficulties in verifying all relevant new information mean that no contract could incorporate all new information, which means renegotiation could be Pareto improving. If parties can benefit from renegotiating, then it seems reasonable to suppose that they will.\textsuperscript{52}

One reason commitment is difficult is that people cannot be bound to a firm. For example, the large shareholder above could quit her job as the supplier of governance. Given this, the parties would again be limited to time-consistent governance levels. Even if the large shareholder could not quit without penalty under her employment contract, as a matter of law such penalties are limited (both by the ability of the large shareholder to seek bankruptcy protection and by the law’s abhorrence of penalty clauses—see, \textit{e.g.}, Hermalin et al., 2007, §\textsuperscript{5}). The ability of key executives to quit gives them a lever with which to force the firm to renegotiate with them. In theory, a firm might seek to develop a reputation for never renegotiating, but as a practical matter achieving such a reputation is daunting and, thus, unlikely.\textsuperscript{53}

Inability to commit or to avoid renegotiation can potentially explain many oddities of governance. For instance, suppose that investors put in place what they believe to be the optimal governance system. Over time, the firm’s CEO proves himself to be of high ability. As such, the CEO gains bargaining power—he has proved himself a rare commodity, which, because he cannot be bound to the firm, provides him market power. One thing he will want is more money. Another potentially is laxer governance (perhaps, \textit{e.g.}, for career-concern issues as discussed earlier). Paying the CEO more is a first-order loss for the investors. Relaxing the governance system, given that it was at an optimum, is a second-order loss. Hence, investors will initially be more yielding to the CEO’s demand for laxer governance than his demand for more pay. Hermalin and Weisbach (1998) use this insight to explain why a long-serving CEO can gain influence over

\textsuperscript{52}This is actually not a universally accepted principle in the literature; some authors have argued that, in some circumstances, there can be off-the-equilibrium-path outcomes that the parties would like to renegotiate, but can’t (\textit{i.e.}, on some off-the-equilibrium paths the parties leave money on the table). See Hermalin et al. (2007, §4.3.2) for a discussion and survey of the relevant literature.

\textsuperscript{53}One factor that makes developing such a reputation daunting is that, while the firm is infinitely lived, the directors—the people actually doing the negotiating—are not. Again theory permits short-lived individuals to carry an infinitely lived firm’s reputation—see, \textit{e.g.}, the relevant sections of the corporate culture chapter in this volume—but it by no means guarantees they can do so. Furthermore, it is possible that directors wouldn’t wish to promote a norm of no renegotiation: Many of them are, themselves, CEOs of other firms; others of them make a non-trivial income serving as directors and could therefore wish to avoid developing a reputation of being too hard-nosed (given that some set of CEOs play a role in their obtaining their board seats). See Adams et al. (2010) for a discussion of the incentives directors have to avoid being seen as “boat rockers.”
the board of directors: At any point in time, the incumbent board believes it is engaging in the optimal level of monitoring. As such, giving in to the CEO’s demand to add a director friendly to him to the board, which will lower monitoring somewhat, is a second-order loss for the incumbent board and, thus, preferable to giving in to the CEO’s demand for more pay. In other words, the CEO’s utility can be increased either through more money or less governance and, on the initial margin, the latter costs the investors less than the former.

7 More Thoughts on Shareholder Control

The famous German banker Carl Fürstenberg (1850–1933) once remarked: “Shareholders are stupid and impertinent—stupid because they give their money to somebody else without any effective control over what this person is doing with it and impertinent because they ask for a dividend as a reward for their stupidity” (quoted in Hellwig, 2000, p. 109). Although exaggerated, as witticisms often are, the remark nonetheless illuminates a basic truth: The legitimacy of shareholder complaints about their lack of control is limited—no one compelled them to purchase equity and they are, in expectation at least, compensated for the risks they bear. Consequently, one should be cautious about extolling the virtues of increasing shareholder control, especially if imposed via regulatory fiat.

In particular, most of the governance relations we observe emerged as the consequence of (at least implicit) bargaining between investors and controlling parties. The general presumption in economics is that private parties will arrive at a better contract than one than can be imposed by the state. There are, to be sure, exceptions to this presumption (see, e.g., Hermalin et al., 2007, §2, for a discussion). But even if these exceptions apply, their existence does not necessarily entail intervention on behalf of the shareholders.

Another reason to be cautious about interventions that strengthen the shareholders’ control is that shareholders are not the only investors in the firm. There are other providers of financial capital, such as the debtholders. Another class of investors are the workers who invest human capital (as noted above and in Tirole, 2001, and Roberts and van den Steen, 2003).

With respect to debtholders, there is the well-known asset substitution problem (Harris and Raviv, 1991): Given limited liability, shareholders have an incentive to want the firm to take on risk, even at the expense
of expected return, as the firm nears bankruptcy. Hermalin and Wallace (1994) provide illustrative empirical evidence on the matter: In the 1980s, stock savings and loans (thrifts) held far riskier portfolios than did mutual savings and loans. Consequently, the probability of a stock savings and loan going bankrupt was 4–1/2 times that of a mutual. This despite the fact that evidence shows stock thrifts were more efficiently operated than mutuals ceteris paribus.

Similarly, at any point in time, the interests of investors of financial capital and those of investors of human capital need not coincide. For instance, changed circumstances can make physical capital more valuable in alternative uses than in its current uses, but such a reallocation of physical capital could severely reduce the return on investment to human capital. If the financial investors have sufficient control of the enterprise, they could decide to reallocate the physical capital at the expense of the human capital investors (the workers). Indeed, in some circumstances, they could use this power to expropriate some of the return to human capital (e.g., threaten reallocation in order to reduce wages). Shleifer and Summers (1988) raise this issue in terms of assessing the social desirability of takeovers. Roberts and van den Steen (2003) present some formal models that explore these issues. In short, when the concerns of other investors are taken into account, it is not obvious that less-than-full shareholder control is a bad thing.

With respect to the concern that shareholders have too little control, I would argue the following:

1. The degree of shareholder control could reflect a rational tradeoff between the benefits of control and the wish to hold a diversified portfolio. (See previous section.)

2. The degree of shareholder control also reflects initial and ongoing negotiations between shareholders and management. (See Section 4.) In particular, shareholders may tradeoff less control in exchange for less managerial remuneration. (Again, see previous section.)

3. The degree to which shareholders lack control may be overstated due to issues with the empirical tests that have sought to gauge the strength of governance. (See Section 5.)

4. Because there are other investors, giving shareholders (or a subset of shareholders) too much control can adversely affect these other investors (including minority shareholders) and, thus, their willingness to invest.
8 Where Does the Field Go Next?

To conclude this chapter, let me briefly outline what I see as promising directions for future research in the area of corporate governance:

New Empirical Approaches: Although nevertheless of value, it is also the case that much of the existing empirical literature in corporate governance is plagued with problems. Identification is an especially difficult problem given that almost all variables of interest are endogenous. New identification strategies would be of great value. For example, the logic of regression discontinuity design (see Lee and Lemieux, 2010, for a survey) could be used: Firms that just satisfy certain regulatory restrictions are likely the firms for which these restrictions are binding and, hence, arguably value reducing. Their behavior and performance should, ceteris paribus, vary from those that satisfy the restrictions “with room to spare.” In the same vein, imposition of restrictions can serve as a “natural” experiment. This will be especially valuable when some firms are “assigned” to the treatment group (must adjust their governance structures to come into compliance) and others are assigned to the control group (do not need to adjust their structures to be in compliance).

Second, interpreting regression results is complicated by the heterogeneity in governance structures. As Demsetz and Lehn (1985), among others, have noted, causal interpretations of regressions between governance structures and firm performance are especially problematic in this regard (even assuming the joint endogeneity issue has been resolved). As I have tried to sketch out above, economic theory offers ways to think about such regressions, but more work is needed both with regard to developing the theory and to its use in the design of empirical tests. The heterogeneity issue is also an issue insofar as it makes the results of cross-sectional regressions suspect; there is strong reason to believe many such regressions will have coefficients that are biased downward. Various hierarchical models and other variations on random-coefficients models should be employed. Such models dictate the use of panel data sets.

Theory: Many aspects of governance represent applications of agency and other contracting theory. There are dimensions to governance, however, that fall outside the general frameworks or that otherwise
exhibit novel features. One direction is to expand the canon beyond the typical single principal-single agent perspective (*e.g.*, shareholders and CEO). Governance involves multiple principals (*e.g.*, large and small shareholders; equity and debtholders; etc.), hierarchical structures (*e.g.*, shareholders-board of directors-management), and teams of agents, among other features. Better modeling of governance with multiple classes of investors, especially taking human-capital investment seriously, is another priority. Theory should also explore the *dynamic* aspects of governance: How do governance institutions evolve over time, optimally and in practice? To what extent is governance a nexus of contracts aimed at controlling the management of the firm in an ever changing environment? To what extent is governance a substitute for formal contracting, reflecting the inherent impossibility of complete contracting over long horizons with unforeseen contingencies? Finally, what do less neo-classical concepts, such as leadership, corporate culture, and various cognitive biases have to say about the structure of corporate governance?

**Perspective:** As noted, a common view in governance research is that the aim of governance is to improve shareholder control. For a number of reasons, it is not clear that should necessarily be the aim of governance. A more appropriate view could be that governance is about balancing the, possibly disparate, needs of investors to have some measure of protection against mismanagement, misallocation, and misappropriation against the various costs of providing these protections.

Given the importance of the topic, governance research continues to be valued. With a large number of open questions, there are plenty of *terrae incognitae* left to be explored.
References


Hermalin

References

46


