

NBER WORKING PAPER SERIES

ASSESSING MANAGERIAL ABILITY:
IMPLICATIONS FOR CORPORATE GOVERNANCE

Benjamin Hermalin
Michael Weisbach

Working Paper 23121
<http://www.nber.org/papers/w23121>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2017

Forthcoming in *The Handbook of the Economics of Corporate Governance*. The authors thank Vivian Fang, Jongha Lim, Yihui Pan, Miriam Schwartz-Ziv, Berk Sensoy, Léa Stern, Luke Taylor, Ralph Walkling, and Tracy Wang for helpful comments on an earlier draft, and Shan Ge for excellent research assistance. Hermalin gratefully acknowledges the financial support of the Thomas & Alison Schneider Distinguished Professorship in Finance and the hospitality of Nuffield College, Oxford, where work on this chapter began. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Benjamin Hermalin and Michael Weisbach. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Assessing Managerial Ability: Implications for Corporate Governance
Benjamin Hermalin and Michael Weisbach
NBER Working Paper No. 23121
February 2017
JEL No. D81,D83,G34,M12

ABSTRACT

A manager's current and potential future employers are continually assessing her or his ability. Such assessment is a crucial component of corporate governance and this chapter provides an overview of the research on that aspect of governance. In particular, we review how assessment generates incentives (both good and bad), generates risks that must be faced by both managers and firms, and affects the contractual relationships between those parties in important ways. Assessment (or learning) proves a key perspective from which to study, evaluate, and possibly even regulate corporate governance. Moreover, because learning is a behavior notoriously subject to systematic biases, this perspective is a natural avenue through which to introduce behavioral and psychological insights into the study of corporate governance.

Benjamin Hermalin
Vice Provost
University of California
200 California Hall, #1500
Berkeley, CA 94720-1500
hermalin@haas.berkeley.edu

Michael Weisbach
Department of Finance
Fisher College of Business
Ohio State University
2100 Neil Ave.
Columbus, OH 43210
and NBER
weisbach.2@osu.edu

CONTENTS

1	INTRODUCTION	1
2	THE BASICS OF LEARNING	5
2.1	The Issue	5
2.2	An Application: A Model of CEO Dismissal/Retention	7
2.3	Empirical Analyses of Management Turnover and Learning	11
3	CAREER CONCERNS	13
3.1	Career Concerns: Theory	14
3.2	Career Concerns: Empirics	20
3.2.1	Evidence from Intertemporal Pattern of Incentives	20
3.2.2	Estimating Learning about CEOs of Public Corporations	21
3.2.3	Evidence from the Money Management Industry	22
4	SIGNAL JAMMING: PROJECT CHOICE AND MANAGERIAL MYOPIA	24
4.1	Project Choice	24
4.1.1	When No Project May be Available	29
4.2	Managerial Myopia	33
4.3	Empirical Analyses of Myopia	37
5	LEARNING ABOUT GOVERNANCE AND CAPITAL MARKETS	39
5.1	Learning Models and Capital Markets	39
5.2	How Learning about Management Affects Capital Markets	40
5.3	Capital Market Inferences about Other Executives and Other Dimensions of Governance	42
6	FURTHER IMPLICATIONS FOR GOVERNANCE	43
6.1	Bargaining Over Oversight	43
6.2	Consequences for Choice of Manager	48
6.3	Consequences for Choice of Information Structures	55
7	HOW NON-BAYESIAN UPDATING AFFECTS GOVERNANCE	59
7.1	How Do People Really Update Their Beliefs?	60
7.2	Implications of Biases for Interpretation of Learning Models	61
7.2.1	The Base-Rate Fallacy	61
7.2.2	The Hot-Hand Fallacy	63
7.2.3	The Fundamental Attribution Bias	64
7.3	Potential Future Research Applying Behavioral Economics to Corporate Governance	64
8	CONCLUSION	66
	APPENDIX A: TECHNICAL DETAILS	68
	REFERENCES	75

1 INTRODUCTION

A good definition of corporate governance is it is “the way in which suppliers of finance . . . assure themselves of getting a return on their investment” (Shleifer and Vishny, 1997, p. 737). There are many dimensions to obtaining such assurance. One, much-studied in the literature, is *control*: aligning the interests of the managers who make decisions with the interests of the investors affected by their decisions.¹ How, for instance, do shareholders or a board of directors keep a CEO from pursuing an agenda that benefits him but not them? Another key dimension, but one much less well-studied, has to do with *assessing* the quality of the decisions being made by managers. Even if managers had the best of intentions *vis-à-vis* their investors, if they are incompetent, pursue poor projects, or adopt ill-conceived strategies, then investors will suffer. The “suppliers of finance” must, therefore, continually assess the ability of their managers. This insight entails a number of questions: for instance, how should a CEO be assessed with respect to her intrinsic ability, competency, and match with her employer? A further, somewhat subtle, but critical question is how does knowing assessment will occur affect the behavior of the relevant actors?

The last question is especially relevant: absent strategic reactions, making inferences about attributes, such as ability, from various correlates, such as firm profits, would be a relatively straightforward and unexciting statistical exercise. However, as has long been known in the physical sciences, the act of observation can directly affect the phenomenon being observed: an insight with important implications for incentive provision and corporate governance, as noted initially by Fama (1980) and Holmstrom (1982) [1999].² Moreover, unlike a sub-atomic particle or an animal in the wild, a manager understands the consequences of being observed. In particular, because the outcome of assessment is unknown in advance, the act of observation exposes the manager to risk—for which he or she will demand compensation—with the amount of risk and, thus, compensation being dependent on the assessment protocol. Furthermore, also unlike a particle or an animal, a manager is a strategic player, who can, therefore, be expected to try to influence how he or she is observed and by whom. In short, assessment—learning—profoundly affects corporate governance by generating incentives (some good, some bad); by creating a tradeoff between the accuracy of the assessment and the risk that imposes on the manager; and by creating an impetus for the manager to try to affect by whom he is observed (and governed).

This chapter presents an overview of the scholarly literature on assessment

¹Smith (1776, p. 700): “The directors of companies, however, being the managers . . . of other peoples money . . . it cannot well be expected, that they should watch over it with the same anxious vigilance [as owners].” In the modern era, there was a revived interest in agency and its relation to governance led by Berle and Means (1932), Williamson (1963), and Jensen and Meckling (1976), among others. See Becht et al. (2003) andermalin (2013) for recent surveys of the relevant literature.

²Holmstrom’s paper was originally published in 1982 as a chapter in a hard-to-find festschrift for Lars Wahlbeck. In the pre-PDF 1980s, a photocopy of the working-paper version of the chapter was a treasured belonging of students and scholars interested in contract theory and governance. In 1999, the *Review of Economic Studies* wisely reprinted the paper.

and learning as it pertains to corporate governance. We begin, in Section 2, with the basics of learning models and their application to the issue of CEO turnover, the decision to dismiss or retain a CEO in light of evidence about her “ability.” At some level, ability is a multi-dimensional and amorphous concept. It can describe a manager’s decision-making or leadership capabilities, the fit between her style and skills with the needs of her firm, or even the fit between her strategic vision and the firm’s capabilities. From a modeling perspective, however, it is possible to act as if ability is a one-dimensional attribute that has a known mapping into the expected profits of the firm (indeed, in many models, a manager’s ability *is* expected profit under her leadership).³ In many models, this ability is a fixed, but *ex ante* unknown, attribute of the manager. How the firm performs under her command provides information—or *signals*—about her ability from which those monitoring her can update their beliefs about her ability via Bayes Theorem. If this updating leads her employers to infer the firm would be more valuable under a different CEO, less any transition costs, then she will be dismissed; otherwise, she is retained. Although the models considered in Section 2 are straightforward from a statistical perspective, they are rich enough to yield insightful comparative statics. Moreover, as we show in the second half of that section, the implications of those comparative statics are consistent with empirical analyses of managerial turnover.

Knowing that they are being monitored, with resulting inferences about their ability affecting their careers (not only retention of current positions, but also their future compensation), managers have an incentive to influence the monitoring process. In many settings, there are actions that they can take to alter the performance signals on which estimates of their abilities are based. For instance, supplying more effort may boost performance. In Section 3, we present the Holmstrom (1999) model, in which a manager attempts to influence the assessment of her ability via the hidden actions she takes (effort she supplies). This being a game, those monitoring her—in particular the market for managers—understand she is supplying effort to influence the performance measure; that is, she is engaged in “signal jamming.” Those monitoring will, consequently, adjust for anticipated effort when estimating managerial ability; in equilibrium, no one is fooled. Although she fools no one, the manager must nonetheless expend this effort because it is expected (were she not to do so, the market would still make the adjustment, to her detriment). Consequently, in equilibrium, career concerns serve to provide the manager an incentive to supply effort; however, as Holmstrom shows, these career concerns do not, generically, lead to her providing the first-best level of effort.

Having set forth the basic model of career concerns, the remainder of Section 3 considers the empirical literature on career concerns. The idea that career concerns serve to motivate workers is very general, and likely applies to nearly every profession. Unfortunately, testing this directly is difficult: to do so would require data on individual outputs and also on their career paths. In

³To be more realistic, one could view ability as that portion of expected profit that is attributable to the manager.

most settings, such data are unavailable. Yet the importance of career concerns on workers' motivations is sufficiently great that the literature has sought various indirect tests or sought to measure the effects in fields where the data are arguably available.

Because career-concern models make predictions about the intertemporal pattern of executive compensation, one empirical approach has been to see whether the data are consistent with those predictions. As reviewed in Section 3.2, the results from these studies are generally consistent with the predictions of career-concern models.

Another empirical approach, also discussed in Section 3.2, is to consider industries in which managerial performance and its consequence for managerial remuneration are readily observed. One such industry is the money management industry. Several papers have documented that, in the private-equity and hedge-fund industries, fund flows respond to performance as Bayesian updating about managers' abilities would suggest they should; indeed, because the fees management receives are usually proportional to money under management, the consequence of even marginally better performance on managers' lifetime incomes is substantial.

As noted, career concerns induce incentives to engage in signal jamming. In the portion of Holmstrom (1999) summarized above, such signal jamming can be seen in a positive light insofar as it creates incentives for managers to supply more effort. As, however, Holmstrom goes on to show, such signal jamming can also be detrimental to the employer's (shareholders') interests; for example, as he shows, it can lead managers to eschew projects with positive expected net-present values. We explore in greater depth the potentially perverse incentives created by career concerns in Section 4. With regard to project choice, Section 4.1 shows that because how the market will update its beliefs about a manager are *ex ante* uncertain, such updating necessarily exposes the manager to risk. The nature of that risk depends not only the underlying risk of the projects themselves, but also what the market knows or infers about their risk. For example, as discussed below, there is a somewhat counter-intuitive result that when the market knows the underlying risk of different projects, the manager can have an incentive to choose the *riskiest* project, regardless of expected return, because the riskier the project, the less information it reveals about the manager's ability, and thus the less the market will update about her ability based on the outcome of the project.

A related idea is that when the market values a firm, it draws inferences about its prospects the same way it does when it evaluates managerial ability; that is, via Bayesian updating of beliefs based on current performance metrics (*e.g.*, earnings). This potentially creates an incentive to signal-jam current performance in the hopes of influencing that updating and, thus, boosting estimated value, which is to say the stock price. Consequently, when managers have motive to care about the stock price, either through ownership or through equity-linked incentive plans, managers have incentives to signal-jam. In some circumstances, such signal jamming means neglecting longer-term investments in favor of ones with more immediate payoffs (that will boost current perfor-

mance). Such a myopic strategy could be at odds with truly maximizing firm value if the longer-term investments have the greater net present value (NPV). Although, in equilibrium, no one is fooled—firm value is correctly estimated—the evaluation process effectively traps managers into behaving in this myopic manner, since if they behave at odds with the market expects—invest in longer-term projects with more inherent value—the market will draw the wrong inference and discount the firm’s value. In Section 4.2, we investigate this using a simplified version of the Stein (1989) model, which originally developed this idea. We then conclude Section 4 by reviewing related empirical work. This work documents that managers appear to prefer *ceteris paribus*, shorter-term investments over longer term investments, especially if they are compensated through stock-based compensation plans, results consistent with the signal-jamming story.

Models of learning have applications to areas of economics beside governance. One is the functioning of capital markets: a number of “anomalies” pertaining to stock returns can be explained by the idea that the market learns about corporate attributes. In Section 5, we review the Pastor and Veronesi (2003) model, in which learning is critical to understanding the distribution of stock returns. We then show that this model can be reinterpreted in terms relevant to corporate governance: the attributes about which the market learns are managerial abilities and fit with their firms. Such a reinterpretation formalizes the notion of “management risk,” which is how uncertainty about an executive’s ability affects the volatility of his company’s stock prices and his company’s cost of debt. When there is less uncertainty about ability, both volatility and the cost of debt are lower. There is empirical support for these ideas: among other results, studies indicate that the less uncertainty there is about the executive, the less volatile is the company’s stock price and its cost of debt.

In addition to changing how managers behave, learning has implications for how owners and boards of directors behave. For instance, how actively owners and boards monitor their executives depends on the potential information has to affect their actions. Monitoring—seeking to obtain information—is intrinsically more valuable when there is more uncertainty about an executive’s ability than when there is little uncertainty. Monitoring of an established CEO is less valuable than monitoring a new CEO. Monitoring a CEO who has a strong track record and for whom there is, therefore, little likelihood, of obtaining signals that would lead to her dismissal is less valuable than monitoring one with a more modest track record. As we discuss in Section 6, such implications have consequences for the structure of the board of directors (how independent it will be) and the firm’s choice of CEO in the first place. Because of the consequences for a CEO’s wellbeing, there are indirect costs to having better monitoring of the CEO: making the CEO’s situation worse necessitates paying her more, which means shareholders won’t always consider more (better) information about managerial performance as an unalloyed good.

Much of the analysis discussed in this chapter and employed in the literature presupposes that learning is rational; in particular, that it is in accord with Bayes Theorem (*i.e.*, is Bayesian learning). The notion that individuals process information rationally and act according to Bayes Theorem has been

challenged by a large body of work in psychology. That literature finds consistent evidence that people are subject to a series of biases that lead them to form beliefs in ways that are inconsistent with Bayes Theorem. Broadly, these biases lead people to place too much weight on new signals and not enough on prior information. Section 7 reviews the evidence that people are in fact “bad Bayesians,” as well as the implications of that for learning models of corporate governance. We argue that, in fact, these departures from rationality make learning-based models of corporate governance more important than they would be if agents were perfectly rational. Since the “action” in learning models comes from agents updating their beliefs, the behavioral biases that real-world actors exhibit suggest that, in reality, they will update more than rational-actor models would suggest. This, in turn, means the models could well *understate* the effects they describe. In other words, learning models appear to be one area of economics in which behavioral biases make the implications of rational-actor models even more important than one might otherwise imagine.

Some technical details, including proofs not given in the text, can be found in the Appendix.

2 THE BASICS OF LEARNING

2.1 THE ISSUE

People constantly evaluate each other, typically with an eye towards making predictions about how they will do in the future. Examples abound: a school evaluates would-be students to see if they will do well if admitted; we evaluate would-be romantic partners to see if they will prove a good match; and employers evaluate employees to ascertain the value they will create going forward. Within the realm of corporate governance, the last of these is especially relevant: shareholders—or their representatives, the board of directors—evaluate executives (existing or potential) to predict their future performance. They do so because they wish to know an executive’s value (*e.g.*, what is the maximum they should pay her); alternatively, they might want to make a decision about retaining her or gambling on putting someone else in charge.

Within economics, such evaluation is modeled via learning models. To wit, the employer’s (shareholders’) *expected* payoff in period t is assumed to be a function of the employee’s (executive’s) then-current ability, $\alpha_t \in \mathcal{A} \subseteq \mathbb{R}$, where \mathcal{A} is the set of possible ability levels.^{4,5} The employer is assumed not to know the employee’s ability, but he knows its relevant statistical properties. Additionally, he has some set of past signals—possibly just past performance—that he can

⁴Generalizing ability to be multi-dimensional (*e.g.*, $\alpha_t \in \mathbb{R}^n$) gains one little: an employer cares about his expected payoff, $B(\alpha_t)$, and one could simply define ability to be $B(\alpha_t)$. The statistical properties of the random variable $B(\alpha_t)$ would be inherited from those of α_t in the usual way. Multiple dimensions could, however, be relevant if possible employee-employer matching across different employers was at issue.

⁵There is also no gain to indexing \mathcal{A} by period: just define $\mathcal{A} = \bigcup_{t=0}^{\infty} \mathcal{A}_t$, where \mathcal{A}_t is the set of possible time- t ability levels.

use to make inferences about ability at time t .

The standard model assumes ability does not change over time.⁶ When ability is time invariant, we simply denote it as α (*i.e.*, without a time index) and assume it is drawn from a known distribution, $F_0 : \mathbb{R} \rightarrow [0, 1]$, with associated density $f_0 : \mathbb{R} \rightarrow \mathbb{R}_+$. Let $\hat{\alpha}_0$ denote the expected value of ability given distribution F_0 . The entities F_0 , f_0 , and $\hat{\alpha}_0$ are known, respectively, as the prior distribution, prior density, and prior of ability (the rationale for the zero subscript will become evident shortly).

As alluded to earlier, the employer observes signals that permit him to update his beliefs about the employee's ability. Specifically, let $s_t \in \mathbb{R}$ denote the signal he observes at time t (*e.g.*, s_t is the realization of profit at time t or an indicator of whether the period- t project was successful).⁷ The signal in any given period is drawn from a conditional distribution $G(\cdot|\alpha) : \mathbb{R} \rightarrow [0, 1]$. Let $g(\cdot|\alpha) : \mathbb{R} \rightarrow \mathbb{R}_+$ denote the corresponding conditional density. Consistent with this notation, the assumption is that, conditional on true ability α , the signal at one time is distributed identically and independently of the signal at any other time. Two relevant examples are:

1. $s_t \in \{0, 1\}$ — the period- t project either fails or succeeds, $g(0|\alpha) = 1 - \alpha$ is the probability it fails, and $g(1|\alpha) = \alpha$ is the probability it succeeds; that is, ability is equivalent to the probability of being successful.
2. $s_t = \alpha + \varepsilon_t$, where $\varepsilon_t \sim \Psi : (-\infty, \infty) \rightarrow [0, 1]$. As the signal could always be redefined as $\tilde{s} = s - \mathbb{E}\{\varepsilon\}$, there is no loss of generality in assuming $\mathbb{E}\{\varepsilon\} = 0$. Observe that $G(s|\alpha) = \Psi(s - \alpha)$.

It follows from Bayes Theorem that

$$f_1(\alpha|s_1) = \frac{g(s_1|\alpha)f_0(\alpha)}{\int_{\mathcal{A}} g(s_1|z)f_0(z)dz}, \quad (1)$$

where the integral in (1) is to be understood to encompass simple summation if \mathcal{A} is finite (*i.e.*, ability is discretely distributed). More generally,

$$f_t(\alpha|s_1, \dots, s_t) = \frac{g(s_t|\alpha)f_{t-1}(\alpha|s_1, \dots, s_{t-1})}{\int_{\mathcal{A}} g(s_t|z)f_{t-1}(z|s_1, \dots, s_{t-1})dz}. \quad (2)$$

In practice, obtaining a closed-form solution for the recursive formula (2) in terms of the primitives, f_0 and g , is difficult, especially if one wants a solution for an arbitrary t . One exception is if the signals are independent Bernoulli trials, as in the first example above. In that case, the probability of n successes in t trials given ability α is

$$\binom{t}{n} \alpha^n (1 - \alpha)^{t-n};$$

⁶See Holmstrom (1999) for an analysis in which ability follows a Gaussian random walk.

⁷Given that ability is uni-dimensional, there is little to be gained by assuming multi-dimensional signals (*e.g.*, $s_t \in \mathbb{R}^m$). If the model is tractable, one can form a sufficient statistic from s_t and, thus, one could just define s_t as that sufficient statistic.

hence, employing Bayes Theorem,

$$f_t(\alpha|n) = \frac{\alpha^n(1-\alpha)^{t-n}f_0(\alpha)}{\int_{\mathcal{A}} z^n(1-z)^{t-n}f_0(z)dz}. \quad (3)$$

So if ability is also the realization of a Bernoulli trial (*i.e.*, $\mathcal{A} = \{\underline{\alpha}, \bar{\alpha}\}$), then

$$f_t(\bar{\alpha}|n) = \frac{\bar{\alpha}^n(1-\bar{\alpha})^{t-n}f_0(\bar{\alpha})}{\bar{\alpha}^n(1-\bar{\alpha})^{t-n}f_0(\bar{\alpha}) + \underline{\alpha}^n(1-\underline{\alpha})^{t-n}f_0(\underline{\alpha})}.$$

As a second example, if ability is drawn from a uniform distribution on $[0, 1]$, then $F_t(\cdot|n)$ is a beta distribution with parameters $n+1$ and $t-n+1$.⁸

Another model that permits a closed-form solution, and which has been widely used in the literature, is (i) to assume that ability is drawn from a normal distribution with mean $\hat{\alpha}_0$ and precision τ_0 (the *precision* of a distribution is the multiplicative inverse of its variance); and (ii) to assume that $s_t = \alpha + \varepsilon_t$ (as in the second illustration above), where ε_t is distributed normally with mean zero and precision η . Then it can be shown (see, *e.g.*, DeGroot, 1970, p. 166) that $F_t(\cdot|s_1, \dots, s_t)$ is a normal distribution with mean

$$\hat{\alpha}_t = \frac{\tau_0\hat{\alpha}_0 + \eta\sum_{i=1}^t s_i}{\tau_0 + t\eta} = \frac{\tau_0\hat{\alpha}_0 + t\eta\bar{s}}{\tau_0 + t\eta}, \quad (4)$$

where \bar{s} is the arithmetic average of the t signals, and precision

$$\tau_t = \tau_0 + t\eta. \quad (5)$$

Because of its reliance on the normal distribution, the literature refers to this as the *normal learning model*.

2.2 AN APPLICATION: A MODEL OF CEO DISMISSAL/RETENTION

As an application of the normal learning model, consider the following. Shareholders (or their representatives, the board of directors) hire a CEO from a pool of *ex ante* identical CEOs. Assume that the ability of any CEO is drawn from a normal distribution with mean 0 and precision τ_0 (each CEO's ability being an independent draw from this common distribution). Setting the mean to zero is purely for convenience and without loss of generality. The timing of moves is:

1. Risk-neutral shareholders (alternatively, the board—here assumed to be perfect agents) set an information policy (more on this below).
2. A CEO is hired from the aforementioned pool of CEOs.

⁸Recall that a uniform distribution is equivalent to a beta distribution with both parameters equal to 1. It can be shown more generally, see DeGroot (1970, p. 160), that if α is distributed beta with parameters ξ and ζ , then $F_t(\cdot|n)$ is a beta distribution with parameters $\xi+n$ and $\zeta+t-n$.

3. A signal about the CEO's ability, s , is realized; its distribution is normal with a mean equal to the CEO's ability and a precision, η , that depends on the information policy determined at stage 1.
4. Based on the signal, the shareholders (board) decide to keep or replace the CEO. The firm incurs a replacement cost of $r > 0$ if it replaces him. A replacement CEO is drawn from the same pool as the original CEO.
5. The shareholders realize a payoff with expected value equal to the ability of the CEO in place at this stage.

Let $\hat{\alpha}_1$ denote the posterior estimate (mean) of CEO ability following the realization of the signal. From (4),

$$\hat{\alpha}_1 = \frac{\eta s}{\tau_0 + \eta}.$$

Observe that $\hat{\alpha}_1$ is the shareholders' expected payoff should they retain the CEO. If, instead, they replace him, the expected payoff generated by the replacement is 0 by assumption. Hence, replacing the initial CEO is rational if and only if

$$-r \geq \hat{\alpha}_1 \iff -\frac{(\tau_0 + \eta)r}{\eta} \geq s. \quad (6)$$

Recall that $s \sim N(\alpha, 1/\eta)$ (the notation $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2). This is a conditional distribution. It is also necessary to know the *unconditional* distribution of s . Because $s = \alpha_0 + \varepsilon$, it is the sum of two normal random variables and, as such, is normally distributed with a mean equal to the sum of their means—in this case, 0—and a variance equal to the sum of their variances—in this case $1/\tau_0 + 1/\eta$; that is, the unconditional distribution of s is

$$s \sim N\left(0, \frac{\eta + \tau_0}{\tau_0 \eta}\right). \quad (7)$$

It is convenient to denote the precision of the unconditional distribution of s :

$$h = \frac{\tau_0 \eta}{\eta + \tau_0}. \quad (8)$$

Observe $\hat{\alpha}_1 = sh/\tau_0$ and so replacement occurs if and only if $s \leq -r\tau_0/h$.

A useful aside. Let Φ and ϕ denote the distribution and density functions of the standard normal (*i.e.*, $N(0, 1)$), respectively. If $x \sim N(\mu, \sigma^2)$, then the distribution and density functions for x can be expressed, respectively, as

$$\Phi\left(\frac{x - \mu}{\sigma}\right) \text{ and } \frac{1}{\sigma}\phi\left(\frac{x - \mu}{\sigma}\right).$$

Employing the aside, the shareholders' expected payoff at stage 1—the value of the firm—is (see Appendix for derivation):

$$\mathbb{E}\{\max\{-r, \hat{\alpha}_1\}\} = -r\Phi\left(-\frac{r\tau_0}{\sqrt{h}}\right) + \frac{\sqrt{h}}{\tau_0}\phi\left(-\frac{r\tau_0}{\sqrt{h}}\right) \quad (9)$$

Some observations about expression (9):

1. Because $\max\{-r, \hat{\alpha}_1\} \geq \hat{\alpha}_1$ and strictly so for a set of positive measure, $\mathbb{E}\{\max\{-r, \hat{\alpha}_1\}\} > \mathbb{E}\{\hat{\alpha}_1\} = 0$; hence, (9) is positive.
2. The positive value of (9) reflects the value of the option to replace the initial CEO. In this regard, observe that the limit of (9) as $r \rightarrow \infty$ equals zero: there is no option if replacing the CEO is impossible (infinitely costly).
3. The derivative of (9) with respect to h is

$$\left(-\frac{r^2\tau_0}{2h^{3/2}} + \frac{1}{2\tau_0\sqrt{h}} + \frac{r^2\tau_0}{2h^{3/2}}\right)\phi\left(-\frac{r\tau_0}{\sqrt{h}}\right) > 0. \quad (10)$$

Because the (conditional) precision of the signal, η , enters into the value of the firm, expression (9), only through h , the last observation implies that the derivative of firm value with respect to η must have the same sign as $dh/d\eta$, which from (8) is positive. We can conclude:

Lemma 1. *Ceteris paribus, the value of the firm (i.e., expression (9)) increases in the precision of the signal of ability, η .*

In light of Lemma 1, one might imagine that shareholders should, for instance, desire very precise information relevant to assessing their CEO's ability, such as highly accurate accounting and close monitoring by the board of directors.⁹ Of course, one needs to take into account the direct costs of obtaining such information, which need to be balanced by the benefits. There are also, as noted by Hermalin and Weisbach (1998, 2012), indirect costs arising from the fact that the CEO must agree to employment in the first place and he may have opposing preferences concerning the quality of information.

As a first pass as to why executives could have opposing preferences, suppose the initially hired CEO's payoff is w if retained and $w - L$ if dismissed, where w is his compensation and $L > 0$ is a loss suffered if dismissed (*e.g.*, a loss of status, costs of relocating, etc.). Normalize CEOs' reservation utilities to zero; that is, a CEO is willing to accept any position that provides him a non-negative expected utility. For purposes here, suppose that the shareholders possess all the bargaining power and can, thus, make take-it-or-leave-it (TIOLI) offers to

⁹This is not the only reason shareholders could prefer greater accuracy. In hidden-action agency models with risk-averse agents, principals benefit from better information systems (see, *e.g.*, Grossman and Hart, 1983; Hermalin and Katz, 1991, 2000; and Kim, 1995). Also, as noted by Cebon and Hermalin (2015), a better informed board can sometimes utilize a better contracting technology than a less well informed one.

CEOs. Because a *replacement* CEO is never fired, she would accept any offer with $w \geq 0$. As the shareholders wish to maximize firm profit, it follows they would offer a replacement CEO $w = 0$.

Before considering the wage offered the initially hired CEO, it is necessary to clarify what he knows about his own ability at time of employment. If his knowledge of his own ability is superior to what a potential employer knows, then bargaining between shareholders and CEO (even TIOLI offers) will convey information to the shareholders about the CEO's ability.¹⁰ For this reason, the literature has, for the most part, focused on information-neutral bargaining, which requires that the employer and employee have identical *ex ante* beliefs about the employee's ability. Although one can debate the literal truth of that assumption, there are reasons to view it as a reasonable approximation. First, one's own assessment of ability is based on past performance, which could well be known to potential employers (this seems especially relevant for executives, whose firms' performance is highly public); second, the uncertainty about a CEO's ability in a particular job is largely uncertainty about the match between him and the firm, which is similarly unknown to both the board and the CEO.¹¹

Returning to the initially hired CEO, what wage is he offered? His expected utility if he accepts employment is

$$w - L\Phi\left(-\frac{r\tau_0}{\sqrt{h}}\right),$$

which reflects that he will be fired if the signal falls below the cutoff given in (6), an event that occurs with probability $\Phi(-r\tau_0/\sqrt{h})$. Because the shareholders get to make a TIOLI offer, the CEO's participation constraint will bind; hence, in equilibrium,

$$w = L\Phi\left(-\frac{r\tau_0}{\sqrt{h}}\right). \quad (11)$$

The righthand side of that expression increases in h and, thus, in η : the greater the precision of the signal, the greater will be the CEO's compensation. As intuition, observe that as $\eta \rightarrow 0$ (the signal is without precision), $h \rightarrow 0$ and, thus, the righthand side of (11) goes to 0 (the argument of Φ goes to $-\infty$): because the signal contains no information, it won't ever be the basis for dismissing the CEO, so he requires no compensation to bear the risk of dismissal. The CEO faces a risk of dismissal only if the signal is informative about his ability and that risk increases with the informativeness of the signal.¹²

It follows that the shareholders bear an indirect cost from a more precise signal: the compensation they must offer the CEO is greater. Taking that into

¹⁰For example, whether the CEO would accept the shareholders' TIOLI offer would depend on his assessment of his ability and, so, the shareholders should update their beliefs about his ability based on his acceptance of their offer.

¹¹This assumption is discussed at length in Holmstrom (1999).

¹²At this juncture, we are admittedly being loose about the relation between the precision of the distribution and informativeness of the signal. However, as Hermalin and Weisbach (2012, Corollary 1) prove, the two concepts coincide in this context.

account, expected firm value is, now, (9) less w or

$$-(r + L)\Phi\left(-\frac{r\tau_0}{\sqrt{h}}\right) + \frac{\sqrt{h}}{\tau_0}\phi\left(-\frac{r\tau_0}{\sqrt{h}}\right) \quad (12)$$

Differentiating (12) with respect to h yields

$$\left(-\frac{Lr\tau_0}{2h^{3/2}} + \frac{1}{2\tau_0\sqrt{h}}\right)\phi\left(-\frac{r\tau_0}{\sqrt{h}}\right);$$

the sign of which is the same as that of $h - Lr\tau_0^2$. It is readily seen that $h = Lr\tau_0^2$ minimizes firm value and, thus, the shareholders are either best off with a signal of minimum precision (within the realm of possible precision levels) or with one of maximum precision (again within the feasible set).

Although a simple model, the analysis above illustrates that employer and employee can have opposing preferences about the precision with which the employee is assessed. Consequently, the employee's compensation will be higher when he is assessed with greater precision than with less precision.

Bibliographic Note. The above model can be seen as simplified version of a more general analysis by Hermalin and Weisbach (2012). Hermalin (2005) uses a similar model to explore, *inter alia*, how trends toward greater monitoring of executives (*e.g.*, via more vigilant boards and tighter reporting requirements) could help to explain a secular trend toward higher executive pay. We return to these and other applications below.

2.3 EMPIRICAL ANALYSES OF MANAGEMENT TURNOVER AND LEARNING

This learning model provides a natural way to understand the relation between management turnover and firm performance. A firm's board is constantly updating its assessment of the firm's CEO and top management team. When the expected profits are lower from the current management than from a potential replacement, a profit-maximizing board should replace the current management. This idea was formally introduced to the governance literature in Hermalin and Weisbach (1998) and provides a framework for understanding the negative empirical relation between firm performance and management turnover.

A large number of papers have documented that there is a positive relation between CEO turnover and poor performance in large corporations, as well as other types of organizations.¹³ In addition, Denis and Denis (1995) document that firm performance generally improves following a CEO turnover, especially a forced turnover. The standard interpretation of this relation is that it measures

¹³Among these are Coughlan and Schmidt (1985), Warner et al. (1988), Weisbach (1988), Barro and Barro (1990), Jensen and Murphy (1990), Blackwell et al. (1994), Huson et al. (2001), Brickley and Van Horn (2002), Engel et al. (2003), Parrino et al. (2003), Eldenburg et al. (2004), Kaplan and Minton (2012), Eisfeldt and Kuhnen (2013), Jenter and Lewellen (2014), and Jenter and Kanaan (2015).

the board’s monitoring; when performance is worse, the board is more likely to find the current CEO unacceptable and, so, replace him.

An important issue in empirical studies of turnover and performance is distinguishing voluntary turnovers (*i.e.*, unforced resignations for reasons unrelated to performance) and involuntary turnovers (*i.e.*, dismissals—perhaps disguised as resignations). Because the true reasons for turnover are often not given publicly, press releases and other public information are often largely or wholly uninformative as to whether a turnover was voluntary or not. Empirical studies have taken different approaches to try to sort voluntary from involuntary managerial change—some ignore the issue, some exclude observations of CEOs at normative retirement age, and some make a detailed effort to distinguish voluntary from involuntary turnover. The algorithm suggested by Parrino (1997) has become more or less the standard approach to assessing whether a particular turnover is “forced”:¹⁴ a plus to it is the turnovers it identifies as forced are likely correctly classified; a minus is it will mis-classify as voluntary those turnovers in which the press fails to report information about pressure exerted on executives to resign. Nonetheless, voluntary turnovers are unlikely to be related to performance and the negative relation between performance and CEO turnover is extremely robust across samples. Therefore, the measured negative relation between turnover and performance probably reflects boards firing CEOs (*i.e.*, the difficulty in distinguishing the two types of turnover merely adds noise to the dependent variable and, thus, is irrelevant beyond its effect on the standard errors).

Are learning models consistent with the data on executive turnover? This is a difficult question to answer, although Taylor (2010) offers evidence that the models could well be consistent with the data. He does so by calibrating a learning model and measuring the extent to which it fits various aspects of the data. In particular, Taylor’s model does a good job of jointly and quantitatively matching the data (see, in particular, Figure 4 of his article). *Inter alia*, Cornelli et al. (2013) also find evidence consistent with a learning model of CEO turnover.

An additional difficulty for empirical work on this topic is disentangling learning-based rationales for dismissal from other rationales. For instance, a prediction of some dynamic moral-hazard models in which the agent enjoys limited-liability protection (*i.e.*, cannot be fined for poor performance) is that the principal will wish to commit to terminate the agent after sufficiently bad performance as a deterrence to shirking *ex ante*.¹⁵ It is not clear how one would

¹⁴Parrino uses a three-step process to identify forced departures. First, all successions for which the *Wall Street Journal* effectively reports that the CEO has been fired are classified as forced. Second, if the CEO is less than 60 and the *Wall Street Journal* fails to report the reason for the departure is death, poor health, or taking another position or if it reports that the CEO is retiring but the announcement is not at least six months before the succession, then this too is classified as forced. Third, the remaining cases are classified as *voluntary* if, as reported in the press, the incumbent takes a comparable position elsewhere or departs for previously undisclosed personal or business reasons that are unrelated to the firm’s activities; otherwise they are classified as involuntary.

¹⁵See Edmans and Gabaix (in press), especially Section 4, for more on such models.

design a test that could distinguish between moral hazard and learning-based models empirically.

One might, at first, think to look at the stock market reaction to a dismissal as a means of distinguishing the two because moral hazard and learning models would seem to have different implications for firm value. Dismissal in a pure moral-hazard model is *ex post* inefficient: all CEOs are identical, so firm value is reduced by the costs of transition (this is why, for dismissal to be a credible threat and, thus, its possibility to create incentives, the board must *commit ex ante* to dismissal should certain outcomes occur). In contrast, in a learning model, the expected value of a replacement CEO to the firm less transition costs exceeds the expected value of the incumbent. Naïvely, then, one might expect the market to respond positively to dismissal if learning explains dismissal and to respond negatively if moral hazard does. This view is, however, potentially incomplete: if there was uncertainty in the market about whether the board was committed to dismissing the CEO for poor performance, then dismissal would be a positive signal about that commitment and, thus, the board's ability to provide strong incentives. Conversely, the information that causes the board to update negatively about the CEO's ability could be information private to it; hence, his dismissal causes the market to revise downward its assessment of the firm's value (since the initial value was calculated under the belief that the incumbent CEO was better than any replacement; a belief that has been revealed to be false).¹⁶ A reasonable summary then is that while one cannot rule out other explanations, one equally cannot claim that learning models are in any way inconsistent with the data on CEO turnover.

3 CAREER CONCERNS

An important way in which assessment affects executives is by influencing their market value. A successful manager is, all else equal, likely to have greater ability than a less successful one. Correspondingly, her services would be in greater demand than her less-successful counterpart, which means her compensation should be greater.

Using the previous section's setup, this point is readily formalized: suppose the signal at time t is the executive's contribution to firm profit (gross of her compensation). That is, s_t is the increment to period t 's profit attributable to the executive. Assume $s_t = \alpha + \varepsilon_t$, where, similar to above, the prior distribution of α is $N(\hat{\alpha}_0, 1/\tau_0)$ and ε_t is i.i.d. $N(0, 1/\eta)$. Given this structure, the executive's expected contribution to period- $t + 1$ profit is $\hat{\alpha}_t$, the value of which is given by (4). If there is a competitive market for the executive's services and her contribution to profits the same across firms, then her pay in period $t + 1$ will be $\hat{\alpha}_t$. From (4), greater past success—higher average s —means greater pay.

A shortcoming of that simple formulation is it misses the possibility that the executive, knowing her compensation is tied to her performance, may undertake

¹⁶See Hermalin and Weisbach (1998) for a more detailed discussion of possible stock-market reactions to dismissals in a learning model.

activities to influence her *perceived* performance. As will be seen, those activities may be to the benefit or detriment of her employer(s).

3.1 CAREER CONCERNS: THEORY

Fama (1980) made the following argument: given a competitive market for managerial talent, the more able a manager appears, the greater will be her compensation.¹⁷ Estimated ability is a function of performance, but now suppose, unlike the analysis to this point, performance is also a function of the manager's actions as well as her innate ability. If so, then because of their effect on the market's inference about her ability and, thus, on her compensation, the manager has incentives to take actions that will cause her firm to perform well. Although compelling, Fama's argument is, as Holmstrom (1999) noted, incomplete and, in some ways, even incorrect.

Holmstrom formalizes the issues by supposing that time- t performance is

$$s_t = e_t + \alpha + \varepsilon_t, \quad (13)$$

where α and ε_t are as above, with the same statistical properties, and $e_t \in \mathbb{R}_+$ is the executive's action (effort) in period t .¹⁸ Holmstrom assumes the executive's utility in period t is $w_t - c(e_t)$, where w_t is her compensation that period and $c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a twice differentiable increasing function. To ensure unique interior maxima, assume that $c'(0) = 0$ and c is strictly convex (for future reference, the latter assumption entails first-order conditions are sufficient as well as necessary). Consistent with the usual notion of cost, $c(0) = 0$.

Assume the executive's action each period is a *hidden action*; that is, known to her, but not observable by anyone else. Were it, instead, observable, then the problem would be of little interest because the parties could contract on her action directly. Although no one but the executive knows her action in period t , e_t , there is a level of effort, \hat{e}_t , that interested parties (*i.e.*, current and potential employers) anticipate she will take. This means that the other interested parties translate performance, s_t , in period t into a signal of ability by subtracting \hat{e}_t from s_t ; call this constructed signal \tilde{s}_t and observe

$$\tilde{s}_t \equiv s_t - \hat{e}_t = \alpha + \varepsilon_t + e_t - \hat{e}_t.$$

¹⁷This is true even if the manager stays with her current employer: to keep her, the current employer must increase her compensation in response to the outside offers her good performance elicits.

¹⁸Conditional on a linear specification such as (13), the assumption of a uni-dimensional action is without loss of generality. The alternative, that

$$s_t = \gamma(\mathbf{e}_t) + \alpha + \varepsilon_t,$$

where $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}$, is effectively equivalent to (13): simply define $e_t = \gamma(\mathbf{e}_t)$. In terms of her mix of actions, the executive's choice must minimize her disutility of action, $C(\mathbf{e}_t)$, subject to her desired level of $\gamma(\mathbf{e}_t)$; hence, the soon-to-be-defined $c(e_t)$ can simply be seen as

$$c(e_t) = \operatorname{argmin}_{\mathbf{e}_t} C(\mathbf{e}_t) \text{ subject to } e_t = \gamma(\mathbf{e}_t).$$

It follows from (4) and (5) that

$$\hat{\alpha}_t = \frac{\tau_0 \hat{\alpha}_0 + \eta \sum_{i=1}^t (\alpha + \varepsilon_i + e_i - \hat{e}_i)}{\tau_0 + t\eta} = \frac{\tau_{t-1} \hat{\alpha}_{t-1} + \eta(\alpha + \varepsilon_t + e_t - \hat{e}_t)}{\tau_{t-1} + \eta}. \quad (14)$$

If the executive plays a pure strategy in equilibrium, then, given players correctly anticipate others' strategies in equilibrium, $\hat{e}_t = e_t$ along the equilibrium path. Of course, part of determining the equilibrium is verifying that, for any purported equilibrium sequence of actions, $\{\hat{e}_t\}_{t=1}^T$, *anticipated* by others, the executive's best response is actually to play the matching actions (*i.e.*, so that $e_t = \hat{e}_t$ is, indeed, a best response to what others expect). Given *accurate* anticipation, observe, from (14), that the executive's actions don't actually influence estimates of her ability.

The expected value of the executive in period t is $\hat{\alpha}_{t-1} + \hat{e}_t$, which is the expected value of s_t . Given a competitive market for her services, that will thus be her compensation at time t . Formally:

$$w_t = \hat{\alpha}_{t-1} + \hat{e}_t. \quad (15)$$

Critically, notice that w_t does *not* depend on the executive's *actual* time- t action.

To begin the analysis, assume the executive's career lasts T periods, T fixed and commonly known. Assume her discount factor from one period to the next is $\delta \in (0, 1)$ (*i.e.*, she values a payoff of u one period hence at δu now).

A consequence of a finite career span is that the game can be solved via backward induction. Consider period T . The executive's utility in that terminal period is $w_T - c(e_T)$. The effort that maximizes that expression is clearly 0; that is, $e_T = 0$. Consequently, in equilibrium, $\hat{e}_T = 0$. From (15), $w_T = \hat{\alpha}_{T-1}$, and, as just noted, it is also the executive's period T utility. Fold back to period $T - 1$, the executive's remaining lifetime expected utility is

$$\delta \mathbb{E}w_T + w_{T-1} - c(e_{T-1}). \quad (16)$$

To determine the executive's action in period $T - 1$, we need to determine $\mathbb{E}w_T$ and how it depends on e_{T-1} . Utilizing (14),

$$\begin{aligned} \mathbb{E}w_T &= \frac{\tau_{T-2} \hat{\alpha}_{T-2} + (e_{T-1} - \hat{e}_{T-1})\eta + \eta \mathbb{E}\{\alpha + \varepsilon_{T-1}\}}{\tau_{T-2} + \eta} \\ &= \hat{\alpha}_{T-2} + \frac{\eta}{\tau_{T-2} + \eta} (e_{T-1} - \hat{e}_{T-1}), \end{aligned}$$

where the second equality follows because the expectation of α after $T-2$ periods is $\hat{\alpha}_{T-2}$ and the expectation of ε_t is zero for all t . Substituting that value for $\mathbb{E}w_T$ back into (16), it follows that the executive's unique best response is the solution to the first-order condition

$$\delta \frac{\eta}{\tau_{T-2} + \eta} - c'(e_{T-1}) = 0. \quad (17)$$

Observe the solution is independent of both the conditional expectation of her ability, $\hat{\alpha}_{T-2}$, and the effort others anticipate, \hat{e}_{T-1} . It follows that, in equilibrium, \hat{e}_{T-1} equals the e_{T-1} that solves (17); call that latter value e_{T-1}^* .

Two observations about the results to this point:

1. Welfare is $e - c(e)$; so, the welfare-maximizing effort, e^{FB} (FB for first best) solves $1 - c'(e) = 0$. Clearly, $e^{\text{FB}} > 0$, so the executive's action is less than first best in the last period of her career. Because $\delta\eta/(\tau + \eta) < 1$, the usual comparative statics entail that $e^{\text{FB}} > e_{T-1}^*$: the executive's action is also less than first best in the penultimate period of her career. Hence, at least in the twilight of an executive's career, career concerns are insufficient to induce optimal actions.
2. From (17), the executive's private marginal return is $\delta\eta/(\tau_{T-2} + \eta)$, which increases in δ and η , but falls in τ_{T-2} . Consequently, the executive takes greater action (closer to optimal) the more patient she is and the greater the precision of the signal; she takes a lesser action (farther from the optimum) the more precise the current estimate of her ability.

Let \mathbb{E}_t denote expectation given information at the *beginning* of period t (*i.e.*, just after $t - 1$ periods). The executive's expected remaining lifetime utility from period t on is

$$\mathbb{E}_t \left\{ \sum_{j=t}^T \delta^{j-1} (w_j - c(e_j)) \right\}$$

For an arbitrary period $t' > t$, the same logic as above yields

$$\begin{aligned} \mathbb{E}_t w_{t'} &= \frac{\tau_{t-1} \hat{\alpha}_{t-1} + \eta \sum_{j=t}^{t'-1} (e_j - \hat{e}_j + \mathbb{E}_t \{\alpha + \varepsilon_j\})}{\tau_{t-1} + (t' - t)\eta} \\ &= \hat{\alpha}_{t-1} + \frac{\eta \sum_{j=t}^{t'-1} (e_j - \hat{e}_j)}{\tau_{t-1} + (t' - t)\eta}. \end{aligned}$$

It follows, therefore, that at period $t < T$, when the executive is choosing her action for that period, her optimization program is equivalent to solving

$$\max_{e_t} -c(e_t) + e_t \sum_{j=1}^{T-t} \delta^j \frac{\eta}{\tau_0 + (t+j-1)\eta} \equiv \max_{e_t} m_t e_t - c(e_t), \quad (18)$$

where m_t , the sum given in (18), is the executive's lifetime marginal return to her effort in period t . The properties of this marginal return are summarized as follows:

Lemma 2. *The executive's lifetime marginal return to effort in period t (*i.e.*, the m_t defined in (18)) is (i) decreasing in t ; (ii) decreasing in the precision of the prior estimate of ability, τ_0 ; (iii) increasing in the precision of the signal, η ; and (iv) increasing with the discount factor, δ .*

The value of m_t determines the executive's unique pure-strategy best response, e_t^* , the solution to the first-order condition $m_t - c'(e) = 0$. Observe m_t is deterministic (it depends on the underlying parameters, not the realization of random variables). The equilibrium is, thus, characterized by $\hat{e}_t = e_t = c'^{-1}(m_t)$.

On the equilibrium path, the executive's actions are perfectly anticipated. In a sense, then, they play no role in employers' (current and potential) estimate of her ability. One might, therefore, ask why she, then, bothers to take actions motivated by influencing the market's beliefs if the market will see through that? The answer is the market subtracts the actions it *anticipates* she will take—recall it does not see what she actually chooses to do—so if she chose not to take those expected actions, then her performance would be worse and the market would subtract its anticipation of her action from that worse performance; her inferred ability would, therefore, be correspondingly low. In other words, she must match what the market anticipates to avoid being judged harshly by it. The manager is like the Red Queen in Lewis Carroll's *Through the Looking Glass*: she must run as fast as possible just to stay still.¹⁹

The properties of $\{m_t\}_{t=1}^T$ set forth in Lemma 2 determine the time path of the executive's actions and how it varies with the underlying parameters. In particular, because $\{m_t\}_{t=1}^T$ is a decreasing sequence, it follows that $\{e_t^*\}_{t=1}^T$ is likewise a decreasing sequence: the actions (efforts) of the executive decline over the course of her career. As seen earlier, they ultimately go to zero and, in the twilight of her career, are below the first best. It is, however, possible that her actions could be above the first best early in her career: if, for instance, $T = 10$, $\delta = 9/10$, and $\tau_0 = \eta = 1$, then $m_1 \approx 1.35 > 1$ —the executive's private marginal return to action in period 1 exceeds the social marginal return.

There are two effects driving the decline in career-based incentives over the executive's life: first, the closer she gets to the end of her career, the fewer periods that her current action affects, which reduces the return to that action; second, as her career progresses she becomes more of a known entity, the weight attached each period's performance in assessing her ability decreases over time, and, thus, the scope for her action to influence assessment also decreases. This last effect can be seen by observing that

$$\hat{\alpha}_t = \frac{\tau_{t-1}}{\tau_{t-1} + \eta} \hat{\alpha}_{t-1} + \frac{\eta}{\tau_{t-1} + \eta} (e_t - \hat{e}_t + \alpha + \varepsilon_t).$$

The first fraction—the weight on the prior going into period t —increases over time because τ_t increases in t , while the second fraction—the weight given current performance—decreases over time for the same reason.

The results in Lemma 2 also entail that the actions of the executive are greater, *ceteris paribus*, the more precise the signal and they are less, *ceteris paribus*, the more precise is the initial prior estimate, $\hat{\alpha}_0$. As a summary:

¹⁹Holmstrom (1999) refers to the *Red Queen effect* as a form of rat racing. Unlike conventional rat-racing models, such as tournaments, the agent in question is not competing with other agents but rather, in essence, with what the market anticipates.

Proposition 1. *In the career-concerns model of this section, there is a pure-strategy subgame-perfect equilibrium in which the executive's strategy is the sequence of actions $\{e_t^*\}_{t=1}^T$ where $e_t^* = c'^{-1}(m_t)$, m_t defined by (18). The market (employers) accurately anticipate the executive's actions. The executive's actions decline over time; although they may initially exceed the first-best level, they ultimately are less than first best and her action is zero in her last period of work. Ceteris paribus, the executive's actions are greater the more precise a signal performance is of her ability and they are less the more precise the initial estimate of her ability is.*

An Infinite Horizon. The model was constructed assuming a known and finite career length for the executive. Such an assumption is not necessary: as an alternative, assume the length of her career is unknown, but that there is a constant probability of it ending for exogenous reasons after each period. Hence, her career has finite length with probability one. As is well known, the probability of her surviving to the next period can be incorporated into the discount factor (*i.e.*, δ reflects both financial and survival discounting). In this variant of the model, one has

$$m_t = \sum_{j=1}^{\infty} \delta^j \frac{\eta}{\tau_0 + (t+j-1)\eta}.$$

It is readily verified that the results stated in Proposition 1 continue to hold. Note, though, that the diminishing-action-over-time result is due, in an infinite horizon model, solely to learning; that is, only the second of the two reasons for a decline in career-based incentives over the executive's life applies. Finally, from Lemma 2(iv), the executive's actions are greater (less) *ceteris paribus* the greater (less) is the exogenous rate of survival from one period to the next; that is, executives who, for exogenous reasons, are more committed to the managerial labor market can be expected to supply greater action *ceteris paribus*.

Ability Follows a Stochastic Process. In Holmstrom (1999), an extension is considered in which executive ability varies over time according to a stochastic process (a Gaussian random walk). In such an environment, the relevance of past performance at time t for predicting future performance at t' diminishes as the difference $t' - t$ grows. Consequently, if the executive's horizon is infinite (as in the last paragraph), there may be no decline in career-based incentives over time: recent performance can be sufficiently important to forecasting future performance that career incentives remain high—indeed, they can remain constant in the limit under some formulations. See Holmstrom (1999) for details.

Contractual Solutions. Although career concerns generate incentives for the executive, they do not—as seen above—generate *optimal* incentives for the executive generically. For the above model, the shareholders could improve matters—at least late in the executive's career—by providing financial incentives

tied to performance. In particular, for the sake of concreteness, suppose that the executive's career is two periods in length (*i.e.*, $T = 2$). It was earlier noted that her action in period 2 is zero (*i.e.*, $e_2^* = 0$) and her action in period 1 is the solution to (17); that is,

$$\delta \underbrace{\frac{\eta}{\tau_0 + \eta}}_{m_1} - c'(e_1^*) = 0. \quad (17')$$

As noted earlier, her marginal return to her action, m_1 , is less than the social marginal return, 1. Assuming the shareholders can contract on performance, s_t , then they increase welfare by offering the executive compensation tied to that performance. For instance, in the second period, a contract that paid the executive s_2 would induce first-best action. The executive would capture all of her contribution to profit, but this is consistent with the underlying presumption that there is a perfectly competitive market for her services. Given such an incentive contract, the market no longer cares about her perceived ability. Hence, in the first period, the executive would have no career concerns. Similar logic dictates, therefore, that she would be offered a contract in the first period in which her compensation equals her contribution to profit, s_1 .

The last paragraph does not invalidate career concerns. First, the ability of incentive contracts to deliver the first-best outcome depends on the executive's being willing to absorb all the risk associated with those contracts and, moreover, a means of compelling her *to pay* her employer when $s_t < 0$. Absent such assumptions, incentive contracts are typically incapable of delivering the first best.²⁰ Moreover, when the executive's compensation is not 100% of her contribution to profit, then contracts offered her in the future can depend on her past performance, which reintroduces career concerns. Second, it could be less than straightforward to write a contract directly contingent on the signal. This could be because the signal, although observed by the market, is not readily measurable in a verifiable way.²¹ This would be true if only market experts (but not judges, juries, or others who would enforce a formal contract) can infer ability from signals (by way of analogy, presumably scouts for professional sports teams can assess ability in ways that non-experts cannot).²² Third, the signal could be *unobservable* to the market and accordingly unverifiable. If,

²⁰See Holmstrom (1979) and Grossman and Hart (1983) for what goes wrong when the executive (agent) is risk averse and Sappington (1983) for what goes wrong when the executive cannot be made to pay her employer.

²¹It is worth noting, however, that the observable-vs.-verifiable distinction is not without controversy; see Hermalin et al. (2007, especially §4) for a discussion of the issues and a survey of the relevant literature. In particular, as Hermalin and Katz (1991) and Maskin and Tirole (1999) show, it is often possible to write contracts in such a way that it is as if they were contingent on observable but unverifiable signals.

²²A possible counter argument is that these very experts could be hired to testify should contractual disputes arise. On the other hand, experts hired for an adversarial process could easily fail to provide accurate or unbiased reports (ironically, there is a literature on the career—reputational—concerns of experts; see Wolinsky, 1993, and subsequent work).

however, the board and CEO can observe it and the CEO's compensation is set according to a bargaining game in which the CEO has some bargaining power (*e.g.*, as in Hermalin and Weisbach, 1998), then the logic of career concerns would still apply.²³

3.2 CAREER CONCERNS: EMPIRICS

Although corporate governance was the impetus for the initial career-concerns literature (Fama, 1980, and Holmstrom, 1999), the subject is much more general: inside of almost any organization the possibility of advancement is what motivates most employees. Standard advice given to young professionals (and others) is that they are always being evaluated, so they should always seek to make a good impression on everyone they meet. Regardless of any explicit pay-for-performance incentives facing employees, they nonetheless wish to persuade others they are high ability. Doing a good job will enhance one's career prospects, and consequently, regardless of any explicit incentive plans, career concerns will provide a reason for employees to do a good job.

Such career-based incentives are crucial to the functioning of most organizations. Unfortunately, they are particularly difficult to measure directly. Consider the marginal return to effort, m_t , defined by (18) above. To know it precisely, one would need to know the relevant discount factor, δ ; the date at which the employee's career ends, T ; the variance of ability, $1/\tau_0$; and the variance of the random shocks to the performance measure, $1/\eta$. Moreover, even if m_t can be estimated, there is no way to measure effort. All of this renders any direct empirical test of career-concern models impossible. Instead, one is obliged to consider secondary hypotheses of the model and test them.

One such secondary hypothesis is that *fixed* components of compensation today, such as salary, should be positively correlated with *past* performance. Under standard (*i.e.*, hidden-action) agency theory, there should be no reason for a salary that is determined today to depend on previous performance.²⁴ In contrast, under a career-concerns model, present compensation should be strongly determined by past performance. Early work exploring that possibility—specifically, Murphy (1986) (but see also Murphy, 1985, and Jensen and Murphy, 1990)—finds evidence of such intertemporal links between performance and compensation. The rest of this section considers other empirical tests.

3.2.1 Evidence from Intertemporal Pattern of Incentives

The career-concerns model's intertemporal implications provide one way to test it; in particular, observe that the incentives due to career concerns diminish over

²³Note if the game were one of *infinite* time, then reputational contracting might serve to replicate formal incentive contracting; see Levin (2003).

²⁴To be precise, this assumes that compensation is determined by formal rather than relational contracting (*i.e.*, this is not a repeated game in which salary today is determined by an implicit promise relating past performance to current salary). It also assumes that there are no wealth effects, due to past contingent pay, that might affect the optimal contract today.

the course of an executive's career. As noted earlier, there are two reasons for this: first, as the executive becomes a known entity, less is inferred from any one period's performance; and, second, the closer she is to retirement, the less time she has to capitalize on her reputation. To offset declining incentives, Gibbons and Murphy (1992) predict that firms will offer those CEOs close to retirement greater *direct* incentives (*i.e.*, performance-contingent pay) than those many years from retirement. Using CEO compensation data drawn from a sample of large US corporations, Gibbons and Murphy test this prediction, finding that the sensitivity of salary and bonus to performance is indeed significantly greater for CEOs at the end of their careers than for CEOs in the midst of them.

Dechow and Sloan (1991) find results complementary to those: to wit, executives at the end of their career are less inclined to make R&D investments than those earlier in their careers. The explanation offered by Dechow and Sloan is that compensation tied to short-run performance measures create perverse incentives for executives to underinvest in projects with long-term payoffs when the executives won't be around to collect any reward from those payoffs later.²⁵ Put slightly differently, consider a positive net-present-value (NPV) project that requires an upfront cost today. An executive at the end of her career has little incentive to pursue such a project if her compensation today is a function of today's profits. But if she has a longer horizon, then she knows tomorrow's pay, which is based on tomorrow's greater profits due to today's investment, will reward her investing. Notice, however, that a career-concerns model would yield a similar pattern: if effort e_{t-2} boosts later signal s_{t-1} , which affects today's compensation w_t , then an executive whose career will last to t will wish to expend effort; whereas one whose career ends at $t - 1$ will not.

There are undoubtedly additional factors affecting the intertemporal pattern of incentives. For instance, a successful executive tends to gain power *vis-à-vis* her board of directors, allowing her to capture greater compensation,²⁶ which provides an alternative (or at least complementary) explanation for the pattern Gibbons and Murphy find. In addition, most performance-based CEO firings occur in the first few years of a CEO's tenure (Allgood and Farrell, 2003). To induce CEOs to accept the risk inherent in such "probation," boards may need to make implicit promises to reward survivors adequately later in their careers.

3.2.2 Estimating Learning about CEOs of Public Corporations

Murphy (1986) addresses the question of which is more empirically relevant: assessing and rewarding CEO ability (the value he adds) or providing him incentives? To do so, he presents both learning and incentive-based models of CEO compensation. The learning and incentive-based models share a number of predictions, but differ in what they predict about how the variance of earnings and the pay-performance relation will vary with experience. Murphy finds empirical

²⁵Some of the models in Section 4 also speak to this finding; see further discussion of Dechow and Sloan (1991) in this context in Section 4.3.

²⁶In this regard, see both Hermalin and Weisbach (1998) and our discussion of it *infra*.

evidence that the relation between a CEO's pay and his performance is strongest early in his tenure. This finding is consistent with the learning model, but not a standard incentive-based one. Early in his career, performance is more informative about ability given that the prior on ability is more diffuse. Therefore, if pay reflects expected ability, it should vary more with performance early in a CEO's tenure, when a given level of performance will lead to larger updates of assessed ability. Consequently, Murphy concludes that assessing and rewarding managerial ability is the primary factor in setting CEO pay.

Boschen and Smith (1995) examine the long-term effect of firm performance on CEO pay. They find that contemporaneous firm performance affects a CEO's income not only today, but in the future. Their estimates indicate that a positive shock to performance raises CEO pay for the following four to five years. Boschen and Smith summarize their findings as "the cumulative response of pay to performance is roughly 10 times that of the contemporaneous response" (p. 577). Although dynamic agency models can yield the prediction that pay today is dependent on performance yesterday (see, *e.g.*, Rogerson, 1985), Boschen and Smith interpret their results as not providing support for such models (see discussion on pages 599–600). Rather, they interpret their results as being more consistent with the career-concerns model. In particular, their finding that the effect of a positive performance shock on compensation eventually fades is consistent with a learning model (consider, *e.g.*, expression (14) above).

Taylor (2013) estimates a structural model of learning about CEO ability. His goal is to evaluate the extent to which CEOs are able to capture rents from this process, as suggested by Bebchuk and Fried (2004). Taylor's model is based on the idea that a CEO and board can enter into an implicit or reputational contract that permits the board to partially insure the CEO against the income risk that assessment imposes on him. His estimates suggest that 50% of any increase in firm value due to favorable news about the CEO is captured by the CEO, whereas the CEO enjoys nearly complete insurance against downward risk. His analysis shows how one might estimate the underlying parameters of a learning model, and use these estimates to evaluate an important question about corporate governance.

3.2.3 Evidence from the Money Management Industry

One industry for which it is possible to measure market-based incentives is the money-management industry. Berk and Green (2004) argue that the market's assessment of a money manager's ability will be reflected by the inflows to her or his fund. To wit, when a fund performs well, the market's assessment of its manager increases and investors correspondingly invest more (presumably, until the perceived abnormal return to managerial skill disappears).

This idea lends itself to empirical analysis. Because the fee structures of most funds are observable and one can estimate the extent to which future inflows respond to current performance, it is possible to calculate the expected fees from incremental flows given a specified performance level. These incremental fees represent the return for having a better managerial reputation insofar as

Table 1: The ratio of indirect to direct incentives and Jensen-Murphy b for different types of funds.²⁷

	Private Equity		Hedge Funds	
Panel A: Averages				
Indirect/Direct	0.63		3.03	
Jensen-Murphy b	<i>0.29</i>		<i>0.66</i>	
Panel B: By Sequence or Age				
	Current Fund Sequence		Fund Age (years)	
	2	3	0	3
Indirect/Direct	0.47	0.32	5.51	3.33
Jensen-Murphy b	<i>0.26</i>	<i>0.23</i>	<i>0.67</i>	<i>0.62</i>
Panel C: By Scalability				
	Buyout	Venture	Not Capacity Constrained	Capacity Constrained
Indirect/Direct	1.24	0.48	3.23	2.53
Jensen-Murphy b	<i>0.43</i>	<i>0.24</i>	<i>0.68</i>	<i>0.60</i>

they reflect the positively updated estimate of the manager's ability. Through this process, one can, therefore, estimate the managers' indirect, market-based, incentives. These incentives are undoubtedly economically important given investors responsiveness to their assessments of managerial ability and the close connection between size of fund and managerial compensation.

Chung et al. (2012) and Lim et al. (2016) perform this exercise for private equity funds and hedge funds, respectively. In each case, the authors estimate the sensitivity of inflows to performance, and calculate the present value of the expected fees that occur with an incremental return. Both private equity and hedge funds have relatively large direct-incentive compensation: 20% profit-sharing plus the additional incentives that derive from management ownership. These papers use this large direct incentive compensation as a benchmark, and also calculate the "indirect Jensen-Murphy b ," the amount managers receive in future income for each additional dollar returned to investors, not including the return on the managers' own investments in the fund (see Jensen and Murphy, 1990).

Table 1 shows that indirect incentives are large even relative to the significant direct incentive compensation in this industry, and that they are a substantial fraction of money returned to investors. For a typical private equity fund, Chung et al. (2012) find that indirect incentives are 63% as large as direct incentives, while Lim et al. (2016) find, for a typical hedge fund, that they are about three times as large.

Moreover, both papers find that the cross-sectional and time-series patterns of indirect incentives are consistent with the assessment framework. In keeping with the theory’s underlying logic, which predicts that assessment effects diminish over a career, indirect incentives are much higher for younger partners than for older ones for both private-equity and hedge funds (see Table 1). The greater uncertainty about younger partnerships’ abilities means the market’s updating about their abilities is more sensitive to performance, hence so too are investment inflows and, thus, future fees.

In addition, fixing ability, funds that rely on *scalable* strategies should be better able than other funds to utilize fresh investments profitably. Hence, inflows will be greater for scalable-strategy funds following good performance, *ceteris paribus*. The market-based incentives for managers of scalable-strategy funds are, correspondingly, stronger. Consistent with this prediction, Chung et al. (2012) find that more scalable buyout funds are associated with higher indirect incentives than less scalable funds. Lim et al. (2016) find a similar result for hedge funds: more scalable strategies have higher indirect incentives than funds with less scalable strategies *ceteris paribus*.

4 SIGNAL JAMMING: PROJECT CHOICE AND MANAGERIAL MYOPIA

In Holmstrom (1999), the executive undertakes an action in hopes of influencing the market’s perception of his ability. As such, the model is among the first examples of a more general economic phenomenon in which an actor has an incentive to influence, through her action, what others infer from public signals. This phenomenon has come to be called *signal jamming*.²⁸ As its somewhat pejorative name suggests, signal jamming can have negative consequences; in this section, we consider the darker side of signal jamming.

4.1 PROJECT CHOICE

Holmstrom (1999) also considers the following model: in each of two periods, the executive chooses a single project p_t from a set of possible projects \mathcal{P}_t , where t again indexes periods. The return, R , to the chosen project is stochastic, distributed normally with a mean of $\alpha + \beta_p + \varepsilon$, where α , drawn from $N(\hat{\alpha}_0, 1/\tau_0)$, again denotes the executive’s ability; β_p is the project-specific expected benefit gross of ability; and $\varepsilon \sim N(0, 1/\eta_p)$, where η_p , the precision, is project specific.

²⁷Jensen-Murphy b (the number in italics) is the present value of the change in managers’ income per dollar returned to investors (other than the managers, in the case of private equity). See Jensen and Murphy (1990) for details. The lefthand columns are from Chung et al. (2012), the righthand from Lim et al. (2016). For details on the private equity numbers, see Tables 5 and 6 of Chung et al. For private equity, the maximum number of potential future funds the general manager (GP) could run is three. For details on the hedge fund numbers, see Tables 4–6 in Lim et al.

²⁸Fudenberg and Tirole (1986) appear to have introduced the term “signal jamming” into economics for the purpose of describing this phenomenon.

It is assumed that the executive and all would-be employers are symmetrically informed about all payoff-relevant variables at the start of the executive's career. In particular, as in the earlier model, the executive initially has no better knowledge of his ability than the market. The sets \mathcal{P}_1 and \mathcal{P}_2 are common knowledge, as are the parameters, (β_p, η_p) , associated with each $p \in \mathcal{P}_1$ and \mathcal{P}_2 .

Assume, other than their potential consequences for future compensation—to be detailed shortly—the executive has no inherent preferences over projects.

As before, there is a competitive market for the executive's services. Consider the second period: because there are no career concerns in the last period, there is no reason for the executive not to choose the project in \mathcal{P}_2 with the greatest expected return, β^* . Competition for the executive's services means his second-period compensation is bid up to expected second-period return, $\beta^* + \hat{\alpha}_1$. Because the ensuing analysis does not depend on the value of β^* , we are free to set it to zero for convenience; so, second-period compensation is just $\hat{\alpha}_1$.

Given the project, p , observers anticipate (or perhaps observe) the executive to take, the realized return from the first-period project, R_1 , is equivalent to the signal $s_1 = R_1 - \beta_p$. Building on previous analysis (in particular (4) and (5) above), it follows that the *posterior* distribution of ability is normal with mean

$$\hat{\alpha}_1 = \frac{\tau_0 \hat{\alpha}_0 + \eta_p (R_1 - \beta_p)}{\tau_0 + \eta_p} \quad (19)$$

and precision $\tau_0 + \eta_p$. In what follows, we also need to know the *ex ante* distribution of $\hat{\alpha}_1$ (*i.e.*, before the realization of R_1). Following an analysis similar to that underlying (7) above, s_1 is the sum of two normal random variables, α and ε_1 . As such, it too is distributed normally, with a mean equal to the sum of their means (*i.e.*, $\hat{\alpha}_0$) and a variance equal to the sum of their variances:

$$\frac{1}{\tau_0} + \frac{1}{\eta_p} = \frac{\tau_0 + \eta_p}{\tau_0 \eta_p} = \text{Var}(s_1). \quad (20)$$

Viewed from an *ex ante* perspective, $\hat{\alpha}_1$ is an affine transformation of the random variable s_1 ; hence, it is distributed normally with a mean of

$$\mathbb{E} \frac{\tau_0 \hat{\alpha}_0 + \eta_p s_1}{\tau_0 + \eta_p} = \frac{\tau_0 \hat{\alpha}_0 + \eta_p \mathbb{E} s_1}{\tau_0 + \eta_p} = \hat{\alpha}_0$$

and a variance of

$$\left(\frac{\eta_p}{\tau_0 + \eta_p} \right)^2 \times \text{Var}(s_1) = \frac{\eta_p}{(\tau_0 + \eta_p) \tau_0}. \quad (21)$$

Unlike the first Holmstrom (1999) model, assume now that the executive is risk averse: his utility in period t is $u(y_t)$, where y_t is his period- t compensation and $u : \mathbb{R} \rightarrow \mathbb{R}$ is strictly increasing and strictly concave. Assuming, as we do, that the executive's only motivation is his expected utility from his market compensation in period 2, the executive's choice of first-period project maximizes $\mathbb{E}u(\hat{\alpha}_1)$. In thinking about his choice, note that a payoff distributed $N(\mu, \sigma_a^2)$

dominates one distributed $N(\mu, \sigma_b^2)$ in the sense of second-order stochastic dominance if $\sigma_b > \sigma_a$;²⁹ that is, fixing the mean, a risk-averse decision maker prefers a draw from a normal distribution with a lower variance than from one with a higher variance. Because of his risk aversion, both the mean and variance of $\hat{\alpha}_1$ matter to him and will, thus, affect his choice of project.

As a first pass, consider a variant of Holmstrom (1999) due to Hermalin (1993): assume the market observes the executive's choice of first-period project; that is, it knows β_p and η_p .³⁰ There is thus no scope for the executive to try to fool the market by choosing a project with a higher project-specific expected benefit than it anticipates. Consequently, the executive cares solely about the risk attributes of the project he chooses. In particular, he wishes to choose among the available projects the one that minimizes (21). It is readily seen that (21) is an increasing function of η_p , which means the executive prefers the project with the least precision, equivalently greatest variance. To summarize:

Proposition 2. *When the market can observe the mean and variance of the project chosen by the executive, the executive will wish to choose the noisiest (highest variance) project ceteris paribus.*

It might at first seem counter-intuitive that a risk-averse executive with career concerns would choose the highest-variance project. The rationale behind this result is that the noisier the signal, the less weight observers assign it in updating their beliefs about the executive. The less weight assigned the signal, the more assigned the prior, which is non-stochastic. Because the expected posterior is always the same, $\hat{\alpha}_0$, assigning weight to the signal is simply adding risk. Admittedly, the risk is also a function of how noisy the signal is; that is, were the weight assigned the signal invariant, then the executive would like less noisy signals. What (21) shows is that the less-weight effect dominates the greater *per se* noise effect.

Because, as Proposition 2 concludes, the executive cares only about about how risky projects are, there is no reason to expect him to choose a project with a high project-specific expected benefit. In other words, an agency problem arises whenever \mathcal{P}_1 contains projects p and p' such that $\beta_p > \beta_{p'}$ and $\eta_p > \eta_{p'}$: the first-period employer would prefer the executive choose the former project, but the executive prefers the latter.

In the original Holmstrom (1999) model, it is assumed that the market *cannot* observe the project chosen (equivalently, does not directly observe its

²⁹See Lemma A.2 in the appendix

³⁰A subtlety is why doesn't the executive's first-period employer write a contract fixing project choice if it's observable? One possibility is that although observers can determine the β_p and η_p of a chosen project, it is not feasible to specify projects with sufficient clarity to serve as contractual contingency (in this regard, see also footnote 21 *supra* and connected discussion). Alternatively, as will be seen, nothing in Hermalin's analysis requires that the set \mathcal{P}_1 be known *ex ante*; hence, it could be infeasible to contract on a specific project or even parameters (β_p, η_p) because whether such project or parameters exist is unknown and no one might know which projects the executive chose *not* to pursue (so, *e.g.*, a contract calling for him to choose the project with the greatest β_p in \mathcal{P}_1 is infeasible).

characteristics, (β_p, η_p)). Let $\hat{p} \in \mathcal{P}_1$ denote the project the market anticipates the executive will take and p the project he actually chooses (of course, in equilibrium, they must be the same). In terms of the analysis above, replace the p s in expression (19) with \hat{p} s; hence,

$$\mathbb{E}\hat{\alpha}_1 = \hat{\alpha}_0 + \frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}}(\beta_p - \beta_{\hat{p}}). \quad (22)$$

The variance of s_1 depends on the project actually chosen; so $\text{Var}(s_1)$ remains as defined in (20). The equivalent of expression (21), the variance of $\hat{\alpha}_1$, is

$$\left(\frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}}\right)^2 \times \text{Var}(s_1) = \left(\frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}}\right)^2 \frac{(\tau_0 + \eta_p)}{\tau_0 \eta_p}. \quad (23)$$

If project-specific expected return were the same for all projects, then the executive would care about the noisiness of the projects only; that is, he would choose the project that minimized (23). It is readily seen that corresponds to the project with the greatest precision (smallest variance). This is opposite to the conclusion reached in Proposition 2: when observers don't directly observe the parameters of the chosen project, the executive cannot affect the weight they assign to the signal versus the prior and, thus, the only relevant effect is the noise of the signal itself. To summarize:

Proposition 3. *When the market cannot observe the variance of the project chosen by the executive and all projects have the same project-specific expected return (i.e., β_p a constant for $p \in \mathcal{P}_1$), the executive will wish to choose the least noisy (lowest variance) project ceteris paribus.*

If project-specific expected return is the same across projects, then the executive's risk preferences don't matter to his employer. But if those returns aren't the same, then an agency problem can arise: the executive's concern about risk can lead him to choose a project other than the one that maximizes β_p . To illustrate this possibility, suppose it is common knowledge that the executive's utility function exhibits constant absolute risk aversion (is CARA), with known coefficient of absolute risk aversion ζ . It follows that his choice of project will maximize³¹

$$\mathbb{E}\hat{\alpha}_1 - \frac{1}{2}\zeta \text{Var}(\hat{\alpha}_1) = \hat{\alpha}_0 + \frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}}(\beta_p - \beta_{\hat{p}}) - \frac{1}{2}\zeta \left(\frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}}\right)^2 \frac{(\tau_0 + \eta_p)}{\tau_0 \eta_p}. \quad (24)$$

³¹Expression (24) is derived as follows: the assumption that utility is CARA entails that $u(y) = K - \exp(-\zeta y)$, where K is an irrelevant constant. It follows that $\mathbb{E}\{u(y)\}$ is K less the moment-generating function (with auxiliary variable $-\zeta$). For a $N(\mu, \sigma^2)$ random variate the moment-generating function with auxiliary variable q is $\exp(q\mu + q^2\sigma^2/2)$; hence,

$$\mathbb{E}\{u(y)\} = K - \exp\left(-\zeta\mu + \frac{1}{2}\zeta^2\sigma^2\right).$$

Optimization programs are invariant with respect to positive monotonic transformations (here, the transformation is $-\log(K - \mathbb{E}\{u(y)\})/\zeta$).

Suppose, to slightly abuse notation, that

$$\mathcal{P}_1 = \{(\beta, \eta) \in \mathbb{R}_+^2 \mid \eta \geq \bar{\eta} - \beta \text{ and } 0 \leq \beta \leq \bar{\beta}\}, \quad (25)$$

where $\bar{\beta}$ and $\bar{\eta}$ are constants such that

$$\bar{\eta}^2 + \tau_0 \bar{\eta} > \frac{\zeta}{2} > (\bar{\eta} - \bar{\beta})^2 + \tau_0(\bar{\eta} - \bar{\beta}) > 0. \quad (26)$$

Given Proposition 3, the only relevant portion of \mathcal{P}_1 is the frontier

$$\{(\beta, \eta) \in \mathbb{R}_+^2 \mid \eta = \bar{\eta} - \beta \text{ and } 0 \leq \beta \leq \bar{\beta}\}.$$

Observe the frontier displays a risk-return tradeoff—on the frontier the greater the project-specific expected return, the greater the variance (the noisier) is the return. Maximizing (24) is, thus, equivalent to solving

$$\max_{\beta \in [0, \bar{\beta}]} \beta - \frac{1}{2} \zeta \frac{\eta_{\hat{p}}}{\tau_0 + \eta_{\hat{p}}} \frac{(\tau_0 + \bar{\eta} - \beta)}{(\bar{\eta} - \beta)\tau_0}.$$

Standard methods reveal the solution is³²

$$\beta^* = \bar{\eta} - \sqrt{\frac{\zeta \eta_{\hat{p}}}{2(\tau_0 + \eta_{\hat{p}})}}. \quad (27)$$

In equilibrium, the market must correctly anticipate the executive's choice. Noting that the frontier can be reexpressed as $\beta = \bar{\eta} - \eta$, the condition for equilibrium then follows from (27):

$$\eta = \sqrt{\frac{\zeta \eta}{2(\tau_0 + \eta)}}.$$

The solution is

$$\eta_e = \frac{1}{2} \left(-\tau_0 + \sqrt{\tau_0^2 + 2\zeta} \right).$$

Given (26), it is readily shown that η_e is in bounds (*i.e.*, $\eta_e \in (\bar{\eta} - \bar{\beta}, \bar{\eta})$). Straightforward calculations reveal that η_e increases in ζ and decreases in τ_0 . In words, the more risk averse the executive, the lower the variance (and hence expected return) of the equilibrium project; and the greater the precision of the prior estimate of his ability, the greater the variance (and hence expected return) of the equilibrium project. To summarize:

Proposition 4. *Suppose the market cannot observe the executive's choice of project. Assume the executive has a CARA utility function and the set of possible projects is described by expressions (25) and (26). Then the executive does not*

³²Because $\eta_{\hat{p}} \in [\bar{\eta} - \bar{\beta}, \bar{\eta}]$, it is readily shown that $\beta^* \in (0, \bar{\beta})$; that is, it is an inbounds interior solution. Second-order conditions are readily verified.

choose the project with the highest project-specific expected return (i.e., he does not choose β) in equilibrium. Furthermore, the expected return of the project he does choose is greater in equilibrium the less risk averse he is or the greater is the precision of the prior estimate of his ability. Correspondingly, the variance of returns from the chosen project is greater in equilibrium the less risk averse is the executive or the greater is the precision of the prior estimate of his ability.

Intuitively, the less risk averse is the executive (the lower is ζ), the more willing he is accept greater risk for a higher expected wage. Although still a victim of the Red Queen effect, the cost to the executive of reducing β below the anticipated level to reduce risk is that he reduces his expected wage; as noted, he is less willing to do this the less risk averse he is, which means the equilibrium β will be greater. The greater the precision with which the prior is estimated, the less weight is placed on actual performance when assessing the executive's ability. Hence, the riskiness of the project matters less, which means he is more willing to accept risk in hopes of inducing a higher assessment of his ability. Again, the Red Queen effect means he fools no one in equilibrium, but because of the greater temptation, the equilibrium β will be greater.

Summary. The analysis presented in this section shows that career concerns *can generate* an agency problem:³³ the executive (agent) is induced to care about the riskiness of the projects, something that is directly immaterial to the shareholders (principal). Were there no career concerns (*e.g.*, the executive lived just one period), then the agency problem would disappear: the executive would have no reason not to choose the project with the greatest project-specific expected return. Ironically, the agency problem could be worse when the executive's choice of project is observable: he then places no weight on expected return, caring only about risk (preferring the riskiest project).³⁴ When the parameters of projects are unobservable to the market, then the Red Queen effect (signal jamming) causes the executive to care somewhat about the expected return of the project he chooses. Nonetheless the agency problem remains and is worse the more risk averse is the executive or the more uncertainty there is about him (the lower the precision of the prior estimate of ability).

4.1.1 When No Project May be Available

Implicit in the analysis to this point is that the executive must choose a project—he cannot claim there is no project to pursue. Suppose, instead, it is plausible for him to make such a claim; that is, it is uncertain what \mathcal{P}_1 may be and there is positive probability it is the empty set (*i.e.*, that $\mathcal{P}_1 = \emptyset$).

Given the above framework, if no project is a plausible claim, then absent any other incentive, the executive will always claim no project existed: if there

³³Section 2.2.2 of Stein (2003) offers a survey of agency problems caused by career concerns.

³⁴Of course, no agency problem would exist if \mathcal{P}_1 is defined by (25) and (26), because the riskiest project would then be the one with the greatest expected return.

is no project, there is no updating beliefs about his ability and his next-period compensation will be $\widehat{\alpha}_0$ with certainty; whereas, no matter what project he pursues (and regardless of the observability or not of its parameters), his next-period compensation will be a random variable with expectation $\widehat{\alpha}_0$. Given his risk aversion, he strictly prefers a sure thing that has a payoff equal to the expectation of a random variable. To summarize:

Proposition 5. *If the career-concerns model set forth above is extended to make it plausible for the executive to claim no project is available (or to claim all available projects are inferior to not undertaking a project), then the executive never undertakes a project in the first period.³⁵*

Proposition 5 indicates that a career concerns can lead to a form of paralysis: to avoid the risk of being seen as bad (*i.e.*, to have an updated assessment of ability less than $\widehat{\alpha}_0$), the executive pursues no innovations.

This result, though, is somewhat hard to reconcile with the celebration of innovative executives such as Steve Jobs. Reconciliation can be achieved if one views managerial ability as referring to an ability to discern good from bad projects rather than, as has so far been considered, an ability to execute projects successfully. To this end, consider the following model: in each of the two periods the manager will have a choice among three projects indexed by 1, 2, or \emptyset . Project p_\emptyset is equivalent to choosing no project. In each period, one of the projects in $\{p_1, p_2\}$ is good (g) and one is bad (b). The prior probability that a given project is good is $1/2$; that is, $\Pr\{p_1 = g, p_2 = b\} = \Pr\{p_1 = b, p_2 = g\} = 1/2$. These probabilities are independent between periods. A bad project pays -2 with certainty, a good project $+1$ with certainty. The project p_\emptyset pays 0 with certainty. Observe, therefore, that choosing a project p_1 or p_2 at random has a lower expected payoff, $-1/2$, than choosing project p_\emptyset . An executive can be able or unable. An able executive acquires a signal $s \in \{1, 2\}$ such that

$$\Pr\{p_s = g|s\} = \Pr\{p_{3-s} = b|s\} = q > \frac{2}{3}$$

$$\text{and } \Pr\{p_s = b|s\} = \Pr\{p_{3-s} = g|s\} = 1 - q < \frac{1}{3}.$$

An unable executive acquires no signal. Whether an executive acquires a signal is her private information. An able executive is able in both periods; that is, an executive who acquires a signal in period 1 knows she will acquire one in period 2. Similarly, an unable executive is unable in both periods: upon not

³⁵A slight caution: although the analysis leading to the proposition is complete if the market observes the mean and precision of any project taken (*i.e.*, as in Hermalin, 1993), it is a bit loose if the market cannot observe those parameters. In that case, there is the question of how observers form their beliefs about the parameters of any project the executive may have undertaken, especially if actually pursuing a project is an out-of-equilibrium act. Even accounting for those issues, it can be shown there is no equilibrium in which the executive pursues a project when he can plausibly claim one does not exist. See Holmstrom and Ricart i Costa (1986) for details.

acquiring a signal in period 1, she knows she won't acquire one in period 2. Let $f_0 \in (0, 1)$ be the commonly held prior probability that an executive is able.

Because

$$1 \times q - 2 \times (1 - q) = 3q - 2 > 0,$$

shareholders would like an able executive to act on her signal (*i.e.*, choose project p_s if she observes signal s). Because

$$1 \times \frac{1}{2} - 2 \times \frac{1}{2} = -\frac{1}{2} < 0,$$

shareholders would like an unable executive to choose project p_\emptyset . Absent career concerns—notably in the second period—there is no reason for the executive not to accede to those preferences. Hence, the executive's second-period market value if known to be able is $3q - 2$ and it's 0 if known to be unable.

Now consider the overall model. There are two equilibria. In one, shareholders believe that executives won't execute a project in the first period (equivalently, choose p_\emptyset) and that any executive who does attempt a project in $\{p_1, p_2\}$ is *unable*. Given such beliefs, both executive types will choose p_\emptyset in the first period: deviating means a second-period wage of 0 (since now seen as unable), whereas choosing p_\emptyset means a second-period wage of $(3q - 2)f_0$, the one-period expected value of an executive of unknown ability. Beliefs are consistent with behavior, so this is indeed an equilibrium.

In the other equilibrium, shareholders believe that both types will execute a project (*i.e.*, choose a $p \in \{p_1, p_2\}$) and that an executive who does *not* is unable. Based on project outcome, shareholders update their beliefs about executive ability:

$$\Pr\{\text{able}|g\} = \frac{qf_0}{qf_0 + \frac{1}{2}(1 - f_0)} \equiv f_1(g) \text{ and}$$

$$\Pr\{\text{able}|b\} = \frac{(1 - q)f_0}{(1 - q)f_0 + \frac{1}{2}(1 - f_0)} \equiv f_1(b).$$

An executive's future compensation based on outcome o is therefore $(3q - 2) \times f_1(o) > 0$; hence, both types do better to pursue a project than not. Beliefs are consistent with behavior, so this is, thus, also an equilibrium.

For this “toy” model, standard refinements (*e.g.*, Cho and Kreps, 1987) cannot be invoked to choose between these equilibria. One could, however, construct a richer model along these lines for which the second equilibrium would be selected. Consequently, we focus on the second equilibrium going forward. Because $f_1(g) > f_1(b)$, an executive who had a successful project in the first period will command greater compensation in the second than one who had an unsuccessful project. As such, the model is consistent with the idea that successful innovators are rewarded. It also yields an equilibrium in which executives are under pressure to pursue a project and will, in fact, do so even if that is contrary to what fully informed shareholders would want.

The structure of this model, in particular the shareholders' ability to discern whether or not a project was pursued, suggests that shareholders would do better to make first-period compensation contingent on choosing a project or not; or even the outcome of a first-period project. For instance, if

$$(3q - 2)f_0 - \frac{1}{2}(1 - f_0) < 0,$$

then the shareholders would do better to force no project in the first period by imposing a sufficiently large fine on any executive who pursues a first-period project than to let the second equilibrium play out. The shareholders can, though, do better than that: let $w_2 = 3q - 2$ denote the second-period compensation of an executive believed to be able and consider a first-period compensation contract $\langle w_g, w_b, w_\emptyset \rangle$, where w_o is compensation if the executive chooses a project with outcome o and w_\emptyset is her compensation if she doesn't choose a project. The objective is to have an equilibrium in which an executive pursues a project in the first period if and only if she learns she is able. Recalling that the executive doesn't know her ability prior to employment, her expected compensation—assuming she pursues a project only if she proves able—is

$$(qw_g + (1 - q)w_b)f_0 + w_\emptyset(1 - f_0). \quad (28)$$

Because, recall, there is a competitive market for managerial talent, (28) must equal the expected revenue generated by the executive, which is $(3q - 2)f_0$ ($= w_2f_0$). If it is an equilibrium for only able executives to pursue projects, then market beliefs following a first-period project must be that the executive is able if she pursues a project regardless of its success or failure. So if an able executive pursues a project, her expected lifetime compensation is

$$w_2 + qw_g + (1 - q)w_b. \quad (29)$$

If an unable executive pursues a project, her expected lifetime compensation is

$$w_2 + \frac{1}{2}w_g + \frac{1}{2}w_b. \quad (30)$$

Regardless of ability, an executive who doesn't pursue a project has a lifetime utility of w_\emptyset . For the contract to induce the correct behavior, (29) cannot be less than w_\emptyset and (30) cannot be greater than w_\emptyset . A necessary condition for those two incentive constraints to be satisfied is that

$$qw_g + (1 - q)w_b \geq \frac{1}{2}w_g + \frac{1}{2}w_b \iff \left(q - \frac{1}{2}\right)(w_g - w_b) \geq 0,$$

from which it follows that a necessary condition is $w_g \geq w_b$ (recall $q > 1/2$). In words, the executive's compensation if she chooses a project must be contingent on its success—with greater compensation due her for success—even though, conditional on choosing a project, nothing she does affects whether the project

is successful (*i.e.*, there is no effort that needs to be induced). As an example, were the shareholders to offer the contract

$$w_0 = 0; w_b = -w_2 \frac{1 + 2q}{2q - 1}; \text{ and } w_g = w_2 \frac{3 - 2q}{2q - 1}, \quad (31)$$

then this would satisfy the relevant constraints and induce a first-period project to be pursued if and only if the executive discovered herself able.

Summary. Although a “toy” model, the preceding illustrates the following possibilities in a setting in which the dimension of managerial ability about which the market is uncertain is the ability to distinguish good from bad projects and shareholders (boards) can write contingent contracts: (i) project failure is directly punished and project success directly rewarded (*i.e.*, from (31), $w_g > w_0 > w_b$) even though, conditional on choosing a project, the executive has no effect on its success or failure; and (ii) innovative (project-choosing) executives are rewarded by the market later in their careers relative to non-innovative executives even if their first-period innovation (project) fails (*i.e.*, $w_2 > 0$).

Bibliographic Note. Holmstrom and Ricart i Costa (1986) and Hermalin (1993) also consider contractual solutions to the issue of inducing executives to pursue projects, but in a context in which ability pertains, as earlier, to the payoff from the project; that is, project payoff is $\alpha + \beta + \varepsilon$. The need to compensate the executive for career risk affects what types of projects the shareholders seek to induce and the contracts they offer. For instance, in Hermalin, under mild assumptions, the shareholders will choose to induce a first-period project that has a lower mean return than the project with the maximum mean return, but is riskier than that project. Holmstrom and Ricart i Costa show that the tradeoffs faced by the shareholders can, in some circumstances, explain credit rationing within the firm.

4.2 MANAGERIAL MYOPIA

In the 1980s, American industry was criticized for “short-termism,” a focus on the next quarter rather than the long run. From the perspective of traditional finance theory, such criticism is puzzling: if, as the theory of efficient markets dictates, the stock price is an accurate reflection of the present value of the firm and if, as those alleging myopia claimed, managers are focused on the stock price, then it is hard to see how there could be scope for costly myopia: if short-termism adversely affected a firm’s long-term interests, then this would *depress* the stock price, an outcome at odds with what myopic managers were supposedly seeking to achieve.

As Narayanan (1985) and Stein (1988, 1989) observed, signal jamming provides a means of reconciling this apparent contradiction. To illustrate the basic idea, consider the following simplified version of Stein (1989): a debt-free firm exists for two periods; at the end of the second, it is liquidated and pays out

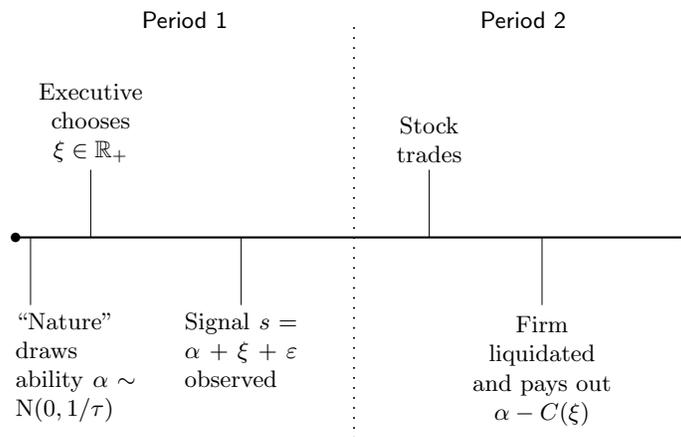


Figure 1: Timing for Simple Version of Stein (1989) Model

$\alpha - C(\xi)$ to the shareholders, where α is the ability of the executive, ξ is a signal-jamming action to be described shortly, and $C(\xi)$ is the cost of that action to the firm.³⁶ Assume $\alpha \sim N(0, 1/\tau)$.³⁷ The first period ends with a public signal $s = \alpha + \xi + \varepsilon$ being observed, where $\varepsilon \sim N(0, 1/\eta)$. Immediately thereafter (*i.e.*, at the beginning of the second period), the firm's stock trades. Assume this is the only point at which trade occurs. At the beginning of the first period, the executive chooses $\xi \in \mathbb{R}_+$. To keep matters straightforward, we will ignore financial discounting between periods. Figure 1 illustrates the timing.

The action ξ should be interpreted as one that appears to boost first-period accounting measures, but at the long-term expense of the firm (*e.g.*, borrowing at a disadvantageous interest rate, extending trade credit in a disadvantageous manner, etc.). That long-term expense is $C(\xi)$. Assume that $C(\cdot)$ is twice differentiable, $C''(\xi) > 0$ for all $\xi \in \mathbb{R}_+$, and that $C(0) = C'(0) = 0$. Observe, the mean-value theorem can be invoked to establish $C'(\xi) > 0$ for all $\xi > 0$ and, again, to establish $C(\xi) > 0$ for all $\xi > 0$. In turn, this implies that firm value is maximized if the executive chooses $\xi = 0$. Interpret, therefore, choosing $\xi > 0$ as engaging in myopic behavior.

Let the market anticipate that the executive chooses $\hat{\xi}$. The posterior esti-

³⁶In Stein (1989), α is an attribute of the firm rather than the executive. For our purposes, the distinction is irrelevant. Thinking of α as ability ties this model to those considered earlier.

³⁷Because α is distributed normally, there is the possibility that the final liquidated value of the firm is negative. This is not a problem insofar as the model can be modified without changing the conclusions to avoid negative values. For instance, at the cost of slightly complicating the analysis, we could assume liquidated value is $W + B(\alpha) - C(\xi)$, where W big enough that $W \geq C(\xi)$ for any equilibrium value of ξ , $B(\cdot)$ increasing, and $\lim_{z \rightarrow -\infty} B(z) \geq 0$. Normalizing $\mathbb{E}\alpha$ to zero is for convenience and without loss of generality.

mate of ability will, thus, be

$$\hat{\alpha} = \frac{(s - \hat{\xi})\eta}{\tau + \eta}$$

(the logic is the same as behind (14) above, recalling that the prior estimate of ability is 0). The estimated *value* of the firm at the end of period one is, thus,

$$\hat{\alpha} - C(\hat{\xi}) = \frac{(s - \hat{\xi})\eta}{\tau + \eta} - C(\hat{\xi}) \equiv V_1.$$

As an initial way of closing the model, suppose that each shareholder knows at the beginning of period one that she will need to liquidate her holdings of the firm at the end of period one with probability $\lambda \in (0, 1)$. Whether she needs to liquidate is *exogenously* determined.³⁸ It follows her expected payoff is her fraction of ownership times

$$\lambda EV_1 + (1 - \lambda)EV_2, \quad (32)$$

where $V_2 = \alpha - C(\xi)$ is final (liquidation) value. We can rewrite (32) as

$$\lambda \left(\frac{(\xi - \hat{\xi})\eta}{\tau + \eta} - C(\hat{\xi}) \right) - (1 - \lambda)C(\xi). \quad (33)$$

Shareholders would like their executive to choose the ξ that maximizes (33). Earlier made assumptions about $C(\cdot)$ ensure that the first-order condition

$$\frac{\lambda\eta}{\tau + \eta} - (1 - \lambda)C'(\xi) = 0 \quad (34)$$

uniquely defines a maximum. Let ξ^* solve (34). Because $C'(0) = 0$, $\xi^* > 0$ if $\lambda > 0$. In equilibrium, the market correctly anticipates that the executive will choose ξ^* , so the equilibrium maximized value of (33) is $-C(\xi^*)$. To summarize:

Proposition 6. *For the version of Stein (1989) presented above, the equilibrium level of myopic action, ξ^* , is positive. It is increasing in the probability that shareholders must sell at the end of period 1 (i.e., λ) and the precision of the signal (η); it is decreasing in the precision of the prior (τ).³⁹ In equilibrium, the initial shareholders bear 100% of the myopic action's cost.*

Given that the shareholders do not, *in equilibrium*, benefit from myopic actions, it might seem odd that they instruct the executive to behave in this way. Once again, the Red Queen logic prevails: because the market *anticipates*

³⁸As will be seen, in equilibrium, market valuation at the end of period 1 will equal the true expected value of the firm at the end of period 2, so a shareholder would be indifferent between selling and retaining her shares. The exogeneity of λ matters for what the shareholders instruct the executive to do. We return to this point later.

³⁹The proof of these comparative-static results is immediate from (34).

myopic behavior, the shareholders would be even worse off were the executive not to behave that way.

Observe that, if the shareholders could commit that the executive would not behave myopically, they would do better: the value of their holdings would increase by $C(\xi^*)$. One way for them to do so would be to commit to never sell at the end of period 1 (*i.e.*, set $\lambda = 0$). From (34), if $\lambda = 0$, then $\xi^* = 0$. This might be achievable if there were a controlling group of shareholders who never face a liquidity event and highly valued control so that they never sell.

What if λ is endogenous, but the shareholders cannot commit to its value? Because, as shown in (33), the shareholders payoff is affine in λ , either $\lambda = 0$ or $\lambda = 1$ is an optimum. If $\lambda = 0$, then, as noted, the optimal $\xi = 0$, which yields an expected payoff of 0. Expression (33) is, then,

$$\lambda \left(\frac{(\xi - 0)\eta}{\tau + \eta} - C(0) \right) - (1 - \lambda)C(\xi).$$

It follows immediately that if the shareholders deviated to $\lambda = 1$ and $\xi > 0$, their payoff would be positive.⁴⁰ So the only equilibrium must be $\lambda = 1$. The value of ξ would be its maximal value (call that $\bar{\xi}$).⁴¹ In equilibrium, no one is fooled, so $\hat{\xi} = \bar{\xi}$.

Proposition 7. *If the shareholders are free to determine whether they sell or not at the end of the first period and their probability of selling is unobservable to the market, then there will be the maximum level of myopic behavior.*

Proposition 7 suggests one reason why securities regulations require controlling parties to commit to a buying or selling pattern in advance.

An Agency Interpretation. To this point, it has been assumed that the (controlling) shareholders instruct the executive as to how myopic to be. Alternatively, one could assume that the executive has complete control over the choice of ξ and is motivated by the fact that he may have to sell at the end of period 1 with probability λ (set either exogenously, as in Proposition 6, or endogenously, as in Proposition 7).

Bibliographic Note. Among the differences between the model above and the one in Stein (1989) is that the latter assumes an infinite horizon, with ability following a random walk. Stein goes on to develop a steady-state equilibrium in which $\xi^* > 0$ is the same every period.

⁴⁰There are some subtleties concerning what the market would make of there being shares for sale when none were expected. A partial out is to assume that there some shareholders, who are not controlling shareholders, who will always sell. There are still questions as to why the price doesn't respond to the unexpected volume, but the points being made in the text would still apply even if we had a richer model of trading and price setting.

⁴¹One could imagine an interpretation in which $\bar{\xi}$ represents the controlling shareholders' looting the firm of its assets. Hence, there is a link between this model and models that worry about the pathologies of tightly controlled firms (see *e.g.*, Johnson et al., 2000).

4.3 EMPIRICAL ANALYSES OF MYOPIA

The traditional view is that, as long as managers are motivated to maximize their firms' stock prices, then they will not make myopic investment decisions:⁴² the stock market will punish them if they fail to make investments that don't have positive NPV. However, as reviewed above, Stein (1989) shows that there is an important distinction between value-maximization and stock-price maximization. Even though, in the long term, an efficient stock market will cause stock prices to coincide with underlying value, in the short term, there can be important deviations between the two. If a manager cares about boosting stock prices in the short term because he plans on selling the stock, then it will be rational for him to undertake myopic investments that boost short-term stock prices at the expense of long-term value. Stein's work, thus, raises an important question: do managers deviate from value-maximization when they compare investments of different horizons? In particular, do they prefer projects with early payoffs more than they should given relative interest rates? If so, what factors affect the magnitude of such myopic preferences in investments?

The idea of myopic investment, while perhaps foreign to the way economists traditionally perceive investment decisions, turns out to reflect how practitioners tend to view the world. The business press often claims that managers regularly sacrifice long-term value to meet short-term earnings goals. Consistent with this notion, survey evidence from Graham et al. (2005) finds that a majority of executives indicated that they would be willing to sacrifice long-term value to meet short-term earnings targets. The Stein model explains why executives would make such sacrifices even if they are compensated with equity-based incentives, and why increasing the magnitude of equity-based incentives could potentially exacerbate rather than alleviate the extent to which investment is myopic.

Aside from survey evidence, identification makes formal tests of myopia models relatively difficult. Managers' compensation plans and ownership, as well as their investment policies, are all determined jointly as a function of the investment opportunities that a firm faces. Nonetheless, there are a number of papers, many of which are in the accounting literature, that provide evidence suggesting that myopia is a real problem facing firms.

Dechow and Sloan (1991) provide evidence consistent with what they refer to as the "horizon problem." Specifically, they document that firms cut back on R&D expenditures investment toward the end of a CEO's tenure. Presumably, decreasing R&D leads to higher current earnings, which come at the expense of future earnings. Executives therefore appear to make myopic investment decisions toward the end of their tenure. Dechow and Sloan attribute these decisions to earnings-based compensation plans and find that this effect is minimized by stock ownership. It is not clear, however, how to interpret this finding in light of Stein (1989); the model predicts that temporary stock ownership will lead to more myopia but permanent stock ownership could offset the myopic tendencies managers acquire through their earnings-based plans.

⁴²See Jensen (1986) for discussion.

Gopalan et al. (2014) provide evidence that firms appear to care about myopia and the horizon of investments. These authors develop a measure of the time dimension to executive compensation, quantifying the extent to which the compensation of executives at a given firm is short term. They find that the timing of executive pay tends to follow that of their firms' investments, so that when firms' investments are longer-term, executive pay is longer-term as well. Presumably, firms are concerned about executives' tendencies to sacrifice valuable long-term investments for short-term profits. Therefore, when firms' more valuable investment opportunities are long term, firms respond by paying their executives with longer-term compensation. When their investments tend to be of a shorter duration, there is less need for longer-term compensation, so firms in this case rely more on shorter-term compensation plans.

Probably the most direct test of myopia theories to date is Edmans et al. (in press). These authors introduce a new measure of short-term incentives: the amount of stock and options scheduled to vest in a given quarter, which in the data is highly correlated with CEO equity sales. There are several aspects of this measure that make it particularly useful to test myopia theories. First, vesting schedules are set years in advance and are unlikely to be jointly determined with investment decisions at the time vesting occurs. Second, they correspond to the key theoretical variable in the myopia models, the fraction of stock that is likely to be sold by executives in the short term. It is this stock that will be sold in the short term that is the underlying determinant of the wedge between value and stock prices in the Stein model.

Edmans et al. (in press) link this measure to reductions in real investment. Their results indicate that there is a sizable decline in growth in R&D plus net capital expenditures in quarters during which CEOs have large amount of equity vesting. Vesting equity is also associated with more positive analyst forecast revisions and earnings guidance during the same quarter. Overall, this paper represents provides clear evidence that incentives to maximize short-term stock valuations can lead to myopic investments.

In many ways, corporate fraud can be viewed through the lens of myopia models, because fraud typically represents seeking benefits today in exchange for uncertain penalties in the future. Incentives that induce myopia in investments are also incentives to commit fraud, given they cause managers to place greater weight on today's payoff relative to future payoffs. So, for example, Burns and Kedia (2006), Efendi et al. (2007), and Johnson et al. (2009) all find that vested options and unrestricted stock increase the likelihood of corporate fraud.⁴³ These findings further suggest that equity-based compensation increases managers' incentives to boost earnings today at the expense of diminished value (or even penalties) tomorrow, consistent with the basic premise of the Stein model.

⁴³It should be noted, however, that Erickson et al. (2006) find that equity-based incentives do *not* increase the likelihood of fraud.

5 LEARNING ABOUT GOVERNANCE AND CAPITAL MARKETS

5.1 LEARNING MODELS AND CAPITAL MARKETS

An influential literature in finance applies learning models to study certain “anomalous” phenomena in capital markets. Two such phenomena are (i) that dividend yields apparently predict future stock returns (Fama and French, 1988), contrary to the efficient markets hypothesis; and (ii) the volatility of stock returns is higher than can be explained by the volatility of a firm’s dividend stream (Shiller, 1981; Leroy and Porter, 1981). To understand these facts, Timmermann (1993) simulates a model with uncertainty in which rational investors learn about the firm over time. He finds that the data from his simulated model are consistent with the empirical findings; in particular, dividends prove to be predictive of future returns and there is “excess” volatility of returns relative to that of dividends.

Pastor and Veronesi (2003, 2009) extend the ideas in Timmermann. The critical aspect of their model is incorporating uncertainty about a firm’s growth into the well-known Gordon growth formula:

$$P = \frac{D}{r - g}, \quad (35)$$

where P is the firm’s stock price, D its dividend, r the interest rate, and g the growth rate of dividends.

In the Pastor and Veronesi model, g is uncertain, but the market can make inferences about it based on observed data; that is, the market estimates growth prospects and continually updates its estimates over time. In this model, an application of Jensen’s inequality to the Gordon growth model leads to

$$P = \mathbb{E} \left\{ \frac{D}{r - g} \right\} > \frac{D}{r - \mathbb{E}g}.$$

Moreover, because the righthand side of (35) is a convex function of g , the greater the uncertainty about g (in the sense of second-order stochastic dominance), the greater the valuation (stock price) of the firm. So, for instance, holding *expected* growth rate constant, a startup with uncertain prospects will be worth more than an established firm for which there is little uncertainty about its rate of growth. Consequently, as the market learns about a firm’s prospects over time—uncertainty about its growth rates is reduced—the model predicts that its valuation should decline *ceteris paribus*. Pastor and Veronesi (2003, 2006) argue that this logic can explain the fact that firms’ market-to-book ratios tend to decline over time, and also how valuations of technology firms with very uncertain growth rates could be so high in the late 1990s.⁴⁴

In addition to influencing the *level* of a firm’s valuation, Pastor and Veronesi also show that the *sensitivity* of a firm’s valuation to news—hence its stock-price

⁴⁴Pastor and Veronesi (2009) is an excellent survey of the literature discussing the way in which learning can also explain a number of other seemingly puzzling stylized facts about capital markets.

volatility—is also affected by market learning. Intuitively, a positive innovation to dividends leads to an increase in valuation, but also to a positive update in the market’s estimate of the growth rate, which further increases valuation. Similarly, a negative innovation leads to a decrease in the assessment of the growth rate and consequently a larger decline in valuation than the news itself would warrant. The updates of the estimates of firms’ growth rates, which serve to magnify the effect of news and increase volatility, will be larger when the market’s prior estimate of the growth rate is more uncertain. Therefore, the Pastor and Veronesi model predicts that when the prior estimate of a firm’s growth rate is more diffuse, the reactions of the firm’s stock to news will be larger, so stock return volatility will be higher.

5.2 HOW LEARNING ABOUT MANAGEMENT AFFECTS CAPITAL MARKETS

Though stated in terms of growth rates, the Pastor and Veronesi model can be readily applied to learning about any other factor that affects a firm’s value; of particular relevance here is learning about the firm’s management. When a firm hires a new executive, his resume will be well known, but the fit between his skills and the firm’s needs are often unclear. For example, JC Penney hired Ron Johnson, a former Apple executive, to be its CEO in 2011. When Johnson’s strategy of adopting policies similar to those used by Apple turned out to decrease rather than increase profits, Johnson was fired. His style did not work as well at an old-fashioned retailing company, like JC Penney, as it had at a technology company like Apple (see *Business Week*, October 22, 2013).

The value of a management team to a particular firm reflects the quality of the match between the team and the firm, something which is revealed over time. In the case of Johnson, the appropriate inference to poor earnings was not only that the current year’s performance was bad, but that the strategy of adopting an Apple-style approach at JC Penney was dubious. Because of its implications for the future, JC Penny’s stock price declined more than it would have with a CEO whose match with the firm was more well-known. In short, uncertainty about a management team, particularly the fit between its approach and the firm, magnifies the effect of news on stock prices, and leads to higher stock return volatility.

Consistent with this idea, Clayton et al. (2005) document that stock-return volatility declines with CEO tenure. Presumably, early in a CEO’s tenure, news conveys, *inter alia*, information about the CEO’s ability (fit). Later in his tenure, when the CEO is more a known quantity, stock prices respond less to news and volatility is, correspondingly, lower.

Pan et al. (2015) examine whether the pattern of volatility over the CEO cycle fits with a variant of the Pastor and Veronesi (2003) model applied to uncertainty about management. Consistent with a learning model, Pan et al. find that the decline in volatility over a CEO’s tenure is convex: it declines more steeply with time earlier in his tenure than later in it. In addition, volatility declines more rapidly when there is more *ex ante* uncertainty about the incoming CEO: if he is younger, an outsider, not an “heir apparent,” or held fewer previous positions.

Pan et al. also find that the magnitude of the changes in value in response to news is larger early in a CEO's tenure, when the news reveals more information about his quality and overall fit with the firm.⁴⁵ Quantitatively, Pan et al. estimate that uncertainty about a new CEO and his policies explains roughly a quarter of overall stock-price volatility at the time of appointment, indicating that uncertainty about management has a substantial effect on volatility.

Arguably, rating agencies have been aware of this for some time insofar as they pay attention to "management risk" and consider it to be an important factor in their evaluations.⁴⁶ If, indeed, management risk is a meaningful factor affecting a firm's overall risk, then the cost of borrowing (*i.e.*, interest rates) should be greater at the time of top management turnover than at other times. Subsequently, as the market learns more about the quality of the management team and its match with the firm, rates should decline to approximately the pre-turnover levels.

Pan et al. (2016) test these predictions by estimating the extent to which the learning about management affects firms' cost of borrowing. Their findings are summarized in Figure 2, which illustrates the way credit default swap (CDS) spreads change around the time of a CEO change.⁴⁷ The announcement of a CEO's departure is associated with an increase in his firm's CDS spread, reflecting a greater estimate by the market of default risk. The CDS spread declines when a successor is announced; it declines further during the new CEO's time in office, approximately back to the pre-turnover level after about three years. Holding other factors constant, the five-year CDS spread is 36 basis points (24% relative to the sample mean) higher when a new CEO takes office than three years into his tenure. Spreads on loans and bond yields also decline following CEO turnovers. These effects are consistent with the idea, set forth above, that markets react to management risk—uncertainty about managerial ability and fit.⁴⁸

⁴⁵These results are robust to sample construction. They hold whether the data consist of the universe of turnovers from Execucomp or just the subsample of turnovers that are likely to have been exogenously determined (*i.e.*, were not the result of the executive's poor performance).

⁴⁶For example, a special document circulated by Moody's about corporate governance states: "[T]here is inherent transition risk in any CEO change and we therefore look to evaluate any changes to strategic initiatives or financial policies that differ from previous expectations, and whether credit metrics or liquidity deteriorates as a result." See Plath (2008).

⁴⁷A firm's CDS spread measures the premium one pays to insure against a potential default on the firm's debt. Because CDSs trade daily, they provide a continually updated, market-based estimate of expected default costs.

⁴⁸An alternative hypothesis to explain these results is that CEO turnovers are more likely at times of greater uncertainty about the firm. This could lead to patterns similar to those reported in Pan et al. (2016). To evaluate the importance of this alternative explanation, these authors also report results for various subsamples of "exogenously determined" turnovers (*e.g.*, due to death). In each case, the results are similar to those described here, suggesting that it is management risk that drives concerns about default risk.

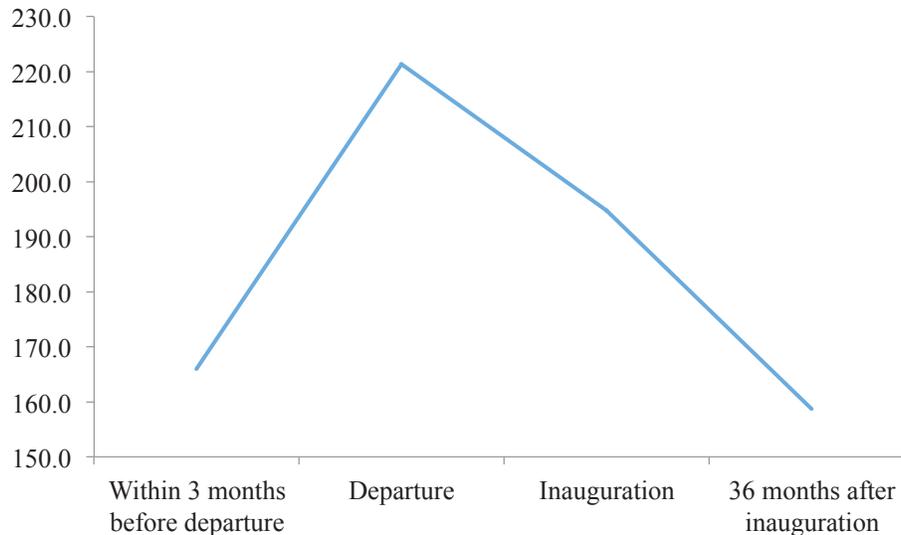


Figure 2: Average CDS Spreads at Key Events Related to CEO Turnovers.

Average CDS spreads (in basis points) at various points around a change in CEO. The turnover sample includes 284 CEO turnovers since 2001, for which the departure announcement and the inauguration occurred in different times. See Pan et al. (2016) for details.

5.3 CAPITAL MARKET INFERENCES ABOUT OTHER EXECUTIVES AND OTHER DIMENSIONS OF GOVERNANCE

Pan et al. (2016) also consider the effect of hiring a new Chief Financial Officer (CFO) on borrowing costs. They find that when there is a new CFO, even absent a change in CEO, effects similar to those seen with CEO turnover hold (although these effects are 36 to 58% smaller in magnitude than for CEOs). The explanation is the same as for CEOs (see Figure 2 and related text). The fact that the effects are smaller with CFO turnover than for CEO turnover suggests that a CEO's ability and match are of greater importance for a firm's future than a CFO's; that is, a CEO can add (or subtract) more value than a CFO.

Stern (2016) uses this approach to understand the importance of boards of directors. She examines the change in volatility around the appointment of directors. Stern's results indicate that there is a spike in volatility around the appointment of directors, followed by a decline as they become more well-known to the market. Uncertainty about the board accounts for between seven and ten percent of overall volatility. Stern also examines how volatility effects vary with circumstances. She finds differences exist; for example, her results imply that when there is a relatively powerful CEO, volatility changes around the arrival of directors are relatively small, consistent with a view that boards matter less when the CEO is more powerful. Overall, this approach is promising and is likely

to be useful in evaluating the importance of different governance attributes.

6 FURTHER IMPLICATIONS FOR GOVERNANCE

6.1 BARGAINING OVER OVERSIGHT

As noted in the introduction, a reaction to being monitored is to try to change how that monitoring is conducted; for example, seeking to change who the monitor is. Hermalin and Weisbach (1998) exploit this insight to explain a number of phenomena in governance, including management's influence over the composition of the board of directors.

The timing and principal assumptions of their model are:

1. The firm has a new CEO with ability α . Assume $\alpha \sim N(0, 1/\tau_0)$.
2. The first realization of earnings, x_1 , occurs. Assume $x_1 \sim N(\alpha, 1/\rho)$. Based on this, the board updates its beliefs about the CEO's ability and may choose to replace him. The ability of any *replacement* CEO is also drawn from $N(0, 1/\tau_0)$.
3. The CEO (incumbent or replacement) bargains with the board over filling board vacancies and over his compensation (wage) w . If bargaining fails, the CEO is fired, a replacement hired, and the stage repeated.
4. The board chooses how intensely to monitor the CEO, which is modeled as the probability, p , it acquires a private signal, y , about the CEO's ability. Assume $y \sim N(\alpha, 1/\eta)$.
5. If the board acquires the signal y , it updates its estimate of the CEO ability, which affects its decision to keep or replace the CEO.
6. The second realization of earnings, x_2 , occurs. Assume $\mathbb{E}x_2 = \alpha$. The random variables $y - \alpha$, $x_1 - \alpha$, and $x_2 - \alpha$ are independent.

The CEO in charge at stage 6 receives some nonpecuniary benefit $b > 0$ (*e.g.*, the additional utility that comes from the power and prestige of being CEO). A CEO dismissed prior to that receives no such benefit. The wage w agreed to in stage 3 is paid *regardless* of whether the CEO survives to stage 6. The CEO is protected by limited liability: he cannot be made to pay the firm, so $w \geq 0$.

As in many models of boards, the board is modeled in a somewhat reduced form: it acts as if its utility (payoff) is $x_2 - \ell d(p)$, where ℓ measures the board's *lack* of independence and $d : [0, 1] \mapsto \mathbb{R}_+$ is a disutility-of-effort function satisfying the usual properties of $d(0) = d'(0) = 0$ and $d''(p) > 0$ for all $p \in [0, 1]$ (so both disutility and marginal disutility are increasing functions). To ensure interior solutions to the board's choice problem *vis-à-vis* p , assume $\lim_{p \rightarrow 1} d'(p) = \infty$.

This is a normal learning model, so the posterior expectation of ability, $\hat{\alpha}'$, given prior $\hat{\alpha}$ and signal s ($s = x_1$ or y) is

$$\hat{\alpha}' = \frac{\tau \hat{\alpha} + \kappa s}{\tau + \kappa}, \tag{36}$$

where τ is the precision of the estimate $\hat{\alpha}$ and $\kappa = \rho$ or η depending on whether the signal is x_1 or y , respectively. Invoking now familiar arguments, the distribution of y given the CEO's estimated ability, $\hat{\alpha}$, is $N(\hat{\alpha}, 1/H)$, where

$$H = \frac{\eta\tau}{\eta + \tau}.$$

The posterior estimate of ability is, note, also the expectation of x_2 .

As usual, the game is solved via backward induction. Consider the decision to retain or replace the CEO at stage 5. By assumption, there is a large pool of *ex ante* identical possible replacements; hence, a replacement has no bargaining power. So he will be offered and accept the minimum possible wage, namely zero. Because $\mathbb{E}\{\alpha\} = 0$, expected earnings under a replacement are also zero; hence, the expected net value of a replacement is 0. So, if it learns y , the board dismisses the incumbent CEO if $\hat{\alpha}' < 0$. Similar to (6) above, this means the board will dismiss the CEO if $y < -\tau\hat{\alpha}/\eta \equiv Y$. Along the lines of the derivation of (9) above, a firm's expected value *if the board will learn y* is⁴⁹

$$V = \Phi(- (Y - \hat{\alpha})\sqrt{H})\hat{\alpha} + \frac{\sqrt{H}}{\tau}\phi((Y - \hat{\alpha})\sqrt{H}). \quad (37)$$

Lemma 3. *The value of the firm if monitoring will occur, V , is increasing in the prior expected ability of the incumbent, $\hat{\alpha}$, but the option value of monitoring, $V - \hat{\alpha}$, is decreasing in $\hat{\alpha}$.*

The board is assumed to choose its monitoring intensity to maximize its expected utility; that is, its choice of p solves

$$\max_{p \in [0,1]} pV + (1 - p) \max\{0, \hat{\alpha}\} - \ell d(p).$$

Observe that should it *not* receive the signal y —an event that happens with probability $1 - p$ —the value of the firm is $\max\{0, \hat{\alpha}\}$, which reflects that it will hire a replacement CEO if and only if its expectation of the incumbent's ability is less than its expectation of a replacement's.⁵⁰ Let P^* denote the solution to that optimization program. Assumptions previously made ensure both that $P^* \in (0, 1)$ and that P^* is unique.

Proposition 8. *The intensity with which the board monitors the CEO, P^* , is*

- (i) *decreasing with its prior estimate of the CEO's ability, $\hat{\alpha}$, whenever $\hat{\alpha} \geq 0$;*
- (ii) *decreasing with the precision of its prior estimate of his ability, τ ;*
- (iii) *decreasing with its collective lack of independence, ℓ ; but*
- (iv) *increasing with the precision (i.e., η) of the performance signal, y .*

⁴⁹Recall the incumbent must be paid w regardless of whether he is retained or not.

⁵⁰Again note that compensation promised the incumbent must be paid regardless of the decision to retain or replace him, so that compensation is a sunk expenditure.

Assume that bargaining between board and incumbent CEO is Nash (1950): their agreement maximizes the product of their surpluses. The limited-liability assumption (*i.e.*, $w \geq 0$) means the board can capture none of the benefit, b , that a new CEO would enjoy; hence, the board's surplus is simply its expected payoff from retaining the incumbent CEO (if that's positive). The incumbent CEO's expected surplus is his *expected* benefit,

$$\left(p\Phi(- (Y - \hat{\alpha})\sqrt{H}) + 1 - p \right) b. \quad (38)$$

Before analyzing the bargaining game, it's worth considering the intensity of monitoring that would maximize *total* surplus. Total surplus is

$$pV + (1 - p) \max\{0, \hat{\alpha}\} - \ell d(p) + \left(p\Phi(- (Y - \hat{\alpha})\sqrt{H}) + 1 - p \right) b.$$

The intensity of monitoring, p , that maximizes total surplus, P^{**} , solves the first-order condition (FOC):

$$\underbrace{V - \max\{0, \hat{\alpha}\} - \ell d'(p)}_{=0 \text{ is FOC determining } P^*} - \left(1 - \Phi(- (Y - \hat{\alpha})\sqrt{H}) \right) b = 0. \quad (39)$$

Because $d'(\cdot)$ is increasing, (39) implies that $P^{**} < P^*$: the joint-surplus-maximizing intensity of monitoring is less than what the board prefers. The reason is that, in maximizing its own payoff, the board ignores the benefit, b , the incumbent CEO can achieve. That benefit is lost if he is fired: limited liability means the board can't capture it from a *new* CEO; so, the marginal *joint* benefit of monitoring is less than the board's own benefit. In sum, we have:

Proposition 9. *The intensity of monitoring, p , chosen by the board is greater than the intensity that would maximize expected total surplus (*i.e.*, $P^* > P^{**}$).*

That $P^{**} < P^*$ reflects, in essence, a commitment problem: to generate more collective surplus, the board and incumbent CEO would like to fix *ex ante* a lower level of monitoring. The one lever they have to do so is to adjust the board's lack of independence, ℓ : from Proposition 8(iii), the more the board lacks independence, the less intensely it monitors. This provides a motive for the board and incumbent CEO to bargain for a less independent board at stage 3. Let ℓ_0 denote the collective lack of independence of those directors continuing as directors. Suppose that bargaining with the CEO results in board vacancies being filled such that the new board's lack of independence is ℓ_1 . If $P^*(\ell)$ denotes the intensity of monitoring chosen by a board with lack of independence ℓ , then the *continuing* directors' expected utility is

$$P^*(\ell_1)V + (1 - P^*(\ell_1)) \max\{0, \hat{\alpha}\} - \ell_0 d(P^*(\ell_1)). \quad (40)$$

Consider now the bargaining between continuing directors and the incumbent CEO. They will agree to a lack of independence, ℓ_1 , and wage, w , that maximizes the product of their surpluses from trade. In light of Proposition 8(iii),

$P^*(\cdot)$, is monotone; hence, there is no loss in treating P^* as the choice variable. Because the CEO gets zero if no agreement is reached, his surplus is

$$\left(P^*\Phi(- (Y - \hat{\alpha})\sqrt{H}) + 1 - P^*\right)b + w. \quad (41)$$

If no agreement is reached, the board hires a replacement CEO. Let U_0 denote the board's expected utility if it hires a replacement. Its surplus is, thus,

$$P^*V + (1 - P^*) \max\{0, \hat{\alpha}\} - \ell_0 d(P^*) - w - U_0. \quad (42)$$

Two more steps are necessary before we can determine the w and P^* that maximize the product of (41) and (42) (*i.e.*, solve the bargaining game): determine what U_0 is and show that the board fires, prior to bargaining, an incumbent CEO for whom $\hat{\alpha} < 0$.

Lemma 4. *The board's outside option in its bargaining with an incumbent CEO is $U_0 = P_0 V_0 - \ell_0 d(P_0)$, where $V_0 = \phi(0)\sqrt{H}/\tau_0$ (*i.e.*, the expected value of the firm under a new CEO) and P_0 maximizes the existing board's (continuing directors') expected utility; that is, P_0 solves the first-order condition $V_0 - \ell_0 d'(p) = 0$ (equivalently, $P_0 = P^*(\ell_0)$ for a new CEO).*

Observe the wage paid a new/replacement CEO is zero.

Proposition 10. *A unique finite cutoff, $A > 0$, exists such that an incumbent CEO is fired prior to bargaining if and only if his estimated ability is less than A . Moreover, A is decreasing in the board's lack of independence (*i.e.*, $\partial A/\partial \ell_0 < 0$).*

The result $A > 0$ reflects that more is known about an incumbent CEO than his replacement, so a replacement has a greater option value than he does; hence, his expected ability needs be greater than a replacement's for him to keep his job. We explore the implications of different option values further in Section 6.2.

Back to the bargaining problem (*i.e.*, maximizing the product of (41) and (42)): Because $\hat{\alpha} > A$, the first-order conditions with respect to p and w are

$$\begin{aligned} (V - \hat{\alpha} - \ell_0 d'(p))(p\Phi b + (1 - p)b + w) \\ + \underbrace{(p(V - \hat{\alpha}) + \hat{\alpha} - \ell_0 d(p) - w - U_0)}_{>0 \text{ given incumbent CEO retained}} \underbrace{(\Phi - 1)b}_{<0} = 0 \end{aligned} \quad (43)$$

and

$$p(V - \hat{\alpha}) + \hat{\alpha} - \ell_0 d(p) - w - U_0 - (p\Phi b + (1 - p)b + w) \leq 0.$$

(Note the arguments of Φ have been suppressed for the sake of readability.) Observe (43) entails $V - \hat{\alpha} - \ell_0 d'(p) > 0$; hence, $p < P^*(\ell_0)$, which means $\ell_1 > \ell_0$: bargaining with a CEO who is retained leads to a less independent board. This establishes:

Proposition 11. *If the continuing directors choose to retain the CEO, then the new board is less independent than the continuing directors (*i.e.*, $\ell_1 > \ell_0$).*

What about the compensation of a CEO who is retained? If the limited-liability constraint binds, then it is zero. If it doesn't bind, then the second first-order condition entails

$$p(V - \hat{\alpha}) + \hat{\alpha} - \ell_0 d'(p) - w - U_0 - (p\Phi b + (1-p)b + w) = 0.$$

Plugging that back into (43), it follows, after some algebra, that

$$V - \hat{\alpha} - \ell_0 d'(p) + (\Phi - 1)b = 0.$$

That, note, is the first-order condition for maximizing joint surplus (*i.e.*, expression (39) above). If, instead, the limited-liability constraint binds, then

$$V - \hat{\alpha} - \ell_0 d'(p) + (\Phi - 1)b < 0.$$

When the limited-liability constraint binds there is more monitoring (equivalently, greater board independence) than if it doesn't bind. In sum,

Proposition 12. *Suppose that the incumbent CEO is retained. If the limited-liability constraint is not binding, then the level of monitoring will maximize the CEO and board's joint surplus. If it is binding, then the level of monitoring will exceed the joint-surplus-maximizing level. Correspondingly, board independence will be greater if the constraint is binding than if it is not binding.*

Empirical Validation. The model analyzed above has a number of empirical predictions, including:

1. A CEO who performs poorly is more likely to be replaced than one who performs well.
2. *Ceteris paribus*, a CEO is more likely to be dismissed for poor performance when the board is more independent (*i.e.*, when ℓ is lower).
3. Strong performance should lead to less independent (higher- ℓ) boards.
4. Board independence declines over the course of a CEO's tenure.

The results of a number of empirical studies are consistent with these predictions. The first prediction has considerable support: a long list of articles plus popular press coverage document the result.⁵¹ Weisbach (1988) provides strong evidence for the second prediction. The last two predictions are consistent with the patterns of *director* turnover documented by Hermalin and Weisbach (1988).

⁵¹For a recent popular-press article see James Surowiecki's "Why CEOs are Getting Fired More," *New Yorker*, November 7, 2016. For discussions of relevant scholarly articles, see Hermalin and Weisbach (1998) or Hermalin (2013).

Summary. The Hermalin and Weisbach (1998) model is one of the first *bargaining models of governance*. In particular, the rarer a commodity a CEO proves to be (*i.e.*, to have higher estimated ability than a replacement), the more leverage he has to extract what he wants from the board of directors. Not surprisingly, he wants greater compensation. But, to the extent he obtains other non-pecuniary benefits from being CEO (the parameter b), he also wants to keep his job. The latter objective makes him seek a board that will be a less vigilant monitor; that is, one that is less likely to dismiss him *ceteris paribus*. Because an incumbent board sees itself as optimally vigilant, it views yielding somewhat on vigilance (agreeing to filling board vacancies with less independent directors—raising ℓ) as a second-order loss; whereas paying the CEO more is a first-order loss. Hence, a CEO has an easier time negotiating a higher ℓ than more pay, at least when he is only marginally more valuable than a replacement.⁵² A “star” CEO, who is significantly more valuable than a replacement, can bargain for both a less independent board (higher ℓ) and greater pay. Observe the two roles played by learning in this model: (i) it is because the board learns that the CEO is a rare commodity that he gains bargaining power; and (ii) because the CEO wants to constrain the board from learning more about him, he uses that bargaining power to lessen the odds it will learn more about him.

6.2 CONSEQUENCES FOR CHOICE OF MANAGER

As reviewed in Hermalin (2005, 2013), among other sources, there have been a number of trends in corporate governance in the past quarter century:

- In many countries, notably the US, there has been a significant shift toward having outside directors be a greater proportion of the board.⁵³ Such directors are generally seen as more independent of management and, thus, more diligent monitors (in the notation of the previous subsection, a more independent board is a lower ℓ board).
- In many countries, again notably the US, there has been a trend toward more external hiring of the CEO (versus appointing someone who rose through the company’s ranks).
- There has also been a trend toward greater CEO compensation, both contingent and non-contingent.
- There has been a trend toward shorter CEO tenures.

A question is how interconnected these trends might be? A further question is what do these trends suggest about governance?

⁵²More valuable accounting for the greater option value a replacement has.

⁵³*Outside directors* are typically defined as directors whose sole involvement with the company is in their role as a director; hence, for example, a university president serving on a corporate board would be an outsider. In contrast, *inside directors* are generally employees of the company (*e.g.*, the CEO is an inside director). See Adams et al. (2010) for details.

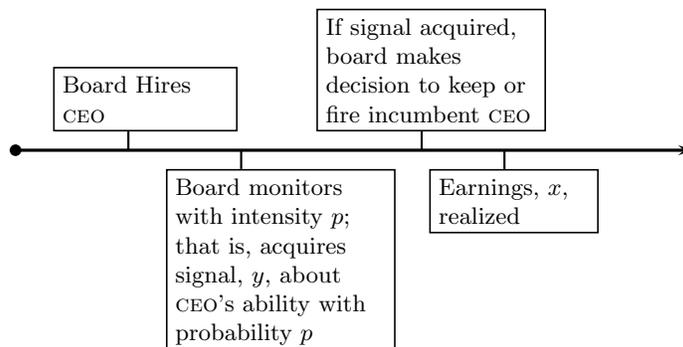


Figure 3: Basic timeline of the Hermalin (2005) model

As discussed earlier, among the roles of the board of directors are hiring a CEO, monitoring (assessing) him, and replacing him if necessary. A key point, though, is that the way in which the board will fulfill the latter two roles has implications for who it wants to hire in the first place. Hermalin (2005) explores this insight via a model that builds on Hermalin and Weisbach (1998). The model’s basic elements are summarized in Figure 3.

As in the previous subsection, the board’s preferences are captured by the utility function $x - \ell d(p)$. As before $\mathbb{E}\{x\} = \alpha$, where α is the ability of the CEO in place at the end (*i.e.*, either the incumbent if retained or his replacement otherwise). Similar to before $\alpha \sim N(\mu, 1/\tau)$. The signal $y \sim N(\alpha, 1/\eta)$. As before, the random variables $y - \alpha$ and $x - \alpha$ are independently distributed.

Assume that the prior estimate of the *incumbent* CEO’s ability (his μ) is strictly positive. Assume, however, the prior estimate of a *replacement* CEO’s ability (his μ) is zero. Setting the latter to zero is a convenient normalization that entails no loss of generality. The higher prior estimate for the incumbent is a way of accounting for firing (transition) costs when replacing one CEO with another (*i.e.*, the r of the Section 2.2 model). An implication of this assumption is that, *absent* any new information about the incumbent CEO, the board would want to retain him: his expected value exceeds his replacement’s.

Given this assumption, there should be no confusion in letting μ now simply denote the prior estimate of the incumbent CEO’s ability.

Via now familiar reasoning, if the board obtains the signal y , then the posterior distribution of ability is normal with mean $\hat{\mu}$ and precision $\hat{\tau}$, where

$$\hat{\mu} = \frac{\tau\mu + \eta y}{\tau + \eta} \text{ and } \hat{\tau} = \tau + \eta.$$

Observe, then

$$\mathbb{E}\{x|y\} = \hat{\mu}, \mathbb{E}\{x|\text{no signal}\} = \mu, \text{ and } \mathbb{E}\{x|\text{replace incumbent}\} = 0.$$

As noted, the board will fire the incumbent CEO only if it gets a signal and, then, only if the signal is such that $\hat{\mu} < 0$. Given the formula for $\hat{\mu}$, the signal

will cause the incumbent CEO to lose his job if and only if

$$\frac{\tau\mu + \eta y}{\tau + \eta} < 0 \text{ equivalently, iff } y < \frac{-\tau\mu}{\eta} \equiv Y.$$

Again building on previous analysis, the *prior* distribution of y is

$$N\left(\mu, \frac{1}{\eta} + \frac{1}{\tau}\right) \equiv N\left(\mu, \frac{1}{H}\right).$$

It follows that the expected value of the firm *if the signal will be observed* is

$$V \equiv \int_{-\infty}^{\infty} \underbrace{\max\left\{0, \frac{\tau\mu + sy}{\tau + s}\right\}}_{\mathbb{E}\{\alpha|y, \text{optimal replacement}\}} \sqrt{\frac{H}{2\pi}} \exp\left(-\frac{H}{2}(y - \mu)^2\right) dy.$$

Consequently, utilizing familiar methods (see, *e.g.*, the derivation of (9) above),

$$\begin{aligned} V &= \left(1 - \Phi((Y - \mu)\sqrt{H})\right)\mu + \frac{\sqrt{H}}{\tau}\phi((Y - \mu)\sqrt{H}) \\ &= \Phi(-(Y - \mu)\sqrt{H})\mu + \frac{\sqrt{H}}{\tau}\phi((Y - \mu)\sqrt{H}). \end{aligned}$$

Observe $\Phi((Y - \mu)\sqrt{H})$ is the probability that the CEO is fired conditional on knowing the signal will be obtained.

If the CEO cannot be monitored (or, equivalently, if no signal is obtained), the value of the firm is just the incumbent CEO's expected ability, μ . Being able to monitor the CEO is valuable insofar as $V > \mu$; that is, monitoring creates an option worth $V - \mu$.

The board chooses its intensity of monitoring, p , to maximize

$$pV + (1 - p)\mu - \ell d(p).$$

Let P^* denote the solution. The assumptions made earlier about $d(\cdot)$ ensure that P^* is unique given V , μ , and ℓ , and that it lies in $(0, 1)$.

Proposition 13. *The intensity with which the board monitors the CEO in equilibrium, P^* , is*

- (i) *decreasing with the prior estimate of the CEO's ability, μ ;*
- (ii) *decreasing with the precision of the prior estimate, τ ; and*
- (iii) *decreasing with the board's lack of independence, ℓ .*

The proof of Proposition 13 is the same as Proposition 8's.

Now extend the model to include the board's initial choice of CEO; in particular, assume the board has a choice between an internal candidate (index I) and an external candidate (index E). Assume it is common knowledge that the

ability of the two candidates are drawn, respectively, from the normal distributions $N(\mu_I, 1/\tau_I)$ and $N(\mu_E, 1/\tau_E)$ (where both μ s are positive). The critical assumption is that $\tau_I > \tau_E$: the board has better information about the internal candidate than the external one. As Hermalin (2005) details, there are many reasons this assumption is empirically valid.

A key result in this extended model is the following:

Proposition 14. *Fix $\tau_I > \tau_E$ and define $\Delta = \mu_I - \mu_E$. Then, fixing μ_I (or μ_E), there exists a minimum $\Delta_m > 0$ such that the external candidate is hired if and only if $\Delta < \Delta_m$. Moreover, Δ_m increases as the board gains independence (i.e., the lower is ℓ); that is, more independent boards are more inclined to hire external candidates, all else equal.*

Intuition: monitoring creates an option value that increases with uncertainty about the CEO's ability (i.e., $\partial V/\partial\tau < 0$ from Proposition 13(ii)). That means the board would strictly prefer, between two CEOs with the same expected ability (same μ s), the one about which it was more uncertain. Hence, the board should be willing to hire the external candidate even if $\mu_E < \mu_I$, provided the difference is not too great. That is, external candidates have an "edge." At the same time, the external candidate's greater option value is *conditional* on the CEO actually being monitored; so, this effect is tempered by how likely the board is to monitor. This, then, explains the "moreover" portion of Proposition 14.

An immediate corollary is:

Corollary 1. *A necessary condition for the internal candidate to be hired is that his estimated ability be strictly greater than that of the external candidate (i.e., that $\mu_I > \mu_E$).*

Observation, including field work, suggests that outside directors are more inclined to monitor than are inside directors. Additionally, an empirical study by Weisbach (1988) finds evidence in support of that hypothesis. If, then, outside directors are more likely to monitor (i.e., a greater proportion of outside directors corresponds to a lower ℓ), then a trend toward greater outsider representation on boards should, in light of Proposition 14, lead to more external candidates being hired as CEOs *ceteris paribus*. Hence, this model can serve to tie together two of the trends noted above: greater proportion of outsiders on the board and increased hiring of CEOs from outside the firm.

Denis et al. (in press) suggest that this result can run in reverse in certain circumstances; namely, that when a firm knows it will employ a CEO about whom it is more uncertain, then it will want to have a board more inclined to monitor. Such a scenario arises with *de novo* firms (e.g., as in the Denis et al., when the firms are spinoffs of existing firms). Denis et al. find evidence to support this hypothesis in their empirical study of 143 spinoffs.

Another implication of the model is the following. Recall that were there no monitoring at all, the initially hired CEO (the incumbent) would never be replaced. In contrast, when there is monitoring the incumbent CEO faces a positive probability of being fired. The more monitoring there is, the greater

the proportion of incumbent CEOs who will be fired. This, then, means that the average length of CEO tenure should fall. So, one gets the following causal chain: more outsider representation on boards entails more monitoring, which entails shorter CEO tenures on average. So the model ties together the trends of a greater proportion of outsiders on the board and shorter average CEO tenures.

There is also an indirect effect that reinforces the effect just identified: greater outsider representation leads to a greater proportion of external hires; externally hired CEOs are more uncertain entities, which means their *ex ante* probability of being fired is greater than internal hires *ceteris paribus*. So this is another path linking greater outsider representation and shorter average tenures:

Proposition 15. *Suppose the board is indifferent between hiring the external candidate or the internal candidate as CEO. Then, if the external candidate is hired,*

- (i) *he is more intensely monitored than the internal candidate would have been (i.e., $P_E^* > P_I^*$);*
- (ii) *he is more likely to be dismissed than the internal candidate would have been conditional on being monitored (i.e., conditional on the signal y being obtained); and, hence,*
- (iii) *he has a shorter expected tenure than the internal candidate would have been expected to have.*

Proof: Result (i) follows from Propositions 13 and 14 given both $\mu_I < \mu_E$ and $\tau_I < \tau_E$. Result (ii) follows if $-(Y - \mu)\sqrt{H} = \mu\tau/\sqrt{H}$ increases in μ and τ . That it increases in μ is immediate. Straightforward differentiation and algebra reveal it's increasing in τ too. Result (iii) is immediate from (i) and (ii). ■

Effort and Compensation. The model so far assumes no effort by the CEO nor does it consider compensation. Suppose, as in earlier models, that the CEO wishes to keep his job. Suppose too that effort influences the board's perception of his ability if it monitors; that is, there is a signal-jamming motive for him to expend effort. Recalling, though, that such signal jamming fools no one in equilibrium—the Red Queen effect holds—so effort makes the CEO directly worse off without truly affecting his probability of retaining his job; consequently, he will demand compensation for this effort. Indeed, even if there were no scope for signal-jamming effort, the fact that greater monitoring puts him at greater risk of losing his job could also lead him to demand greater compensation. To explore these issues consider the modified model of Figure 4.

Make the standard assumptions about the CEO's disutility-of-effort function: it is twice differentiable; $k(0) = k'(0) = 0$; and it exhibits increasing *marginal* disutility of effort. It follows that $k'(e) > 0$ for all $e > 0$. Assume the benefit of effort e , $\beta(e)$, is realized *after* the board makes its keep or fire decision, but its realization does not depend on that decision (i.e., the firm gets $\beta(e)$ whether

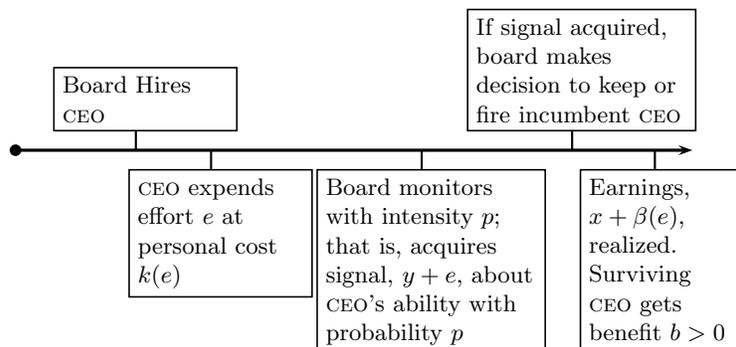


Figure 4: Timeline of expanded model with CEO effort

the initially hired CEO keeps or loses his job). Because the properties of $\beta(\cdot)$ are irrelevant to the analysis at hand, we will not focus on it in what follows. As with previous signal-jamming models, the board cannot observe y separately from e ; instead, it observes $\tilde{y} \equiv y + e$. As noted, the board will *anticipate* some level of effort from the CEO. Denote that anticipated level by \hat{e} .

As before, the CEO is fired if the best estimate of y is too low. Specifically, following the prior analysis, he's fired if

$$\tilde{y} - \hat{e} < -\frac{\tau\mu}{\eta} \equiv Y.$$

We can rewrite that cutoff rule as the CEO loses his job if $y < Y + \hat{e} - e$. The CEO chooses his effort to maximize his expected utility; that is, to maximize

$$\left(P^* \Phi \left(- (Y + \hat{e} - e - \mu) \sqrt{H} \right) + (1 - P^*) \right) b - k(e).$$

The first-order condition for maximizing that expression is

$$bP^* \phi \left(- (Y + \hat{e} - e - \mu) \sqrt{H} \right) \sqrt{H} - k'(e) = 0.$$

As has been noted repeatedly, in a pure-strategy equilibrium, \hat{e} must equal the e the CEO chooses. Hence, equilibrium effort, e^* , must satisfy

$$bP^* \phi \left(- (Y - \mu) \sqrt{H} \right) \sqrt{H} - k'(e^*) = 0. \quad (44)$$

The assumed properties of $k(\cdot)$ ensure that a unique e^* exists that solves (44).

Expression (44) is only a *necessary* condition; existence of a pure-strategy equilibrium also requires that

$$e^* \in \operatorname{argmax}_e \left(P^* \Phi \left(- (Y + e^* - e - \mu) \sqrt{H} \right) + (1 - P^*) \right) b - k(e). \quad (45)$$

Additional assumptions on $k(\cdot)$ are needed to ensure (45) holds. Fortunately, such assumptions are plausible (see Hermalin, 2005, for details) and henceforth maintained. A key result for this extended model is the following:

Proposition 16. *Assume the game with CEO effort has a pure-strategy equilibrium. Then the following comparative statics hold:*

- (i) *the lower the CEO's estimated ability, the more effort he expends in equilibrium; and*
- (ii) *the less independent is the board (i.e., the greater is ℓ), the less effort the CEO expends in equilibrium.*

Given that, conditional on employment, externally hired CEOs tend to have lower μ , Proposition 16 suggests that external CEOs will expend more effort in equilibrium than CEOs hired from the inside. The proposition also suggests that a trend toward greater board independence (a growing trend for greater outsider representation) could lead to CEOs' expending more effort.

Now consider compensation. A CEO's equilibrium utility is

$$\left(P^* \Phi(- (Y - \mu) \sqrt{H}) + (1 - P^*) \right) b - k(e^*).$$

The derivative of that with respect to ℓ , the lack-of-independence parameter, is

$$\underbrace{\frac{\partial P^*}{\partial \ell}}_{-} \underbrace{(\Phi - 1)b - k'(e^*)}_{-} \underbrace{\frac{\partial e^*}{\partial \ell}}_{-} > 0$$

given Propositions 13 and 16. So a CEO's equilibrium utility *falls* if ℓ *falls*; that is, if the board is *more* independent. People typically require compensation if their utility is lowered. Combined, these insights yield:

Proposition 17. *If CEOs with similar attributes enjoy equal expected utility in the equilibrium of the CEO market, then, controlling for attributes, CEOs who work for more independent boards will receive greater compensation than CEOs who work for less independent boards.*

It is not necessary to introduce effort into the model to identify a relation between board independence and executive compensation. To see this, return to the version of the model *without* effort. Assume a CEO's expected utility is

$$w + \left(P^* \Phi(- (Y - \mu) \sqrt{H}) + (1 - P^*) \right) b, \tag{46}$$

where w denotes compensation. Assume further that a CEO accepts a job if and only if his expected utility (i.e., expression (46)) exceeds his reservation utility, \underline{U} . If we assume that the firms have all the bargaining power and can, thus, hold a CEO to his reservation utility, then

$$w = \underline{U} - \left(P^* \Phi(- (Y - \mu) \sqrt{H}) + (1 - P^*) \right) b.$$

The derivative with respect to ℓ is

$$\frac{\partial w}{\partial \ell} = -b \underbrace{\frac{\partial P^*}{\partial \ell}}_{<0} (\Phi - 1) < 0.$$

In words: *the less independent the board, the less the CEO's equilibrium pay.*

Summary. The causal links just identified are readily summarized as follows:

- More independent boards (greater outsider representation) means more monitoring of CEOs (*i.e.*, higher values of p in equilibrium).
- More monitoring of CEOs leads to:
 - More effort by CEOs
 - More hiring of external CEOs
 - Shorter average tenures for CEOs
- More hiring of external CEOs leads to:
 - More effort by CEOs
 - Shorter average tenures for CEOs
- More effort by CEOs and shorter average tenures for CEOs *both* lead to greater compensation.

6.3 CONSEQUENCES FOR CHOICE OF INFORMATION STRUCTURES

As touched upon in Section 2.2, shareholders and their executives can have divergent interests with respect to the quality of information used in assessing executives' ability. Hermalin and Weisbach (2012) provide a more in-depth analysis of that issue, which we briefly review here.

The basic timing and assumptions of Hermalin and Weisbach (2012) are:

1. The firm's owners determine the quality of information—what Hermalin and Weisbach call a *disclosure regime*. Let \mathcal{D} denote a disclosure regime.
2. The owners negotiate with a potential CEO as to his compensation, w , and (in equilibrium) hire him.
3. The owners learn information, the quality of which depends on the disclosure regime.
4. The owners base an action (*e.g.*, firing the CEO) on what was learned.
5. Payoffs are realized.

This is a very general model that pertains to a number of situations, including:

- Information pertains to the CEO's ability. The consequent action is whether to retain or fire him. As in Section 2.2, the CEO suffers a loss if fired.
- Information pertains to the firm's prospects. The consequent action is whether to invest additional resources into the firm or to take resources out. The CEO prefers to manage more resources than fewer, *ceteris paribus*.

- Information learned by the owners lessens an informational advantage held by the CEO. The consequent action is the owners adjust the CEO's compensation plan, resulting in the CEO's loss of information rents.
- Information refers to how informative firm performance is about the actions the CEO has taken. The consequent action is the owners adjust the CEO's compensation plan, resulting in his loss of quasi-rents.

Let $\mathcal{D} \succ \mathcal{D}'$ denote that disclosure regime \mathcal{D} is more informative than regime \mathcal{D}' with respect to some recognized notion of informativeness, such as Blackwell informativeness.

Let the owners' payoff be $\pi(\mathcal{D}) - w$ and the CEO's $U(\mathcal{D}) + v(w)$, where $v(\cdot)$ is an increasing function. Assume the CEO is protected by limited liability: $w \geq 0$.

The following condition is key to the Hermalin and Weisbach (2012) analysis:

Condition 1. *If \mathcal{D} and \mathcal{D}' are two disclosure regimes such that $\mathcal{D} \succ \mathcal{D}'$, then $\pi(\mathcal{D}) > \pi(\mathcal{D}')$ and $U(\mathcal{D}) < U(\mathcal{D}')$.*⁵⁴

In words: owners prefer a more informative disclosure regime to a less informative one, whereas a CEO prefers a less informative regime. Recall that Condition 1 followed from first principles in the model in Section 2.2 above. Indeed, Hermalin and Weisbach show that in many situations of interest (including those just listed) the condition follows from first principles.

Hermalin and Weisbach (2012) assume that bargaining between owners and the CEO is generalized Nash; to wit, compensation, w , is set to maximize

$$\theta \log (\pi(\mathcal{D}) - w) + (1 - \theta) \log (U(\mathcal{D}) + v(w) - \underline{u}),$$

where $\theta \in [0, 1]$ is the owners' bargaining power and \underline{u} is the CEO's reservation utility (the value of his outside option).

The first result is⁵⁵

Proposition 18. *Assume wage bargaining is generalized Nash and Condition 1 holds. Then the CEO's compensation, as determined by the bargaining process, is non-decreasing in how informative is the disclosure regime.*

The proof is a straightforward application of standard comparative static methods (see Hermalin and Weisbach, 2012, pp. 223–224 for details) and so omitted. As noted in Section 2.2, there is a cost to owners of better information (beyond any direct cost of producing it) because they must compensate their executive for the disutility it imposes on him.

Although, as noted, Hermalin and Weisbach (2012) consider a wide variety of models that satisfy Condition 1, we limit attention to a learning model, in

⁵⁴In Hermalin and Weisbach (2012), the condition is slightly weaker insofar as they allow $\pi(\mathcal{D}) \geq \pi(\mathcal{D}')$; that greater generality is not needed for our purposes here.

⁵⁵For the purposes of this chapter, we ignore the possibility of corner solutions; see Hermalin and Weisbach (2012) for an analysis that accounts for possible corner solutions.

keeping with the theme of this chapter.⁵⁶ Suppose that the owners' payoff is

$$x\gamma(a) - c(a), \tag{47}$$

where a is an action taken by the owners from a known set \mathbf{A} and x is a stochastic return such that $\mathbb{E}\{x\} = \alpha$. As above, α is an unknown parameter. Consistent with previous analysis, it can be considered the CEO's ability, although Hermalin and Weisbach allow for other interpretations. As before, the information owners receive permit them to form an unbiased estimate of α , $\hat{\alpha}$. Although Hermalin and Weisbach allow for a number of different functions $\gamma(\cdot)$ and $c(\cdot)$, one that fits with earlier analysis is that $a \in \{0, 1\}$, keep or fire the CEO, respectively; $c(1) > c(0)$, reflecting firing costs; and $\gamma(a) = 1 - a$, so if the CEO is fired, the owners expect the return a replacement generates (normalized to zero), and otherwise they expect the return equal to their best estimate of his ability, $\hat{\alpha}$.

Define

$$a^*(\hat{\alpha}) = \operatorname{argmax}_{a \in \mathbf{A}} \hat{\alpha}\gamma(a) - c(a); \tag{48}$$

that is, $a^*(\hat{\alpha})$ is the owners' best response to their information.⁵⁷ Define

$$\Pi(\hat{\alpha}) = \hat{\alpha}\gamma(a^*(\hat{\alpha})) - c(a^*(\hat{\alpha}));$$

that is, $\Pi(\hat{\alpha})$ is the owners' expected payoff conditional on their estimate of the parameter α .

Lemma 5. *The payoff function $\Pi(\cdot)$ is convex.*

Lemma 5 entails that the owners are risk *loving* with respect to the estimator $\hat{\alpha}$, which recall is a random variable *ex ante*. The (unconditional) mean of $\hat{\alpha}$ is $\mathbb{E}\{\alpha\}$, so invariant across disclosure regimes; hence, owners prefer, *ceteris paribus*, a disclosure regime in which the distribution of $\hat{\alpha}$ is riskier, in the sense of second-order stochastic dominance, than a regime in which the distribution is less risky. Critically, an estimator based on better information in the Blackwell sense is a riskier estimator in the sense of second-order stochastic dominance (see Baker, 2006); intuitively, receiving no information—minimal informativeness—makes $\hat{\alpha}$ invariant, it remains equal to the prior. Receiving information makes $\hat{\alpha}$ vary, which means risk. It therefore follows from Lemma 5 that

Proposition 19. *If \mathcal{D} is a more informative disclosure regime, in the Blackwell sense, than \mathcal{D}' , then the owners prefer \mathcal{D} to \mathcal{D}' , ceteris paribus.*

Proposition 19 establishes half of Condition 1. To establish the other half, more structure is required. Among the possible set of assumptions, suppose that the CEO loses L if dismissed (similar to Section 2.2). Observe that the owners

⁵⁶Among the other models they consider are a managerial myopia model with empire building; a hidden-information agency model; and a hidden-action agency model.

⁵⁷Generically, $a^*(\hat{\alpha})$ is unique, so we treat it as such (*i.e.*, we assume an equality in (48) rather the technically more correct \in).

maximize (47) by choosing $a = 1$ if and only if $\hat{\alpha} < -c(1)$. Let $F(\cdot|\mathcal{D})$ denote the distribution function of the estimator $\hat{\alpha}$ conditional on the disclosure regime being \mathcal{D} . It follows that $U(\mathcal{D}) = -LF(-c(1)|\mathcal{D})$ in this version of the model.

We also require a bit more structure on the relations between disclosure regimes: suppose that the set of disclosure regimes can be ordered by the dispersive order: for any two disclosure regimes \mathcal{D} and \mathcal{D}' either $F(\cdot|\mathcal{D})$ dominates $F(\cdot|\mathcal{D}')$ in the dispersive order (denoted $F(\cdot|\mathcal{D}) \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D}')$) or the other way around.⁵⁸ Because all distributions of $\hat{\alpha}$ have the same mean, $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ implies $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{ssd}} F(\cdot|\mathcal{D})$, where $\stackrel{\geq}{\text{ssd}}$ denotes second-order stochastic dominance (Shaked and Shanthikumar, 1994, Theorem 2.B.10). It follows from Lemma 5 that $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ implies the owners prefer regime \mathcal{D} to regime \mathcal{D}' . Under reasonable conditions, the CEO has the opposite preferences:⁵⁹

Proposition 20. *Suppose the median of the estimate $\hat{\alpha}$ coincides with the mean and exceeds $-c(1)$. Then $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ implies the owners prefer regime \mathcal{D} to regime \mathcal{D}' and the CEO prefers \mathcal{D}' to \mathcal{D} . In other words, Condition 1 holds.*

Intuitively, $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ means that $F(\cdot|\mathcal{D})$ has a fatter left tail than does $F(\cdot|\mathcal{D}')$. Because it is left-tail outcomes that get him fired, the CEO prefers thinner left tails to fatter left tails, *ceteris paribus*.

Finally, it's worth tying this more general analysis back to the normal learning model employed elsewhere in this chapter. As often assumed above, let $\alpha \sim N(0, 1/\tau)$ and assume the owners receive a signal $s \sim N(\alpha, 1/\eta)$. Because $0 > -c(1)$ and the median of the distribution of $\hat{\alpha} = 0 = \mathbb{E}\{\hat{\alpha}\}$, the suppositions in Proposition 20 are met. Consider two disclosure regimes: in one $s \sim N(\alpha, 1/\eta)$, in the other $s \sim N(\alpha, 1/\eta')$, where $\eta > \eta'$; that is, the first is more informative than the second. We wish to show that the distribution of $\hat{\alpha}$ under the first regime is dominated by the second in the sense of the dispersive order. Recall (see, *e.g.*, (4)) that $\hat{\alpha} = \eta s / (\tau + \eta)$. Hence

$$\text{Var}(\hat{\alpha}) = \frac{\eta^2}{(\tau + \eta)^2} \text{Var}(s) = \frac{\eta^2}{(\tau + \eta)^2} \times \left(\frac{1}{\tau} + \frac{1}{\eta} \right) = \frac{\eta}{\tau(\tau + \eta)}$$

(note this is the unconditional variance of $\hat{\alpha}$). Simple calculus reveals that $\partial \text{Var}(\hat{\alpha}) / \partial \eta > 0$. In light of Lemma A.3 in the appendix, it follows that

$$N\left(0, \frac{\eta'}{\tau(\tau + \eta')}\right) \stackrel{\geq}{\text{disp}} N\left(0, \frac{\eta}{\tau(\tau + \eta)}\right).$$

Proposition 20 therefore entails:

⁵⁸ $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ if

$$F^{-1}(\xi|\mathcal{D}') - F^{-1}(\xi'|\mathcal{D}') < F^{-1}(\xi|\mathcal{D}) - F^{-1}(\xi'|\mathcal{D})$$

whenever $1 > \xi > \xi' > 0$, where F^{-1} is the inverse of the distribution function.

⁵⁹The requirement that the mean and median of the estimator coincide is a property of many estimation procedures. The condition that $\mathbb{E}\{\hat{\alpha}\}$, which recall equals $\mathbb{E}\{\alpha\}$, exceed $-c(1)$ should be non-controversial because the alternative would mean the owners would wish to fire the CEO in the *absence* of information, which begs the question of why they would have hired him in the first place.

Corollary 2. *In a CEO-dismissal model, if $\hat{\alpha}$ derives from a normal-learning model, then owners prefer more precise signals (i.e., higher η s), while the CEO prefers less precise signals, ceteris paribus; that is, Condition 1 holds.*

Bibliographic Note. As noted, Hermalin and Weisbach (2012) consider the tradeoff reflected in Condition 1 across a number of models relevant to governance. Additionally, they study how firm characteristics affect the choice of disclosure regime. In particular, they show that larger firms will tend, *ceteris paribus*, to have better disclosure regimes and, thus, greater executive compensation. Moreover, via a general equilibrium analysis of the CEO labor market, they show firms with better disclosure regimes employ more able CEOs. They also explore the consequences of mandated (*e.g.*, by legislation) reforms of disclosure and its consequences for managerial compensation and well being.

7 HOW NON-BAYESIAN UPDATING AFFECTS GOVERNANCE

A maintained assumption of most learning models is that individuals incorporate new information rationally; that is, they update their beliefs according to Bayes Law. There is, however, a large body of research in psychology that convincingly demonstrates that people often hold beliefs or take actions that are inconsistent with their having properly employed Bayes Law to account for new evidence.⁶⁰

If, as the psychological evidence suggests, people fail to apply Bayes Law correctly (are “bad Bayesians”), what does this mean for the realism of the learning-based models previously discussed? One view is that they are invalidated, since their key underlying assumption is wrong. This is too strong a response: although the average person may incorrectly calculate posterior probabilities in the face of new information, these models do not rely on precise calculations; rather, they rely on the direction in which beliefs are updated. As long as bad news shifts beliefs about ability downward and good news shifts them up, then the spirit of the models is preserved. Moreover, the models allow us to say what the consequence of particular cognitive biases will be. Specifically, keep in mind that the sign of the bias when updating beliefs is generally understood: in particular, people put too much weight on their most immediate observations and experiences and too little on base rates (as we discuss below). A consequence is that people are likely reacting to current signals more than rational Bayesian updating would imply. In a sense, the models’ implications are likely to be greater in magnitude in reality than on paper. For this reason, the behavioral biases documented by psychologists imply that the models previously considered could *understate* the true importance of the effects they characterize.

This section expands on that idea; in particular, we explore the implications of biased (non-Bayesian) updating for learning-based models of governance. We first summarize some of the literature on behavioral biases. Then we discuss the

⁶⁰Some good introductions and overviews of this literature are Gilovich (1991), Plous (1993), and Kahneman (2011).

way that the behavioral biases potentially affect learning models of governance, as well as the way in which behavioral analysis could be incorporated into this analysis in future research. Finally, we suggest some ways in which this type of behavioral analysis could be incorporated into future research.

7.1 HOW DO PEOPLE REALLY UPDATE THEIR BELIEFS?

The psychology literature has documented a number of biases or fallacies that pertain to how individuals tend to depart from rationality in their decision-making. Three such departures are especially relevant in the context of this chapter: the *base-rate fallacy*, the *“hot-hand” fallacy*, and the *fundamental-attribution bias*.⁶¹

The base-rate fallacy is a tendency to *underweight* base rates; in terms of Bayes Law, when agents receive a signal, they revise their beliefs by more than Bayes Law would have them do. Numerous experiments have given test subjects information about the population (the base rate), and then subsequent information that can be used to answer a question. This is especially relevant with regard to experiments considering tests for rare diseases: tell the subjects that the prevalence of some disease is, say, one in 10,000 in the population and there is a test for that disease that has only a one-percent false positive rate and a very high (perhaps even perfect) true positive rate. Then ask the subjects how likely is it that a patient who tests positive has the disease. The subjects’ guesses are usually very high, often over 90%.⁶² The true answer, however, is less than one percent.⁶³ In other words, individuals underweight the base rate (the remarkably low prevalence of the disease) and place too much weight on new information (the signal—the test result).⁶⁴

The hot-hand fallacy, somewhat relatedly, is a tendency to forecast future success based on a recent run of success. Gilovich et al. (1985) originally documented this tendency using data from professional basketball games. Their

⁶¹See Kahneman (2011) for, *inter alia*, an overview of these biases.

⁶²This has been the experience of one of the authors who has routinely run this experiment in his first-year MBA course.

⁶³If p is the true positive rate, then Bayes Law yields a posterior probability of having the disease based on a positive test as

$$\frac{p \times \frac{1}{10,000}}{p \times \frac{1}{10,000} + \frac{9999}{10,000} \times \frac{1}{100}} < \frac{1}{1 + 99.99} \approx .0099.$$

⁶⁴A related fallacy is the *conjunction fallacy* (Tversky and Kahneman, 1983): Linda is 31, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable? 1) Linda is a bank teller or 2) Linda is a bank teller and active in the feminist movement. In the Tversky and Kahneman study, the majority of subjects chose the second option even though the event in that option is a proper subset of the event in the first; that is, if \mathcal{B} is the event Linda is a bank teller and \mathcal{F} the event she’s a feminist, then the first option is \mathcal{B} and the second is $\mathcal{B} \cap \mathcal{F} \subseteq \mathcal{B}$ and, necessarily, $\Pr\{\mathcal{B} \cap \mathcal{F}\} \leq \Pr\{\mathcal{B}\}$.

statistical analysis reveals that a player’s likelihood of making a given shot is unrelated to whether he made his most recent shots. Nevertheless, coaches, players, and fans all believe strongly that the opposite is true, that the probability of a player making his next shot increases substantially if he has hit his last two or three shots. Gilovich et al. argue that this misperception is evidence of a bias in which people over-extrapolate the most recent events and underweight other relevant data.

This finding was very controversial, and has led to a large literature reëxamining Gilovich et al.’s conclusions. However, as emphasized by Alter and Oppenheimer (2006), much of this literature is actually about the statistics of basketball, using more sophisticated estimation strategies to see whether, in fact, the likelihood of making shots is independent. Little of the subsequent research is on the more important issue (at least to non-basketball fans) of whether people systematically and substantially overestimate the implications of past success on future success (see, *e.g.*, Bocskocsky et al., 2014). Alter and Oppenheimer argue that regardless of whether there is serial correlation in shot-making success, agents dramatically overstate the implication that making one shot has on the probability of making the next. Moreover, data unrelated to sports confirm that people tend to believe that one success increases the probability of future success far more than is warranted by data. Therefore, there appears to be a systematic tendency to overestimate the implications of recent events on the likelihood of future events.⁶⁵

The fundamental attribution bias is to assign more credit or blame to an individual and less to her circumstances than is appropriate. This behavior was first documented in experiments by Jones and Harris (1967), and the term coined by Ross (1977). This bias represents a departure from Bayesian inference because the rational inference from a person’s action is to update one’s prior about his ability based on his actions conditional on the circumstances the person faces. However, people fail, on average, to adjust appropriately for circumstances outside one’s control and give people too much credit or blame than they deserve for a particular outcome.

7.2 IMPLICATIONS OF BIASES FOR INTERPRETATION OF LEARNING MODELS

7.2.1 The Base-Rate Fallacy

It is relatively straightforward to incorporate the base-rate fallacy into learning models of governance. As was shown repeatedly above (consider, *e.g.*, expressions (4) and (19)), the true Bayesian estimate of ability is a weighted average of new information and the prior (for instance, in (19), the weights are—dropping subscripts—

$$\frac{\eta}{\tau + \eta} \text{ and } \frac{\tau}{\tau + \eta},$$

⁶⁵A related bias is the *gambler’s fallacy*: the tendency of people to think that luck will reverse itself. If a gambler has a run of bad luck, people will think s/he is “due” some good luck, even though the random events are, just that, random, and do not systematically reverse themselves.

respectively). A shift in weight toward new information over the prior will increase a CEO's incentive to jam the signal. For example, consider, expression (34): if $\eta/(\tau+\eta)$ increases, then the CEO's marginal return to the signal-jamming action, ξ , is greater and so his choice—the solution to (34)—will be greater. So if we think of the base-rate fallacy as meaning those assessing the CEO have a downward bias in their estimate of τ (or, equivalently, an upward bias in their estimate of η), then the CEO, if he understands this bias (or simply suffers it himself), will be more inclined to jam the signal than were this a world of unbiased Bayesians. This insight has a number of implications *vis-à-vis* models that assume actors are perfect Bayesians: if the base-rate fallacy is a prevalent bias, then 1) the career-concern incentives facing managers would be greater; 2) the distortions in investment identified by Holmstrom (1999) could be larger; 3) the reaction (positive or negative) to any news releases in terms of estimates of managerial ability would be greater; and 4) under at least some assumptions, it would, as just suggested, mean managers' motives to take myopic investments are bigger.

The base-rate fallacy should be kept in mind when interpreting empirical work about learning. For example, consider the estimates of indirect incentives in the private equity and hedge fund industries summarized in Table 1. The mechanism presumed to underlie these estimates is that a fund's performance leads the market to update its assessment of the ability of that fund's manager, which affects inflows into the fund. The fees on these new inflows (or outflows) determine the incremental pay the managers receive from their performance. The magnitude of this effect clearly depends on the extent to which the market updates their assessment of the manager's ability. So if the base-rate fallacy leads the market to have larger updates for a given level of performance than Bayes Law would imply, there would be more inflows to the fund for that level of performance and consequently higher indirect incentives.

When interpreting these estimates, it is impossible to disentangle rational Bayesian updating from an overreaction due to behavioral biases. There are, however, reasons to suspect behavioral biases are present. The magnitude of the estimates of indirect incentives is very large, potentially larger than one might expect only from rational updating. For private equity funds, Chung et al. (2012) find that indirect incentives are about the same size as the 20% carried interest, which by itself is substantial source of incentives for private-equity general partners. For hedge funds, for which moving capital is easier than with private-equity funds, Lim et al. (2016) find that, on average, indirect incentives are several times as large as they are for private-equity funds. For young funds, indirect incentives are so large that an additional dollar of wealth returned to investors appears to increase the fund manager's lifetime personal wealth by more than a dollar. It is possible that these effects could be generated by rational updating, after all a small difference in estimated ability between fund managers could lead a rational investor to switch funds. However, the estimated effects are sufficiently large that it seems likely that some of the incentives could be a result of investors overreacting to fund performance.

Relatedly, the magnitude of the investors' responses to performance in terms

of fund inflows and outflows is very large. The strong relation between inflows and performance, together with the lack of abnormal performance for mutual funds, has long been considered an important puzzle in finance (see *e.g.*, Berk and Green, 2004). It is possible that some of this reaction could be explained by the base-rate fallacy: investors place too much weight on recent performance and consequently adjust their portfolios more than they should in response to a shock to performance. In fact, Lim et al. (2016) find that inflows respond strongly to monthly performance; it is difficult to imagine that a rational update of ability could change so rapidly in response to short-term fluctuations in performance.⁶⁶

7.2.2 The Hot-Hand Fallacy

In one sense, the hot-hand fallacy is an extreme version of the base-rate fallacy. In basketball, for instance, it turns out that a player's field-goal percentage is roughly constant over time (*e.g.*, over a season or two) and there is little to no serial correlation within a particular game. Hence, the right "model" is that each shot a player takes has an independent probability p of being good, so p should be the probability with which one should expect the next shot to hit, even if the player has enjoyed a run of success (just as one knows a fair coin will land heads with probability $1/2$ no matter how many heads in a row have just been tossed). The hot-hand fallacy is, in essence, the false belief that there is something to be learned when there is, in fact, nothing to learn—those subject to the fallacy can be seen as assigning a weight less than one to the prior (base rate) even though, here, one is the correct weight to assign.

Another way to view the hot-hand fallacy is that people are using the wrong stochastic model. Suppose that, as is standardly assumed in the literature, a CEO's ability is fixed over time. If, instead, one mistakenly believed that it followed a random walk (or other process with serial correlation), then one would believe that there is more to be learned from recent events than there truly is. In this regard, it is worth recalling, as noted earlier, that Holmstrom (1999) considers an extension of the basic career-concerns model that encompasses abilities that follow a random walk. Although the resulting formulae are different in this version of the model than the one with fixed ability, the basic conclusions discussed above continue to hold—*at least if abilities truly follow a random walk and this is commonly understood*. What could prove interesting is to explore the deviations in beliefs, efforts, compensation, and so forth when people are subject to the hot-hand fallacy (falsely believe, *e.g.*, ability follows a random walk). To the best of our knowledge, no one has extended the career-concern model to allow for the possibility that some or all observers are using an erroneous model of the underlying stochastic process.

⁶⁶The published version does not contain these monthly regressions. They are, however, reported in the working paper version, NBER WP 18903.

7.2.3 The Fundamental Attribution Bias

The fundamental-attribution bias is somewhat trickier. If there is only one signal of performance, then there is no means of identifying what is attributable to the manager and what to circumstances. In some settings, though, inferences could be made about circumstances by comparing the performance of a given firm to others in the same industry. Differences in relative performance could help identify the degree to which the firm's success (or lack thereof) should be attributable to its CEO and the degree attributable to circumstances beyond his control.⁶⁷ The extent to which markets appropriately adjust for circumstances beyond the CEO's control is somewhat unclear: although there is some evidence that relative performance is taken into account (Gibbons and Murphy, 1990), there is also evidence that it is not (Bertrand and Mullainathan, 2001).

Another issue is that if managers know they will be assessed in a relative way, then, consistent with the general point that the act of observation can distort incentives, such assessment could lead to suboptimal behavior. For instance, Scharfstein and Stein (1990) present a model in which relative-performance assessment results in herding in investment decisions: as with wildebeests on the Serengeti, standing out from the herd is a good way to get picked off by predators. But while herding can be good for wildebeests and CEOs, it can be quite suboptimal from the perspective of shareholders.

Despite the difficulties of identifying managerial ability from circumstances, there are reasons to suspect that markets give too much credit to managers. In the management literature, Khurana (2002a,b) has made a compelling case that there is a misguided tendency to treat a number of CEOs as superstars. In addition, experimental work suggests that subjects have a tendency to overweight the importance of the leader to a team's success or failure (see, in particular, Weber et al., 2001).

7.3 POTENTIAL FUTURE RESEARCH APPLYING BEHAVIORAL ECONOMICS TO CORPORATE GOVERNANCE

Although, as discussed above, there is a way in which departures from Bayesian updating don't change the basic point of models that assume wholly rational actors, it is nonetheless worth speculating about what the implications for behavioral economics might be for learning-based models of corporate governance.

Note, first, that it is not necessarily obvious what the effect of these biases will be. For example, the base-rate fallacy causes observers to overweight current performance, which could increase a young executive's incentives relative to what they would be if observers were perfect Bayesians. On the other hand, if the past will be effectively forgotten, the long-run effects that current effort has on future compensation will be less than a model based on Bayesian updating would predict, which could reduce a young executive's incentives relative to the model. Hence, one avenue of research is to determine the theoretical

⁶⁷Wasserman et al. (2010) find evidence that who the CEO is does matter, but this effect varies considerably across industries.

consequences if the relevant parties employ updating/inference rules exhibiting well-documented biases (*e.g.*, will they boost or lessen managerial effort early in a career).

A second set of questions stem from the insight that if these biases are predictable and lead others to behave sub-optimally, then ought there not be ways for one to make money off of them? That is, why can't savvy players effectively "print money" at the expense of the less savvy? One possible answer is to imagine that there is some process—as yet unmodeled in this context (to the best of our knowledge)—that causes competition among the savvy players seeking to capture such rents to somehow drive the relevant institutions to an equilibrium in which it is *as if* the institutions were run by rational actors.

Alternatively, it might actually be profitable in some ways to have a board made up of bad Bayesians. To wit, consider equation (17) and, to eliminate the distortion introduced by discounting, set $\delta = 1$. If all boards suffered from the base-rate fallacy so severely that they mistakenly set $\tau_t = 0$ (*i.e.*, their estimate of ability is wholly governed by current performance), then a CEO's return to effort is 1, matching the social return: in a world of biased boards, the first best attains. Admittedly, this is a highly incomplete analysis—and certainly one can also imagine reasons why a board of bad Bayesians is bad—but it serves to illustrate one way in which cognitive biases could be a plus.

Another way cognitive biases could be a plus is by serving as a commitment device.⁶⁸ To illustrate this possibility, another well-known learning-related bias is *confirmation bias*, the tendency to look for or overweight evidence confirming a decision and to ignore or underweight contrary evidence. Consider two boards, one with confirmation bias, one without. Suppose, as in Section 4.1, CEOs decide whether or not to pursue projects. The CEO working for the board without bias will, for the reasons detailed in conjunction with Proposition 5, never undertake a project. The story could be different for the CEO working for the board with bias. If this board wishes to confirm the wisdom of its selection of the CEO, it will tend to overlook or at least underweight a project failure, while noticing and overweighting a project success. Although choosing a project still exposes the CEO to risk, the magnitude of the risk-adjusted upside can exceed the magnitude of the risk-adjusted downside, which means he could wish to undertake projects. In short, the boards confirmation bias could commit it to being more lenient toward failure, which is necessary if the CEO is to have an incentive to pursue projects at all. Again, this is a highly incomplete analysis—and certainly one can also imagine reasons why a lenient board is bad—but it is suggestive of possible lines of future research.

Empirically, there is much work to be done. Even assuming rationality, there is a dearth of empirical work related to learning and its effect on governance. Identifying the extent to which observed learning represents rational Bayesian updating and to which it represents overreaction because of behavioral biases is

⁶⁸The discussion that follows is somewhat related to a strain of the leadership literature that deals with delegation. See, *e.g.*, Rotemberg and Saloner (1993, 1994, 2000), Van den Steen (2005), and Blanes i Vidal and Möller (2007).

clearly a very important (and difficult) question.

The existence of behavioral biases raises an important methodological issue regarding updating. Structural estimation, which has become popular recently, takes the underlying modeling as a maintained assumption while estimating the underlying model parameters. The fact that agents are likely to be overreacting to new information, suggests that the approach of *assuming* that Bayes Law holds as one of the maintained assumptions is potentially problematic.

Another important issue is that biases are typically measured via lab experiments using a somewhat random draw of local undergraduates. Do cognitive biases evidenced by students playing for small stakes in psychology labs necessarily have bearing in a corporate world in which governance mistakes have the potential to destroy billions of dollars? Put slightly differently, Bayes Rule is not a secret formula and the incentives to get matters right are huge, so shouldn't boards and others overcome these biases? As noted, there are some hints in the empirical literature that the answer may nonetheless be "no," but obviously how accurately the relevant actors adjust their beliefs is an empirical question, one in need of more research.

In summary, there is a natural complementarity between the learning and assessment approach to governance and behavioral economics. Although, as noted, it is possible that market forces drive the relevant actors to behave as if they are good Bayesians, there are also many reasons to suspect that managerial ability is assessed in such a way that the relevant actors are updating their beliefs in a decidedly non-Bayesian manner. If so, then our understanding of the corresponding governance phenomena will be vastly enriched by incorporating the various insights arising from research on cognitive biases.

8 CONCLUSION

Research in corporate governance has historically focused on issues of control: getting those with decision-making authority to make decisions that are as aligned as possible with the interests of investors. Without minimizing the importance of these agency problems, this chapter has sought to make the case that such a focus is incomplete. An equally important role of corporate governance is to ensure that those with decision-making authority are the best people to hold such authority. This means assessing the capabilities of decision makers—that is, management—and the strategies they are deploying. In the past few decades, a research program has emerged among corporate governance scholars that seeks to understand not only such assessment directly, but also its ancillary effects on the behavior of managers, their compensation, and the structure of governance itself.

One consequence of assessment is that the actor being assessed can be tempted to take actions to affect that assessment. Starting with Holmstrom (1982) [1999], it has been recognized that assessment creates incentives. The incentives can be positive; they can, for example, induce a manager to work harder. They can also be negative; they can, for example, lead a manager to eschew a positive NPV project or behave in other myopic ways. In this chapter,

we have reviewed a number of models that reflect the good, the bad, and the ugly of the incentives created by assessment.

Beyond its effect on managerial behavior, assessment is also important because it affects the desirability of the manager's job, particularly how risky it is. Because what others will ultimately infer about him or her is uncertain *ex ante*, a manager is necessarily exposed to risk. People demand additional compensation to bear risk, so a cost of intense monitoring of a manager is the increased compensation s/he will demand.

Assessment also has implications for how governance evolves. Because of the downside risk in being evaluated, a manager has reasons to seek less vigilant monitors, which in turn has implications for how independent the board of directors will be. Because the ability to learn about a manager and dismiss him if he is a poor performer creates an option value, employers (boards) can have an incentive to hire executives about which there is uncertainty because that raises the option value. How valuable that option value is, however, depends on how vigilant a monitor the board is. Hence, there are relationships between board independence (propensity to monitor) and the sort of executive employed. Boards and executives may also wish to bargain about the kind of information available to the board. In short, as discussed above, assessment entails a number of consequences for who is hired, how she is monitored, who monitors her, and the information that will be used in assessing her. An assessment perspective on governance can thus prove essential for gaining insights into phenomena such as executive selection and turnover; the evolution of executive salaries and other forms of non-contingent compensation; and changes in the balance of power between boards and CEOs.

Not only is there assessment by the board (employers), but also the market more generally. The more uncertainty there is about a CEO, the more influential any signal about his ability will be. As discussed above, this can help to explain the volatility of stock returns over time: the more established a CEO, the less of an impact any performance metric has in assessing his ability, and thus the less the market will react to such news.

Although by no means comprehensive, our discussion of the extant literature on the assessment-based approach to corporate governance illustrates that this approach has already yielded considerable insights. Yet it is not a spent force: it remains a rich approach for future research. This is especially true for empirical work. Probably the greatest challenge to empirical work in this area is data given the key variables in the models tend to be unobservable to an outsider: the assessment itself, managerial effort, managerial career paths, and the question of whether an observed action is distortionary (how, for instance, does one tell if a particular investment is myopic)? Yet, the rewards to empirical work are likely to be large, since the effects the models describe seem *a priori* very important. Acting to improve others' opinions of oneself is a constant element of everyday life, and it seems likely that such actions provide important motivations and indirect consequences in ways we do not fully understand.

A striking feature of the models in this literature is how well suited they are to testing: as a rule, they offer sharp time-series and cross-sectional predictions.

These predictions occur, in large part, because Bayes Theorem is very clear on the way updating occurs. The magnitude of updating should be negatively related to the precision of the prior information, positively related to the amount of information in the signal, and decrease in magnitude over time given that beliefs about ability are, necessarily, more fixed for old-timers than newcomers. These predictions fall out of from virtually every model based on updating and, with appropriate data and clever empirical designs, are likely to lend themselves to interesting and important empirical work.

One potential shortcoming of the current literature is that it has ignored the well-documented fact that most people are “bad Bayesians”—they fail, in systematic and predictable ways, to update their beliefs in a manner consistent with Bayes Theorem. In particular, the psychology literature has documented in numerous experiments that people tend to overweight signals relative to prior information, leading them to update their beliefs when they receive new information by more than is warranted. As we discussed in detail above (see in particular the previous section), recognition of this fact opens a number of avenues of research, both empirical and theoretical. If, as the psychological literature suggests, actors react to signals more than they should, then many of the effects identified in models with wholly rational actors are magnified. For this reason, we believe that assessment-based models are one area of economics where behavioral considerations are likely to cause the rational models to understate the effects they describe. A critical issue in this regard is determining empirically how far real-world belief formation deviates from the Bayesian ideal in the context of governance.

In sum, learning and assessment complement agency theory as tools for the study of corporate governance; in particular, they are a means to provide a unifying framework to a number of disparate phenomena in governance. Further, it is an approach amenable to empirical analysis and, indeed, a rich empirical literature has started to emerge that builds implicitly and explicitly on it. Finally, it offers natural ways to incorporate behavioral insights into the study of governance and contracting.

APPENDIX A: TECHNICAL DETAILS

Derivation of expression (9). Observe

$$\begin{aligned} \mathbb{E}\{\max\{-r, \hat{\alpha}_1\}\} &= \int_{-\infty}^{\infty} \max\left\{-r, \frac{\eta s}{\tau_0 + \eta}\right\} \sqrt{h} \phi(s\sqrt{h}) ds \\ &= -r \int_{-\infty}^{-r\tau_0/h} \sqrt{h} \phi(s\sqrt{h}) ds + \int_{-r\tau_0/h}^{\infty} \frac{sh}{\tau_0} \sqrt{h} \phi(s\sqrt{h}) ds. \end{aligned}$$

The last step in deriving (9) follows from the following lemma:

Lemma A.1. *If ϕ is the density function of a standard normal, then*

$$\int_a^b x\phi(x)dx = \phi(a) - \phi(b).$$

Proof: Recall $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$. Hence,

$$\begin{aligned} \int_a^b x\phi(x)dx &= \int_a^b x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \Big|_a^b \\ &= \phi(a) - \phi(b). \end{aligned}$$

■

Lemma A.2. *$N(\mu, \sigma_a^2)$ dominates $N(\mu, \sigma_b^2)$ in the sense of second-order stochastic dominance if $\sigma_b > \sigma_a$.*

Proof: Recall that if $x \sim N(\mu, \sigma^2)$, then its cumulative distribution function can be written as $\Phi((x-\mu)/\sigma)$, where Φ is the cumulative distribution function of the standard normal (*i.e.*, $N(0, 1)$). Recall, too, that

$$\Phi(x) = 1 - \Phi(-x) \tag{49}$$

for all x (the tails have equal mass).

As a preliminary,

$$\begin{aligned} &\int_{-\infty}^{\infty} \left(\Phi\left(\frac{x-\mu}{\sigma_a}\right) - \Phi\left(\frac{x-\mu}{\sigma_b}\right) \right) dx = \\ &\int_{-\infty}^{\mu} \left(\Phi\left(\frac{x-\mu}{\sigma_a}\right) - \Phi\left(\frac{x-\mu}{\sigma_b}\right) \right) dx - \int_{\mu}^{\infty} \left(\Phi\left(-\frac{x-\mu}{\sigma_a}\right) - \Phi\left(-\frac{x-\mu}{\sigma_b}\right) \right) dx \\ &= \int_{-\infty}^{\mu} \left(\Phi\left(\frac{x-\mu}{\sigma_a}\right) - \Phi\left(\frac{x-\mu}{\sigma_b}\right) \right) dx - \int_{-\infty}^{\mu} \left(\Phi\left(\frac{z-\mu}{\sigma_a}\right) - \Phi\left(\frac{z-\mu}{\sigma_b}\right) \right) dz \\ &= 0, \end{aligned} \tag{50}$$

where the first equality follows from (49) and the second from making the change of variables $z = 2\mu - x$.

Recall distribution F dominates G in the sense of second-order stochastic dominance if they have the same mean and

$$\int_{-\infty}^y (F(x) - G(x))dx \leq 0$$

for all y . Because Φ is an increasing function and $(x-\mu)/\sigma_a < (x-\mu)/\sigma_b$ for all $x < \mu$,

$$\int_{-\infty}^y \left(\Phi\left(\frac{x-\mu}{\sigma_a}\right) - \Phi\left(\frac{x-\mu}{\sigma_b}\right) \right) dx < 0 \tag{51}$$

for all $y \leq \mu$. The lefthand side of (51) goes to zero as $y \rightarrow \infty$ by (50). Hence the lemma must hold if the lefthand side of (51) is monotone in y for $y > \mu$. The derivative of that lefthand side is

$$\Phi\left(\frac{y-\mu}{\sigma_a}\right) - \Phi\left(\frac{y-\mu}{\sigma_b}\right) > 0$$

for all y , where the sign follows because Φ is increasing and $(y-\mu)/\sigma_a > (y-\mu)/\sigma_b$ for all $y > \mu$. ■

Proof of Lemma 2: Claims (ii) and (iv) are immediate. Claim (iii) follows because

$$\frac{d}{d\eta} m_t = \sum_{j=1}^{T-t} \delta^j \frac{\tau_0}{(\tau_0 + (t+j-1)\eta)^2} > 0.$$

For claim (i), it was shown earlier that $m_{T-1} > m_T = 0$. Consider $t < T-1$. Observe

$$m_t = \delta^{-t} \sum_{i=t+1}^T \delta^j \frac{\eta}{\tau_{i-1}}$$

if one reindexes (note the use of (5)); hence,

$$\begin{aligned} m_t - m_{t+1} &= \delta \frac{\eta}{\tau_t} - \delta^{-(t+1)} (1-\delta) \sum_{i=t+2}^T \delta^j \frac{\eta}{\tau_{i-1}} \\ &> \delta \frac{\eta}{\tau_t} - \delta^{-(t+1)} (1-\delta) \frac{\eta}{\tau_t} \sum_{i=t+2}^T \delta^j = \delta^{T-t} \frac{\eta}{\tau_t} > 0, \end{aligned}$$

where the first inequality follows because $\tau_{t''} > \tau_{t'}$ if $t'' > t'$.⁶⁹ ■

Proof of Lemma 3: Noting that

$$(Y - \hat{\alpha})\sqrt{H} = -\frac{\tau + \eta}{\eta} \hat{\alpha} \sqrt{H} = -\hat{\alpha} \frac{\tau}{\sqrt{H}}$$

and utilizing Lemma A.1, it is readily seen that $\partial V / \partial \hat{\alpha} = \Phi(-(Y - \hat{\alpha})\sqrt{H}) > 0$. Consequently,

$$\frac{\partial(V - \hat{\alpha})}{\partial \hat{\alpha}} = \Phi(-(Y - \hat{\alpha})\sqrt{H}) - 1 < 0. \quad (52)$$

■

⁶⁹Recall

$$\sum_{k=n}^N \delta^k = \frac{\delta^n - \delta^{N+1}}{1 - \delta}.$$

Proof of Proposition 8: The first-order condition for maximizing the board's objective function is

$$V - \max\{0, \hat{\alpha}\} - \ell d'(p) = 0. \quad (53)$$

Given the usual comparative statics, part (i) follows immediately from Lemma 3. Part (ii) follows, via the usual comparative statics, if $\partial V/\partial \tau < 0$. Recall $(Y - \hat{\alpha})\sqrt{H} = -\hat{\alpha}\tau/\sqrt{H}$, so, from (37),

$$V = \Phi\left(\hat{\alpha}\frac{\tau}{\sqrt{H}}\right) + \frac{\sqrt{H}}{\tau}\phi\left(-\hat{\alpha}\frac{\tau}{\sqrt{H}}\right).$$

Using Lemma A.1 and the fact that $\phi(\cdot)$ is symmetric about 0, it is readily shown that

$$\frac{\partial V}{\partial \tau} = \phi\left(-\hat{\alpha}\frac{\tau}{\sqrt{H}}\right)\frac{\partial\sqrt{H}/\tau}{\partial \tau} = -\phi\left(-\hat{\alpha}\frac{\tau}{\sqrt{H}}\right)\frac{\eta + 2\tau}{2\eta\tau^3}H^{3/2} < 0.$$

Hence, by the usual comparative statics, part (ii) follows. Part (iii) is immediate from (53) (a shift up in marginal cost reduces the equilibrium action). Finally, it is readily shown that

$$\frac{\partial V}{\partial s} = \frac{\tau}{2(\tau + \eta)\sqrt{H}}\phi\left(-\hat{\alpha}\frac{\tau}{\sqrt{H}}\right) > 0.$$

So part (iv) follows from the usual comparative statics. ■

Proof of Lemma 4: Consider bargaining with a new CEO. If this bargaining is unsuccessful, the board can hire yet another CEO, who is *ex ante* identical to the new CEO. Hence, Nash bargaining entails choosing P^* and w to maximize

$$(P^*V_0 - \ell_0 d(P^*) - w - U_0) \times \left(\underbrace{(P^*\Phi(0) + 1 - P^*)}_{=1/2} b + w \right).$$

The respective first-order conditions are

$$(V_0 - \ell_0 d'(P^*)) \left(\left(1 - \frac{1}{2}P^*\right) + w \right) - \frac{1}{2}(P^*V_0 - \ell_0 d(P^*) - w - U_0)b = 0 \quad (54)$$

and

$$P^*V_0 - \ell_0 d(P^*) - w - U_0 - \left(\left(1 - \frac{1}{2}P^*\right) + w \right) \leq 0. \quad (55)$$

In equilibrium, one new CEO must yield the board the same utility as any other new CEO; hence,

$$P^*V_0 - \ell_0 d(P^*) - w = U_0.$$

Consequently, (55) implies that $w = 0$. Plugging all this back into (54), we have $V_0 - \ell_0 d'(P^*) = 0$. That is the first-order condition for the utility-maximizing intensity of monitoring for a board with lack of independence ℓ_0 . Because P^* is

monotone in ℓ , it follows that $\ell_1 = \ell_0$. ■

Proof of Proposition 10: Let U^E denote the equilibrium expected utility of the board if it bargains with the incumbent CEO. Recalling that $P^*(\ell_0)$ is the board's most preferred level of monitoring, it must be that

$$U^E \leq P^*(\ell_0)V + (1 - P^*(\ell_0)) \max\{0, \hat{\alpha}\} - \ell_0 d(P^*(\ell_0)). \quad (56)$$

Invoking the envelope theorem, the derivative of the righthand side (RHS) of (56) with respect to $\hat{\alpha}$ is

$$\frac{\partial \text{RHS}}{\partial \hat{\alpha}} = \begin{cases} P^*(\ell_0) \frac{\partial V}{\partial \hat{\alpha}} > 0, & \text{if } \hat{\alpha} < 0 \\ 1 - P^*(\ell_0) + P^*(\ell_0)\Phi > 0, & \text{if } \hat{\alpha} > 0 \end{cases},$$

where the second line utilizes (52). So the RHS of (56) is increasing in $\hat{\alpha}$; moreover, it tends to positive infinity as $\hat{\alpha}$ does. Differentiating the RHS of (56) with respect to τ , again invoking the envelope theorem, yields

$$P^*(\ell_0) \left(-1 + \frac{1}{2} \frac{\eta}{\tau + \eta} \right) \frac{\sqrt{H}}{\tau^2} \phi < 0. \quad (57)$$

Suppose $\hat{\alpha} = 0$. Because $\tau > \tau_0$ (first-period performance means more is known about an incumbent than a replacement), the inequality in (57) implies

$$P^*(\ell_0)V - \ell_0 d(P^*(\ell_0)) < P_0 V_0 - \ell_0 d(P_0) \equiv U_0.$$

Hence, $U_0 > U^E$ if $\hat{\alpha} = 0$. Because the RHS of (56) is continuous and increasing without bound, but is less than U_0 for $\hat{\alpha} = 0$, it follows that A exists and is unique.

Turning to the “moreover” part: we’ve just established that

$$P^*(\ell_0)(V(A) - A) + A - \ell_0 d(P^*(\ell_0)) - U_0 = 0.$$

This holds for all values of ℓ_0 , so it is an identity. Differentiating that last expression with respect to ℓ_0 , using the envelope theorem, yields

$$(P^*\Phi + 1 - \Phi) \frac{\partial A}{\partial \ell_0} - d(P^*(\ell_0)) + d(P_0) = 0.$$

Proposition 8 implies $P^*(\ell_0) < P_0$; so, the last expression entails $\partial A / \partial \ell_0 < 0$. ■

Proof of Proposition 14: For $t \in \{E, I\}$ define

$$W_t = P_t^* V_t + (1 - P_t^*) \mu_t - \ell d(P_t^*).$$

Observe we can write W_I as

$$W_I(\Delta) = P_I^* \left(\overbrace{(\mu_E + \Delta)}^{\mu_I} \Phi(- (Y_I - (\mu_E + \Delta)) \sqrt{H_I}) \right. \\ \left. + \frac{\sqrt{H_I}}{\tau_I} \phi(- (Y_I - (\mu_E + \Delta)) \sqrt{H_I}) \right) \\ + (1 - P_I^*)(\mu_E + \Delta) - \ell d(P_I^*).$$

Note that the two terms in boxes sum to V_I (the value of the firm if the internal candidate is hired). Next, keeping in mind (i) $\partial V/\partial Y = 0$; ⁷⁰ (ii) $\partial V/\partial \mu = \Phi$; ⁷¹ and (iii) the envelope theorem, differentiating $W_I(\Delta)$ with respect to Δ yields:

$$W_I'(\Delta) = P_I^* \Phi + (1 - P_I^*) > 0. \quad (58)$$

If $W_I(0) < W_E$ and there exists Δ such that $W_I(\Delta) > W_E$, then Δ_m must exist by continuity. Using the envelope theorem, we have $\partial W/\partial \tau = P^* \partial V/\partial \tau < 0$, where the inequality follows from Proposition 13. So, since, $\tau_I > \tau_E$, $W_I(0) < W_E$. Because $V - \mu > 0$, $W_I(\Delta) \geq \mu_E + \Delta - \ell d(0)$; hence, $\lim_{\Delta \uparrow \infty} W_I(\Delta) = \infty$.

Turning to the “moreover” part: $0 \equiv W_I(\Delta_m) - W_E$. Differentiating that identity with respect to ℓ , using envelope theorem, yields

$$0 = W_I'(\Delta_m) \frac{\partial \Delta_m}{\partial \ell} - d(P_I^*) + d(P_E^*).$$

Because $\mu_I > \mu_E$ and $\tau_I > \tau_E$, Proposition 13 entails $P_I^* < P_E^*$; so, because $d(\cdot)$ is increasing, $0 > W_I'(\Delta_m) \partial \Delta_m / \partial \ell$. Given (58), this implies $\partial \Delta_m / \partial \ell < 0$. ■

⁷⁰Proof: recall $Y = -\tau\mu/\eta$ and

$$V = \int_Y^\infty \frac{\tau\mu + \eta y}{\tau + \eta} \sqrt{\frac{H}{2\pi}} \exp\left(-\frac{H}{2}(y - \mu)^2\right) dy.$$

So

$$\frac{\partial V}{\partial Y} = -\frac{\overbrace{\tau\mu + \eta Y}^{=0}}{\tau + \eta} \sqrt{\frac{H}{2\pi}} \exp\left(-\frac{H}{2}(Y - \mu)^2\right) dy = 0.$$

⁷¹Proof: recall V can also be written

$$V = \Phi(- (Y - \mu)\sqrt{H})\mu + \frac{\sqrt{H}}{\tau} \phi((Y - \mu)\sqrt{H}).$$

Differentiating with respect μ (keeping in mind the last footnote; *i.e.*, $\partial V/\partial Y = 0$ and Lemma A.1):

$$\frac{\partial V}{\partial \mu} = \Phi(- (Y - \mu)\sqrt{H}) + \phi(- (Y - \mu)\sqrt{H})\mu + \frac{\sqrt{H}}{\tau} \phi((Y - \mu)\sqrt{H})(Y - \mu)\sqrt{H}$$

Simple algebra (recall $Y = -\tau\mu/\eta$) shows the last two terms sum to zero.

Proof of Proposition 16: It is immediate from (44) that the greater is $bP^*\phi(- (Y - \mu)\sqrt{H})\sqrt{H}$, the greater e^* must be. Observe

$$bP^*\phi(- (Y - \mu)\sqrt{H})\sqrt{H} = bP^*\phi\left(\frac{\tau}{\sqrt{H}}\mu\right)\sqrt{H}. \quad (59)$$

If μ decreases, the ϕ term increases given $\mu > 0$ (the peak of the standard normal density occurs at zero and the density increases as one moves toward zero from right of zero). Additionally, from Proposition 13, if μ decreases, P^* increases. Putting the two effects together, a *decrease* in μ increases the value of (59). This means e^* will be greater, which establishes result (i).

Consider the righthand side of (59): the only part of it that depends on ℓ is P^* . From Proposition 13, P^* is decreasing in ℓ . This establishes result (ii). ■

Proof of Lemma 5: Consider $\hat{\alpha} > \hat{\alpha}'$. Fix $\zeta \in (0, 1)$ and define

$$\hat{\alpha}_\zeta = \zeta\hat{\alpha} + (1 - \zeta)\hat{\alpha}'.$$

We need to show

$$\Pi(\hat{\alpha}_\zeta) \leq \zeta\Pi(\hat{\alpha}) + (1 - \zeta)\Pi(\hat{\alpha}'). \quad (60)$$

By definition of a maximum

$$\Pi(z) \geq z\gamma(a^*(\hat{\alpha}_\zeta)) - c(a^*(\hat{\alpha}_\zeta)) = \Pi(\hat{\alpha}_\zeta) + \gamma(a^*(\hat{\alpha}_\zeta))(z - \hat{\alpha}_\zeta). \quad (61)$$

Consider (61) with $z = \hat{\alpha}$ and multiply through by ζ . Do the same with $z = \hat{\alpha}'$, but multiply through by $1 - \zeta$. Adding those two expressions yields

$$\zeta\Pi(\hat{\alpha}) + (1 - \zeta)\Pi(\hat{\alpha}') \geq \Pi(\hat{\alpha}_\zeta) + \gamma(a^*(\hat{\alpha}_\zeta))(\zeta\hat{\alpha} + (1 - \zeta)\hat{\alpha}' - \hat{\alpha}_\zeta) = \Pi(\hat{\alpha}_\zeta);$$

that is, (60), as was to be shown. ■

Proof of Proposition 20: The conclusion about the owners was shown in the text. The claim about the CEO follows if $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ implies $F(-c(1)|\mathcal{D}) > F(-c(1)|\mathcal{D}')$. The assumption that $F(\cdot|\mathcal{D}') \stackrel{\geq}{\text{disp}} F(\cdot|\mathcal{D})$ implies

$$F^{-1}\left(\frac{1}{2}|\mathcal{D}'\right) - F^{-1}(\xi|\mathcal{D}') < F^{-1}\left(\frac{1}{2}|\mathcal{D}\right) - F^{-1}(\xi|\mathcal{D}) \quad (62)$$

for all $\xi < 1/2$. Because the mean and median coincide, $F^{-1}(1/2|\mathcal{D}') = F^{-1}(1/2|\mathcal{D})$. Hence, (62) implies, for all $\xi < 1/2$, that

$$F^{-1}(\xi|\mathcal{D}) < F^{-1}(\xi|\mathcal{D}') \quad (63)$$

Because $-c(1) < \mathbb{E}\{\hat{\alpha}\}$, $F(-c(1)|\mathcal{D}') < 1/2$. So substituting $F(-c(1)|\mathcal{D}')$ for ξ in (63) yields

$$F^{-1}(F(-c(1)|\mathcal{D}')|\mathcal{D}) < -c(1).$$

Because $F(\cdot|\mathcal{D})$ is increasing function, this last inequality implies $F(-c(1)|\mathcal{D}') < F(-c(1)|\mathcal{D})$, as was to be shown. ■

Lemma A.3. *Let $X \sim N(\mu, \sigma_X^2)$ and $Y \sim N(\mu, \sigma_Y^2)$, where $\sigma_X^2 < \sigma_Y^2$. Then the distribution of X dominates the distribution of Y in the dispersive order.*

Proof: Let F_X and F_Y denote the distribution functions of X and Y , respectively. To show $F_X \stackrel{\geq}{\text{disp}} F_Y$, we need to show

$$F_X^{-1}(\xi'') - F_X^{-1}(\xi') < F_Y^{-1}(\xi'') - F_Y^{-1}(\xi') \quad (64)$$

if $0 < \xi' < \xi'' < 1$. Expression (64) is equivalent to

$$F_Y^{-1}(\xi') - F_X^{-1}(\xi') < F_Y^{-1}(\xi'') - F_X^{-1}(\xi'');$$

hence $F_X \stackrel{\geq}{\text{disp}} F_Y$ if and only if $F_Y^{-1}(\xi) - F_X^{-1}(\xi)$ is increasing in ξ . Replacing ξ with $F_X(z)$ reveals that $F_X \stackrel{\geq}{\text{disp}} F_Y$ if and only if $F_Y^{-1}(F_X(z)) - z$ is increasing in z . Recall

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma_X}\right) \quad \text{and} \quad F_Y(y) = \Phi\left(\frac{y - \mu}{\sigma_Y}\right).$$

Observe, therefore, $F_Y^{-1}(\xi) = \mu + \sigma_Y \Phi^{-1}(\xi)$. Consequently,

$$F_Y^{-1}(F_X(z)) - z = \mu + \sigma_Y \Phi^{-1}\left(\Phi\left(\frac{z - \mu}{\sigma_X}\right)\right) - z = \frac{\sigma_Y - \sigma_X}{\sigma_X}(z - \mu),$$

which is increasing in z because $\sigma_Y > \sigma_X$. ■

REFERENCES

- Adams, Renée B., Benjamin E. Hermalin, and Michael S. Weisbach,** “The Role of Boards of Directors in Corporate Governance: A Conceptual Framework & Survey,” *Journal of Economic Literature*, March 2010, 48 (1), 59–108.
- Allgood, Sam and Kathleen A. Farrell,** “The Match between the CEO and Firm,” *Journal of Business*, 2003, 76, 317–341.
- Alter, Adam L. and Daniel M. Oppenheimer,** “From a Fixation on Sports to an Exploration of Mechanism: The Past, Present, and Future of hot Hand Research,” *Thinking and Reasoning*, 2006, 12, 431–444.
- Baker, Erin,** “Increasing Risk and Increasing Informativeness: Equivalence Theorems,” *Operations Research*, January–February 2006, 54 (1), 26–36.
- Barro, Jason R. and Robert J. Barro,** “Pay, Performance, and Turnover of Bank CEOs,” *Journal of Labor Economics*, October 1990, 8 (4), 448–481.
- Bebchuk, Lucian Arye and Jesse M. Fried,** *Pay without Performance*, Cambridge, MA: Harvard University Press, 2004.

-
- Becht, Marco, Patrick Bolton, and Ailsa Röell**, “Corporate Governance and Control,” in George Constantinides, Milton Harris, and René Stulz, eds., *The Handbook of the Economics of Finance*, Amsterdam: North-Holland, 2003.
- Berk, Jonathan and Richard Green**, “Mutual Fund Flows and Performance in Rational Markets,” *Journal of Political Economy*, 2004, *112*, 1269–1295.
- Berle, Adolph A. and Gardiner C. Means**, *The Modern Corporation and Private Property*, New York: MacMillan, 1932.
- Bertrand, Marianne and Sendhil Mullainathan**, “Are CEOs Rewarded for Luck? The Ones Without Principals Are,” *Quarterly Journal of Economics*, 2001, *116*, 901–932.
- Blackwell, David W., James A. Brickley, and Michael S. Weisbach**, “Accounting Information and Internal Performance Evaluation: Evidence from Texas Banks,” *Journal of Accounting and Economics*, May 1994, *17*, 331–358.
- Blanes i Vidal, Jordi and Marc Möller**, “When Should Leaders Share Information with Their Subordinates?,” *Journal of Economics & Management Strategy*, Summer 2007, *16* (2), 251–283.
- Bocskocsky, Andrew, John Ezekowitz, and Carolyn Stein**, “The Hot Hand: A New Approach to an Old ‘Fallacy’,” in “8th Annual MIT Sloan Sports Analytics Conference” 2014.
- Boschen, John F. and Kimberly J. Smith**, “You Can Pay Me Later: The Dynamic Response of Executive Compensation to Firm Performance,” *Journal of Business*, 1995, *68*, 577–608.
- Brickley, James A. and R. Lawrence Van Horn**, “Managerial Incentives in Nonprofit Organizations: Evidence from Hospitals,” *Journal of Law and Economics*, 2002, *45* (1, Part 1), 227–249.
- Burns, Natasha and Simi Kedia**, “The Impact of Performance-based Compensation on Misreporting,” *Journal of Financial Economics*, 2006, *79*, 35–67.
- Cebon, Peter and Benjamin E. Hermalin**, “When Less Is More: The Benefits of Limits on Executive Pay,” *Review of Financial Studies*, 2015, *28* (6), 1667–1700.
- Cho, In-Koo and David Kreps**, “Signaling Games and Stable Equilibria,” *Quarterly Journal of Economics*, 1987, *102*, 179–222.
- Chung, Ji-Woong, Berk A. Sensoy, Léa H. Stern, and Michael S. Weisbach**, “Pay for Performance from Future Fund Flows: The Case of Private Equity,” *Review of Financial Studies*, 2012, *25*, 3259–3304.

-
- Clayton, Matthew C., Jay C. Hartzell, and Joshua Rosenberg**, “The Impact of CEO Turnover on Equity Volatility,” *Journal of Business*, 2005, 78, 1779–1808.
- Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist**, “Monitoring Managers: Does It Matter?,” *Journal of Finance*, 2013, 68 (2), 431–481.
- Coughlan, Anne T. and Ronald M. Schmidt**, “Executive Compensation, Managerial Turnover, and Firm Performance: An Empirical Investigation,” *Journal of Accounting and Economics*, 1985, 7 (1/2/3), 43–66.
- Dechow, Patricia M. and Richard G. Sloan**, “Executive Incentives and the Horizon Problem: An Empirical Investigation,” *Journal of Accounting and Economics*, March 1991, 14 (1), 51–89.
- DeGroot, Morris H.**, *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.
- Denis, David J. and Diane K. Denis**, “Performance Changes Following Top Management Dismissals,” *Journal of Finance*, 1995, 50, 1029–1057.
- , – , and **Mark D. Walker**, “CEO Assessment and the Structure of Newly Formed Boards,” *Review of Financial Studies*, in press.
- Edmans, Alex and Xavier Gabaix**, “Executive Compensation: A Modern Primer,” *Journal of Economic Literature*, in press.
- , **Vivian W. Fang, and Katharina A. Lewellen**, “Equity Vesting and Investment,” *Review of Financial Studies*, in press.
- Efendi, Jap, Anup Srivastava, and Edward P. Swanson**, “Why Do Corporate Managers Misstate Financial Statements? The Role of Option Compensation and Other Factors,” *Journal of Financial Economics*, 2007, 85, 667–708.
- Eisfeldt, Andrea and Camelia Kuhnen**, “CEO Turnover in a Competitive Assignment Framework,” *Journal of Financial Economics*, 2013, 109 (2), 351–372.
- Eldenburg, Leslie, Benjamin E. Hermalin, Michael S. Weisbach, and Marta Wosińska**, “Governance, Performance Objectives and Organizational Form: Evidence from Hospitals,” *Journal of Corporate Finance*, September 2004, 10 (4), 527–548.
- Engel, Ellen, Rachel Hayes, and Xue Wang**, “CEO Turnover and Properties of Accounting Information,” *Journal of Accounting and Economics*, 2003, 36 (1–3), 197–226.

- Erickson, Merle, Michelle Hanlon, and Edward L. Maydew**, “Is There a Link Between Equity Linked Incentives and Accounting Fraud?,” *Journal of Accounting Research*, 2006, *44*, 113–143.
- Fama, Eugene F.**, “Agency Problems and the Theory of the Firm,” *Journal of Political Economy*, 1980, *88* (2), 288–307.
- and **Kenneth R. French**, “Dividend Yields and Expected Stock Returns,” *Journal of Financial Economics*, 1988, *22*, 3–26.
- Fudenberg, Drew and Jean Tirole**, “A ‘Signal-Jamming’ Theory of Predation,” *RAND Journal of Economics*, Autumn 1986, *17* (3), 366–376.
- Gibbons, Robert S. and Kevin J. Murphy**, “Relative Performance Evaluation for Chief Executive Officers,” *Industrial and Labor Relations Review*, 1990, *43*, 30S–51S.
- and —, “Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence,” *Journal of Political Economy*, 1992, *100*, 468–505.
- Gilovich, Thomas**, *How We Know What Isn’t So: The Fallibility of Human Reason in Everyday Life*, New York: The Free Press, 1991.
- , **Robert Vallone, and Amos Tversky**, “The Hot Hand in Basketball: On the Misperception of Random Sequences,” *Cognitive Psychology*, 1985, *17*, 295–314.
- Gopalan, Radhakrishnan, Todd Milbourn, Fenghua Song, and Anjan V. Thakor**, “Duration of Executive Compensation,” *Journal of Finance*, 2014, *69* (6), 2777–2817.
- Graham, John, Campbell Harvey, and Shiva Rajgopal**, “The Economic Implications of Corporate Financial Reporting,” *Journal of Accounting and Economics*, 2005, *40*, 3–73.
- Grossman, Sanford J. and Oliver D. Hart**, “An Analysis of the Principal-Agent Problem,” *Econometrica*, 1983, *51*, 7–46.
- Hermalin, Benjamin E.**, “Managerial Preferences Concerning Risky Projects,” *Journal of Law, Economics, & Organization*, 1993, *9*, 127–135.
- , “Trends in Corporate Governance,” *Journal of Finance*, 2005, *60* (5), 2351–2384.
- , “Corporate Governance,” in Robert S. Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton, NJ: Princeton University Press, 2013.
- and **Michael L. Katz**, “Moral Hazard and Verifiability: The Effects of Renegotiation in Agency,” *Econometrica*, November 1991, *59* (6), 1735–1753.

-
- **and** –, “Corporate Diversification and Agency,” in Peter J. Hammond and Gareth D. Myles, eds., *Incentives, Organization, and Public Economics: Papers in Honour of Sir James Mirrlees*, Oxford: Oxford University Press, 2000.
 - **and Michael S. Weisbach**, “The Determinants of Board Composition,” *RAND Journal of Economics*, Winter 1988, 19 (4), 589–606.
 - **and** –, “Endogenously Chosen Boards of Directors and Their Monitoring of the CEO,” *American Economic Review*, March 1998, 88 (1), 96–118.
 - **and** –, “Information Disclosure and Corporate Governance,” *Journal of Finance*, February 2012, 67 (1), 195–233.
 - **, Avery W. Katz, and Richard Craswell**, “Contract Law,” in A. Mitchell Polinsky and Steven Shavell, eds., *Handbook of Law and Economics*, Vol. 1, Amsterdam: North-Holland, 2007.
 - Holmstrom, Bengt**, “Moral Hazard and Observability,” *Bell Journal of Economics*, Spring 1979, 10 (1), 74–91.
 - **,** “Managerial Incentive Problems—A Dynamic Perspective,” *Review of Economic Studies*, January 1999, 66 (226), 169–182.
 - **and Joan Ricart i Costa**, “Managerial Incentives and Capital Management,” *Quarterly Journal of Economics*, November 1986, 101 (4), 835–860.
 - Huson, Mark R., Robert Parrino, and Laura T. Starks**, “Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective,” *Journal of Finance*, December 2001, 55 (6), 2265–2297.
 - Jensen, Michael C.**, “Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers,” *American Economic Review*, May 1986, 76 (2), 323–329. Papers and Proceedings of the Ninety-Eighth Annual Meeting of the American Economic Association.
 - **and Kevin J. Murphy**, “Performance Pay and Top-Management Incentives,” *Journal of Political Economy*, 1990, 98, 225–264.
 - **and William H. Meckling**, “Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure,” *Journal of Financial Economics*, 1976, 3, 305–360.
 - Jenter, Dirk and Fadi Kanaan**, “CEO Turnover and Relative Performance Evaluation,” *Journal of Finance*, 2015, 70 (5), 2155–2184.
 - **and Katharina Lewellen**, “Performance-induced CEO Turnover,” 2014. Stanford Working Paper.

-
- Johnson, Shane A., Harley E. Ryan, and Yisong S. Tian**, “Managerial Incentives and Corporate Fraud: The Source of Incentives Matter,” *Review of Finance*, 2009, 13, 115–145.
- Johnson, Simon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer**, “Tunneling,” *American Economic Review*, May 2000, 90 (2), 22–27.
- Jones, Edward E. and Victor A. Harris**, “The Attribution of Attitudes,” *Journal of Experimental Social Psychology*, 1967, 3, 1–24.
- Kahneman, Daniel**, *Thinking, Fast and Slow*, New York: Farrar, Strauss, and Giroux, 2011.
- Kaplan, Steven N. and Bernadette A. Minton**, “How has CEO Turnover Changed?,” *International Review of Finance*, March 2012, 12 (1), 57–87.
- Khurana, Rakesh**, “The Curse of the Superstar CEO,” *Harvard Business Review*, September 2002, 80 (9), 60–65.
- , *Searching for a Corporate Savior: The Irrational Quest for Charismatic CEOs*, Princeton, NJ: Princeton University Press, 2002.
- Kim, Son Ku**, “Efficiency of an Information System in an Agency Model,” *Econometrica*, January 1995, 63 (1), 89–102.
- Leroy, Stephen and Richard D. Porter**, “The Present Value Relation: Tests Based in Implied Variance Bounds,” *Econometrica*, 1981, 49, 555–574.
- Levin, Jonathan**, “Relational Incentive Contracts,” *American Economic Review*, June 2003, 93 (3), 835–857.
- Lim, Jongha, Berk A. Sensoy, and Michael S. Weisbach**, “Indirect Incentives of Hedge Fund Managers,” *Journal of Finance*, 2016, 71, 871–918.
- Maskin, Eric and Jean Tirole**, “Unforeseen Contingencies and Incomplete Contracts,” *Review of Economic Studies*, January 1999, 66 (1), 83–114.
- Murphy, Kevin J.**, “Corporate Performance and Managerial Remuneration,” *Journal of Accounting and Economics*, 1985, 7 (1), 11–42.
- , “Incentives, Learning and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts,” *RAND Journal of Economics*, 1986, 17 (1), 59–76.
- Narayanan, M.P.**, “Managerial Incentives for Short-Term Results,” *The Journal of Finance*, 1985, 40 (5), 1469–1484.
- Nash, John F.**, “The Bargaining Problem,” *Econometrica*, April 1950, 18 (2), 155–162.

- Pan, Yihui, Tracy Yue Wang, and Michael S. Weisbach**, “Learning about CEO Ability and Stock Return Volatility,” *Review of Financial Studies*, 2015, 28, 1623–1666.
- , – , and – , “How Management Risk Affects Corporate Debt,” 2016. Working Paper, Ohio State University.
- Parrino, Robert**, “CEO Turnover and Outside Succession: A Cross-Sectional Analysis,” *Journal of Financial Economics*, 1997, 46, 165–197.
- , **Richard W. Sias, and Laura T. Starks**, “Voting With Their Feet: Institutional Ownership Changes Around Forced CEO Turnover,” *Journal of Financial Economics*, 2003, 68 (1), 3–46.
- Pastor, Lubos and Pietro Veronesi**, “Stock Valuation and Learning about Profitability,” *Journal of Finance*, 2003, 58, 1749–1789.
- and – , “Was there a NASDAQ Bubble in the late 1990s?,” *Journal of Financial Economics*, 2006, 81, 61–100.
- and – , “Learning in Financial Markets,” *Annual Review of Financial Economics*, 2009, 1, 361–81.
- Plath, Christian**, “Analyzing Credit and Governance Implications of Management Succession Planning,” Technical Report, Moody’s Investors Service 2008. Available at <http://ssrn.com/abstract=1285082>.
- Plous, Scott**, *The Psychology of Judgment and Decision Making*, New York: McGraw-Hill, 1993.
- Rogerson, William P.**, “Repeated Moral Hazard,” *Econometrica*, January 1985, 53 (1), 69–76.
- Ross, Lee**, “The Intuitive Psychologist and his Shortcomings: Distortions in the Attribution Process,” in Leonard Berkowitz, ed., *Advances in Experimental Social Psychology*, Vol. 10, New York: Academic Press, 1977, pp. 173–220.
- Rotemberg, Julio J. and Garth Saloner**, “Leadership Style and Incentives,” *Management Science*, November 1993, 39 (11), 1299–1318.
- and – , “Benefits of a Narrow Business Strategy,” *American Economic Review*, December 1994, 84 (5), 1330–1349.
- and – , “Visionaries, Managers, and Strategic Direction,” *RAND Journal of Economics*, 2000, 31, 693–716.
- Sappington, David E.M.**, “Limited Liability Contracts between Principal and Agent,” *Journal of Economic Theory*, February 1983, 29 (1), 1–21.
- Scharfstein, David S. and Jeremy C. Stein**, “Herd Behavior and Investment,” *The American Economic Review*, June 1990, 80 (3), 465–479.

-
- Shaked, Moshe and J. George Shanthikumar**, *Stochastic Orders and their Applications*, San Diego: Academic Press, 1994.
- Shiller, Robert**, “Do Stock Prices Move too Much to be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 1981, *71*, 421–436.
- Shleifer, Andrei and Robert W. Vishny**, “A Survey of Corporate Governance,” *Journal of Finance*, June 1997, *52* (2), 737–783.
- Smith, Adam**, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Indianapolis: Liberty Press, 1776.
- Stein, Jeremy C.**, “Takeover Threats and Managerial Myopia,” *Journal of Political Economy*, February 1988, *96* (1), 61–80.
- , “Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior,” *Quarterly Journal of Economics*, November 1989, *104* (4), 655–669.
- , “Agency, Information and Corporate Investment,” in George M. Constantinides, Milton Harris, and René Stulz, eds., *Handbook of the Economics of Finance*, Amsterdam: North-Holland, 2003.
- Stern, Léa**, “A Learning Based Approach to Evaluating Boards of Directors,” 2016. Working Paper, University of Washington.
- Taylor, Lucian A.**, “Why Are CEOs Rarely Fired? Evidence from Structural Estimation,” *Journal of Finance*, 2010, *65* (6), 2051–2087.
- , “CEO Wage Dynamics: Estimates from a Learning Model,” *Journal of Financial Economics*, 2013, *108*, 79–98.
- Timmermann, Allan G.**, “How Learning in Financial Markets Generates Excess Volatility and Predictability,” *Quarterly Journal of Economics*, 1993, *108*, 1135–1145.
- Tversky, Amos and Daniel Kahneman**, “Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment,” *Psychological Review*, October 1983, *90* (4), 293–315.
- Van den Steen, Eric**, “Organizational Beliefs and Managerial Vision,” *Journal of Law, Economics, & Organization*, Spring 2005, *21* (1), 256–283.
- Warner, Jerold B., Ross L. Watts, and Karen H. Wruck**, “Stock Prices and Top-Management Changes,” *Journal of Financial Economics*, Jan./March 1988, *20* (1/2), 461–492.
- Wasserman, Noam, Bharat Anand, and Nitin Nohria**, “When Does Leadership Matter? A Contingent Opportunities View of CEO Leadership,” in Nitin Nohria and Rakesh Khurana, eds., *Handbook of Leadership Theory and Practice*, Boston: Harvard Business School Publishing, 2010.

-
- Weber, Roberto A., Colin F. Camerer, Yuval Rottenstreich, and Marc Knez**, “The Illusion of Leadership: Misattribution of Cause in Coordination Games,” *Organization Science*, September-October 2001, *12* (5), 582–598.
- Weisbach, Michael S.**, “Outside Directors and CEO Turnover,” *Journal of Financial Economics*, Jan./March 1988, *20* (1/2), 431–460.
- Williamson, Oliver E.**, “Managerial Discretion and Business Behavior,” *American Economic Review*, December 1963, *53* (5), 1032–1057.
- Wolinsky, Asher**, “Competition in a Market for Informed Experts’ Services,” *RAND Journal of Economics*, Autumn 1993, *24* (3), 380–398.