How do Independent Boards Affect Shareholder Value? Evidence from a Structural Estimation*

Francesco Celentano

University of Lausanne

ABSTRACT

This paper evaluates the effects of board independence on shareholder value by estimating a dynamic model. The model features learning and monitoring of the CEO's ability, costly turnover, and bargaining over the CEO's compensation. Board independence is chosen endogenously taking into account the CEO's disutility from independence and the CEO's perceived ability, as well as the costs of monitoring and firing the CEO. I find evidence that board monitoring contributes to shareholder value, but is inhibited due to the CEO's pay capturing most of the combined bargaining surplus from the CEO and the board. If the board and the CEO had equal bargaining power, board independence would increase by 56%, board's learning of the CEO's ability would occur faster since monitoring would increase by three times, resulting in a reduction of the CEO's average career earnings of 44% and an increase in the firm's profitability by 3%. I also find that more important than regulations requiring minimum levels of board independence is providing the right incentives for the board to monitor.

*I am extremely indebted to my advisors Dean Corbae, Oliver Levine, Antonio Mello and Erwan Quintin for invaluable guidance and support during my PhD. I also thank Brian Chang, Jason Choi, Theodosios Dimopoulos, Andrew Ellul, Boris Nikolov, Dmitry Orlov, Charlotte Ostergaard, Marco Pagano, Mark Rempel and Randall Wright and many seminar participants at the University of Wisconsin-Madison, University of Naples Federico II, Virginia Tech, University of Lausanne, BI Norwegian Business School, Tilburg University, ESCP Business School, HEC Montreal, University of Vienna, and European Econometric Society Winter Meetings 2021 in Barcelona for helpful comments. I am grateful to Florian Peters and Alex Wagner for providing data on CEO forced turnover from 1993 to 2016. All errors are my own.

1 Introduction

Boards of directors have arguably the most crucial decision-making role within the governance of companies. They are supposed to perform critical functions such as monitoring and advising the Chief Executive Officer (CEO) over time as well as making decisions about replacing the CEO if the firm is performing poorly or is involved in a scandal. Since the decisions that boards take can have a considerable impact on shareholder value, it is important that boards are comprised of individuals that maximize its effectiveness. For this reason, in the last decades, investors, regulators, and academic researchers have shown an increasing interest in boards and in particular in their composition.

For years, policy-makers and academics focused on a specific characteristic of board composition, board independence. The reason behind this attention is that, according to the conventional wisdom, a board is more effective, the more it is made up of independent directors, that is directors with no significant financial or personal ties to the firm and to the management. Because of lack of conflicts of interest, independent directors are expected to vigilant and prone to replace the CEO if the company is performing poorly. While this concept seems appealing and intuitive, the empirical evidence about the relation between board independence and firm value is ambiguous and weak (see the surveys by Baghat and Black (1999) and Adams (2017)).

In this paper, I quantify how much board independence increases shareholder value by developing a dynamic model and structurally estimating it. The model features a board of directors that learns about the CEO's unobservable ability, meaning the capacity to generate profits higher than the industry average. Learning occurs both passively, by observing the realization of the firm profits, and actively, by exerting effort to collect additional soft information about the CEO. Following the literature about boards, I define the latter type of learning as monitoring throughout the rest of the paper. The board decides to fire the CEO if her perceived ability falls below a certain threshold. However, dismissing the CEO entails a cost to the shareholders, which corresponds to the cost of replacing her, as well as a directors' cost, which reflects not only any uncompensated effort and stress from the CEO succession process, but also the fact that forcing out a manager may hurt their careers. The cost to the directors creates a misalignment of incentives between shareholders and the board about the decision of firing the CEO: since it affects only the directors, this cost might induce the board to retain some CEOs whom the shareholders would rather see fired. The model also features bargaining between the board and the CEO over future compensation. Moreover, the board chooses its future composition taking into account the CEO's perceived ability. Specifically, each period the directors decide about the future board independence, that is the fraction of independent directors who will sit on the board. This decision is affected by a second type of agency conflict: the board chooses its future independence to maximize the total surplus of the match with the CEO rather than the shareholder value. This is because the board also takes into account the CEO's utility, as the key manager of the firm. Hence, when the board decides about its independence, it considers not only the benefits of having a more independent board, but also the fact that the CEO has a preference for less board independence.

The key feature of the model is that a more independent board has not only higher incentive to monitor the CEO but also a lower cost firing her. This implies that both the speed of learning and the firing frequency are positively affected by board independence. Consequently, both the bargaining outcome and the CEO's tenure at the firm are affected by board independence through these learning and firing channels.

Thus, the model suggests two potential reasons for why an independent board might contribute to the shareholder value: by increasing the speed of learning, and by decreasing the level of entrenchment of the CEO. Estimates of the model's parameters dictate the relative strength of these two channels.

I estimate the model to assess the quantitative effects of board independence on the shareholder value and the CEO's ability to extract rents. The parameter estimates show that to rationalize the actual low percentage of fired CEOs, the board has to behave as if firing the CEO would be very costly. Further, the model estimate of the CEO's bargaining power must be very large to fit the observed data. This result can be explained by the fact that the share of the bargaining surplus that the CEO is able to obtain strongly affects the board's incentives of monitoring the CEO and hence to stay independent.

I then use the estimates of the structural model to study how the model outcomes are affected by the different agency frictions and the board monitoring. I find that removing the dependent directors' additional cost of firing the CEO increases the shareholder value by reducing the CEO's average earnings. However, the board has lower incentives to monitor since board independence doesn't directly affect the degree of CEO's entrenchment, causing the firm's profitability to fall. Eliminating the board monitoring decreases both the shareholder value and the CEO's earnings since the board is not able to distinguish a good CEO from a bad one, and it bases all its decision upon the observed realization of the firm profits. If the board doesn't have the capacity of monitoring, eliminating the dependent director's additional cost improves the situation by decreasing the CEO's rent extraction. Furthermore, I show that increasing the board's bargaining power, so that the board and the CEO obtain the same share of the surplus, increases the shareholder value from a higher firm's profitability. Interestingly, this result is driven by the fact that the board increases its monitoring intensity since it is able to obtain a higher return from investing in monitoring rather than by simply the lower CEO's bargaining power. Overall, the results highlight how board monitoring plays an important role in value creation. Finally, I show that regulations introducing a minimum threshold numbers of independent directors have a modest impact. In sum, creating the condition for rigorous monitoring of the CEO's performance appears to be more important than setting rules over the ratio of independent to dependent directors.

This paper contributes to the large empirical literature[1] about boards of directors and board independence by structurally estimating a model with endogenous board composition via Simulated Method of Moments (SMM). To my knowledge, this study represents the first attempt at using a dynamic model to evaluate the impact of board independence. As emphasized by Adams et al. (2010) since there is no reason to assume that the board structure is

[1] See, for example, Weisbach (1988), Yermack (1996), Coles et al. (2008), Adams and Ferreira (2009), Ravina and Sapienza (2009), Fahlenbrach et al. (2010), Ahern and Dittmar (2012), Guthrie et al. (2012) Coles et al. (2014), Amihud and Stoyanov (2017), Stern (2019) and Graham et al. (2020).

exogenous, this endogeneity creates estimation problems and challenges for empirical studies. Moreover, the CEO's pay is the outcome of negotiations between the board and the CEO and, as is the decision about board independence and about firing the CEO, depends on the CEO's perceived ability which is an object without obvious measures in the data. These challenges lend themselves to a structural estimation approach since this infers unobservable quantities from endogenous patterns without requiring an instrument. This approach allows to assess not only the qualitative effects but also magnitude of the model elements, and allows for the possibility to explore counterfactuals. In addition, another contribution to the empirical literature is the classification of independent directors that I adopt. I consider as dependent directors not only those that are not classified as independent according to the Institutional Shareholder Services (ISS) database, but also independent directors with more than nine years of tenure on the board of the same firm. As I explain in detail in the identification section, this threshold corresponds to the one the ISS considers a director's independence to be potentially compromised, and is close that adopted by several countries outside the US that have implemented tenure related requirements in the last decades.

This paper fits in the theoretical[2] literature on boards by extending the model of Hermalin and Weisbach (1998) to a fully dynamic setup with persistent earnings in a similar way to Taylor (2010). In addition to the work by Taylor (2010), my paper endogeneizes monitoring, board composition, and CEO's pay. Another significant difference between this paper and Hermalin and Weisbach (1998) is the bargaining protocol. While in Hermalin and Weisbach (1998) both the CEO's pay and board independence are chosen via Nash bargaining, in my paper independence is chosen to maximize the total surplus while the CEO's pay is set to split the surplus according to Nash bargaining. This implies that there is not necessarily a trade-off between CEO's pay and board composition: according to how good the CEO is perceived to be, higher board independence can positively affect the CEO's compensation.

Finally, my paper relates to recent works in structural corporate finance on corporate

[2]See, for example, Hirshleifer and Thakor (1994), Hermalin and Weisbach (1998), Hermalin (2005), Song and Thakor (2006), Adams and Ferreira (2007), Harris and Raviv (2006), Laux (2008), Baldenius et al. (2014), Fisman et al. (2014), Levit and Malenko (2016), Garlappi et al. (2017), Chemmanur and Fedaseyeu (2018), and Donaldson et al. (2020).

governance such as Taylor (2010, 2013), and Page. (2018). Specifically, the main differences between my work and these papers are the endogenous monitoring and its effects on the bargained compensation and the firing decision, as well as the board endogenous composition. Another difference is the moment selection for the SMM since I use statistics related not only to the firm's performance and to CEO's characteristics but also to board independence. In addition, my paper relates to Matveyev (2016) who finds that increasing directors' pay would induce less desirable firms to attract a better pool of directors.

The remainder of the paper is organized as follows. Section 2 presents and analyzes the model. Section 3 describes the data and the estimation procedure and presents the empirical results. Section 4 presents the counterfactual experiments, and Section 5 concludes. The Appendix provides technical details.

2 Model

I present a dynamic model of learning and bargaining in which board independence is endogenously chosen. First, I describe each part of the model in detail and then I discuss the model's solution.

The model features a firm that lives for an infinite number of periods, a board of directors, a CEO, and a pool of potential replacement CEOs.[3] All parties are risk neutral and share a discount factor $\beta \in (0, 1)$. Time is discrete and indexed by t, and one period corresponds to one year. Each period the board decides its level of independence in the next period, as the fraction of independent directors at $t + 1$. The board acts as an agent of the shareholders and therefore makes choices to maximize the total surplus. The total surplus is the sum of the surplus that a board brings on the table given its independence level and the surplus created by the characteristics of a particular CEO. Board independence lowers the board's cost of monitoring as well as the cost of firing the CEO. However, the board also need to considers that the CEO dislikes board independence. Hence, when the board chooses its

[3] The shareholders don't have any active role, that is, they simply ratify the slate of directors put forward by the company even though, in practice, they could vote against a director nominee.

level of independence, it trades-off the costs of monitoring and firing and the CEO's disutility from board independence. If at some point in time the board is not able to find a level of independence such that the total surplus is positive, then the board fires the CEO. After deciding the level of independence and whether the CEO is fired or not, the board bargains with the CEO (either the current or a new CEO) over the CEO compensation, and then decides how intensely to monitor. In sum, both the board's decisions and the bargaining outcome are affected by the current board independence and the board's beliefs about the CEO's ability whose actual value is unknown to both parties.

Production Technology

The firm has book assets constant over time and equal to B[4]. The firm generates gross profits (profits before CEO pay) equal to $Y_t B$, where Y_t is the firm profitability rate and equals

$$Y_t = x_t + y_t - \mathbb{1}\{turnover_t\}c_{firm}. \tag{1}$$

The first component, x_t, is an industry profitability rate and its realization is observable. The first element of the third component, $\mathbb{1}\{turnover_t\}$, is an indicator that equals one if a CEO turnover occurs either because the CEO is fired or because leaves the firm at time t, and zero otherwise. The second element, c_{firm}, is a parameter representing the cost of a CEO turnover to the shareholders, meaning how much the gross profitability decreases in the event of a turnover. This cost includes severance or retirement packages, executive search fees, disruption costs, and any other CEO turnover costs that affect the profits of the firm. The second component, y_t, is a firm specific profitability rate and follows an autoregressive process of order 1,

$$y_t = \phi\alpha + (1 - \phi)y_{t-1} + \epsilon_t. \tag{2}$$

[4] For tractability, I assume that all the cash flows generated by the firm are paid out to the shareholders if positive and refinanced with external finance if negative.

The persistence parameter, ϕ, is assumed to be in the range between zero and one in that the process is mean-reverting around α, the current CEO's ability level. Neither the CEO nor the board know the actual value of α, which is fixed throughout the CEO's tenure[5]. Given this characterization of a CEO's ability, a good CEO, that is a CEO with a high α, is the one who is able to achieve a long-run average profitability higher than industry's competitors. Thus, since the firm's specific profitability is persistent, a good (bad) CEO might have a long lasting positive (negative) effect on the firm's profitability, and, at the same time, a good (bad) CEO might have low (high) earnings during the beginning of its career.

The shock ϵ_t is an independently and normally distributed shock with mean 0 and variance σ_ϵ^2. The realization of this shock is not observable, hence, when a high firm's specific profitability is observed, neither the CEO nor the board can be sure whether this is due to a high ability CEO (i.e. one with a high α) or luck (i.e. a high realization of ϵ_t).

Each period the firm pays the CEO a fraction of its assets, w_t, which is bargained in the previous period $t-1$. The cash flows that the firm generates at time t are given by

$$D_t = (Y_t - w_t)B. \tag{3}$$

Board of Directors

Each period the board learns about the CEO's ability by observing the realization of the firm's specific profitability, y_t, and an additional latent signal, z_t. This second signal is normally distributed with mean equal to the CEO's ability level, α, and variance σ_z^2 [6]. It represents additional information that is unrelated to current profitability, and includes information held privately in the firm, such as CEO's specific actions and choices, performance of individual project, and CEO's strategic plans, as well as public information, such

[5] This assumption can be justified by noting that the uncertainty about a CEO's ability in a particular job at a specific firm is largely uncertainty about the match between her and that company which is unknown to both the board and the CEO. The adopted notion of ability is close to one of synergy between a CEO and a firm. To simplify notation, I don't use firm or CEO specific subscripts.

[6] In other words, the draws of z_t are independent and identically distributed conditional on the value of the CEO's ability, α.

as firm growth prospects, stock market prices, discretionary earnings accruals, and media coverage[7]. Unlike the earnings signal, the additional signal is not observed by the board every period with probability one (except in the first period as I explain below). Specifically, the probability of observing the additional signal is given by a Bernoulli distribution with success probability

$$\mathbb{P}(z_t \text{ is received}) = 1 - e^{-p_t} \tag{4}$$

which implies that the probability of receiving z_t is increasing and concave in the monitoring intensity chosen by the board at time $t-1$, p_t. The time t cost of monitoring for the board is given by $\theta(1-i_t)p_{t+1}B$, where the parameter $\theta > 0$ captures how much the board distastes monitoring the CEO and $i_t \in [0,1)$[8] is the percentage of independent directors chosen at time $t-1$. Therefore, this cost is increasing in the percentage of dependent directors, $1-i_t$, since dependent directors' careers are tied to the CEO's own career, and so they rarely find it in their interest to monitor the CEO. Accordingly, this cost should be thought of as a personal cost to monitor rather than a cost to get access to information. The monitoring cost is also increasing in the amount of assets of the firm, B. This relation captures the fact that it is difficult for the board to assess the added value of the CEO of a large firm since large firms are usually more complex and they are involved in multiple types of businesses.

The board chooses the monitoring intensity, p_{t+1}, to maximize its lifetime utility V_t:

$$V_t = \max_{\{p_{t+1+s}\}_{s=0}^{\infty}} E_t\left[\sum_{s=0}^{\infty} \beta^s v_{t+s}\right] \tag{5}$$

where v_t is the per-period utility of the board, which is given by

$$v_t = D_t - \theta(1-i_t)p_{t+1}B - \mathbb{1}\{fire_t\}(c_{board,0} + c_{board,1}(1-i_t)^2)B. \tag{6}$$

[7]Cornelli et al. (2013) empirically document the importance of soft information and that it plays a much larger role than performance metrics for directors' decisions.

[8]This fraction is bounded below one since in the overwhelming majority of US public firms the CEO has a seat on the board.

The first component represents the directors' preference for higher firm cash flows. It reflects both the fact that directors often have shares in the firm and that firms with good cash flows help the directors' reputation. The third component represents an additional personal cost that the all directors incur every time a CEO is dismissed[9]. This cost is a private cost to the board members that is not reflected in the firm profits. The first element, $\mathbb{1}\{fire_t\}$, is an indicator function that equals one if the CEO is fired in period t, and zero otherwise. The board personal cost is increasing in the amount of assets of the firm since, the bigger the firm, the greater the responsibility of the board when a CEO is dismissed.

The board's firing cost is made of a fixed and a variable part. The fixed cost term, $c_{board,0}$, captures how much all the directors are affected by the departure of the CEO. It represents the reputation cost that all the board members incur for dismissing the CEO plus any uncompensated effort and stress from the succession process. The variable cost term, $c_{board,1}(1-i_t)^2$, represents the additional cost that the dependent directors incur if the board decides to fire the CEO. It includes all the personal costs that the dependent directors incur by losing an ally both within and outside the firm. This term is increasing and convex in the fraction of dependent directors since the more dependent a board is, the more each dependent director's reputation suffer if the board fires the CEO. The board's cost of firing represents the misalignment of incentives between the board and the shareholders with regards to the decision of firing the CEO. Thus, the fraction of independent directors has two effects on the board: it creates incentives to monitor more intensely the CEO, and it makes more likely the dismissal of an under-performing CEO.

Before a new CEO is matched with the firm, all the replacement CEOs look ex-ante identical. The distribution of ability of any replacement CEO before being matched with a firm is normally distributed with mean μ_0 and variance σ_0^2. Alternatively, one can think that the CEOs' talent pool is $\mathcal{N}(\mu_0, \sigma_0^2)$. Before bargaining, but after a new CEO has been matched with the firm, the board receives an additional signal z_t about the new manager with probability one. This initial additional signal represents the information about the new

[9] In Taylor (2010) and in other subsequent papers, this cost has been shown to be important to explain the low CEO percentage of fired CEOs.

CEO that the board gathers from interviews, references as well as CEO's past experiences. This assumption implies that when the board and the CEO negotiate for the first time, the board has updated its beliefs. I define the updated initial mean as $\mu_{s,0}$ and the updated initial variance as $\sigma^2_{s,0}$. After the first period, the board updates its beliefs about the CEO's ability with Bayes rule using the two signals so that in each period the board has beliefs characterized by a Normal distribution with mean μ_t and variance σ^2_t. The belief distribution is assumed to be known by both parties since the CEO observes her performance, has private conversations with the directors and has access to private and public performance evaluation letters.

The board' utility conditional on the CEO staying is

$$V_{stay}(\mu, \sigma, i, \tau, w) = \max_{p'} \left\{ \frac{\phi\mu}{1-\beta(1-\phi)} - w - \theta(1-i)p' + \beta E[V(\mu', \sigma', i', \tau', w')|p'] \right\} \quad (7)$$

subject to Bayesian updating of the beliefs about the CEO's ability[10]. The first term reflects the board preference for a good CEO. After the incumbent CEO has left the firm but before a new CEO is matched with the firm, the board's utility can be written as

$$V_{leave}(\mu, \sigma, i, \tau, w) = E[V_{stay}(\mu_{s,0}, \sigma_{s,0}, i, 0, w)] - c_{firm} \quad (8)$$

and, similarly, after the incumbent CEO has been fired but before a new CEO is matched with the firm, the board's utility is

$$V_{fire}(\mu, \sigma, i, \tau, w) = E[V_{stay}(\mu_{s,0}, \sigma_{s,0}, i, 0, w)] - c_{firm} - c_{board,0} - c_{board,1}(1-i)^2 \quad (9)$$

Notice that the industry profitability rate, x_t, drops out of the expression. The amount of the firm book assets, B, does not appear since all the above utilities are expressed per unit of firm assets. Furthermore, the CEO's pay is a state variable in all the three different scenarios since it is paid to the current CEO whether she stays, she leaves or she is fired. Finally, the continuation value on the right-hand side of equation (7) equals $V_{leave}(\mu', \sigma', i', \tau', w')$ if the CEO

[10] For details about the board learning and the derivation of the value functions, see the appendix.

leaves, $V_{fire}(\mu', \sigma', i', \tau', w')$ if the CEO does not leave but gets fired, and $V_{stay}(\mu', \sigma', i', \tau', w')$ if she doesn't leave and is not fired.

CEO

Each CEO works up to $\bar{\tau}$ periods for the firm and every period she voluntarily might leave the firm, i.e. either quits or retires, with an exogenous probability $q(\tau)$[11]. Unlike the board, she does not make any active decision and her role is bargaining with the board over future wage. The CEO has a quasi-linear per period utility of the form

$$u_t = (w_t - \gamma i_t^2)B \tag{10}$$

that is, the CEO has a linear utility from the pay and a convex disutility from board independence. The CEO dislikes a more independent board for two reasons. First, the more independent the board is, the less the CEO enjoys non-pecuniary private benefits. Second, the more the board is comprised of dependent directors, the better the board can advise the CEO since, being in most cases insiders of the firm, dependent directors have subject matter expertise and experience to help the CEO, and the CEO feels more comfortable communicating with them (similarly to the advising channel explored in Adams and Ferreira (2007)). Altogether, this term captures the CEO's disutility from board independence that is not reflected in the factors discussed before, that is, the board monitoring and the higher probability of being fired. The CEO's disutility for board independence is affected by the level of firm assets since the bigger the firm, the higher will be the level of control of a more independent board (or, the more the CEO would benefit from the advice of the board).

The value of the CEO's outside option, that is her utility if she leaves the firm or she is fired, is simply equal to the wage she bargained over at time $t-1$ [12]. The CEO's lifetime utility if she stays is given by

[11] The CEO's decision of leaving the firm can be easily endogeneized but I leave it as exogenous to simplify the analysis.

[12] In other words, I abstract from liquidation or retirement options and other factors that could make the CEO's outside option greater than simply her wage.

$$U_{stay}(\mu, \sigma, i, \tau, w) = w - \gamma i^2 + \beta E[U(\mu', \sigma', i', \tau + 1, w')|p'] \tag{11}$$

subject to Bayesian updating of the beliefs about the CEO's ability. Since the CEO's utility is expressed per unit of firm assets, B, I refer to w_t as the CEO's pay hereafter even though technically the actual CEO's compensation is $w_t B$. As for the board, the continuation value on the right-hand side of equation (11) equals $U_{leave}(\mu', \sigma', i', \tau', w') = w'$ if the CEO leaves, $U_{fire}(\mu', \sigma', i', \tau', w') = w'$ if the CEO does not leave but gets fired, and $U_{stay}(\mu', \sigma', i', \tau', w')$ if the CEO does not leave and is not fired.

Independence Choice and Bargaining Problem

At time t the board chooses the time $t+1$ board independence, i_{t+1}, that maximizes the total surplus of the match with the incumbent CEO, that is the sum of the surplus brought by the board and the current CEO, subject to the constraint that only a certain fraction δ of the board members can change each period, $|i_{t+1} - i_t| < \delta$. This restriction captures the fact that in reality, because of several factors such as firm bylaws, regulations and time constraints, the overwhelming majority of firms appoint a limited fraction of new directors, if they appoint at all. If there is no feasible i_{t+1} such that the total surplus is positive, the board fires the CEO. Once the future board independence has been decided, the board negotiates with either the incumbent or the new CEO over her time $t+1$ compensation, w_{t+1}, via Nash bargaining. Hence, the future compensation splits the surplus that is maximized by the future board independence decision. I denote by $\lambda \in [0, 1]$ the bargaining power of the board. During the bargaining, the CEO's outside option is her utility if she is fired, U_{fire}, while the board's threat point if no agreement is reached with the CEO is its utility if the CEO is fired, V_{fire}.

To simplify this problem and to derive the expression of the total surplus, I assume that, because of competition for scarce managerial talent[13], after a CEO turnover has occurred

[13]Since I am abstracting from firm size considerations, this assumption seems consistent with what it is observed in practice: in their first year in office, CEOs usually receive higher than average stock compensation and higher bonuses including signing bonuses.

but before the firm has been matched with a new CEO, the board offers to all the potential CEO candidates a minimum wage $w_p(i_t)$ and an initial level of independence equal to the current one, i_t. Hence, once a new CEO is matched with the firm, if the wage that splits the total surplus is greater than $w_p(i_t)$, the CEO obtains a compensation above the minimum, and a board with the level of independence that maximizes the total surplus. Otherwise, the new CEO obtains the minimum wage $w_p(i_t)$ and the board independence stays constant for one period[14] [15]. The CEO's minimum wage in the first period is the compensation level such that the expected value of the board's utility before being matched with a new CEO net of the sunk costs (i.e. the wage due to the previous CEO and the turnover costs) is zero. Mathematically,

$$w_p(i) = \{w' : E[V_{stay}(\mu_{s,0}, \sigma_{s,0}, i, 0, w; w')] + w = 0\} \tag{12}$$

Using this assumption, the Nash bargaining problem over the future CEO's pay can be written as

$$\max_{w'} \quad \Big[S_{board}\Big]^{\lambda}\Big[S_{CEO}\Big]^{1-\lambda} \tag{13}$$

$$\text{subject to} \quad S_{board} = V_{stay}(\mu, \sigma, i, \tau, w; w') - (-w - c_{firm} - c_{board,0} - c_{board,1}(1-i)^2) \tag{14}$$

$$S_{CEO} = U_{stay}(\mu, \sigma, i, \tau, w; w') - w \tag{15}$$

Except in the first period (when a new CEO is hired), the CEO is not fired if the total surplus is positive. The time t total surplus is

$$S(\mu, \sigma, i, \tau) = max\Big\{0, S_{stay}(\mu, \sigma, i, \tau)\Big\}$$

[14]This constraint is similar to the promise keeping constraint used in the optimal contract literature.

[15]An implicit assumption of this setting is that the CEO cannot be fired immediately after being matched with the firm, i.e. in the first period.

where

$$S_{stay}(\mu,\sigma,i,\tau) = \max_{i'}\Big\{S_{board} + S_{CEO}\Big\}$$

$$= \max_{i'}\Big\{\frac{\phi\mu}{1-\beta(1-\phi)} - \theta(1-i)p' - \gamma i^2 + (1-\beta)c_{firm} + (1-\beta(1-q(\tau)))c_{board,0}$$

$$+ ((1-i)^2 - \beta(1-q(\tau))(1-i')^2)c_{board,1} + \beta(1-q(\tau))E[S(\mu',\sigma',i',\tau')|p']\Big\} \tag{16}$$

$$\text{subject to} \quad |i'-i| < \delta. \tag{17}$$

The surplus expression is independent of wage, since the latter enters linearly in both the CEO's and the board's utilities.

Timing

The sequence of events is summarized as follows.

0) The firm is led by a CEO whose perceived ability is normally distributed with mean μ_t and variance σ_t^2, and that has spent τ_t periods working for the firm. The board is comprised of a fraction of independent directors i_t.

1) Given the beliefs about the incumbent CEO, her tenure, and the initial level of independence i_t, the board chooses the future board independence, i_{t+1}, subject to the constraint that $|i_{t+1} - i_t| \leq \delta$, to maximize the total surplus of the match with the incumbent CEO. If there is no feasible i_{t+1} such that the total surplus is positive, the CEO is fired.

2) The board bargains with the CEO (either the current or the new one) over her future wage, w_{t+1}.

3) Given the outcome of the previous stages, the board sets the intensity of monitoring, p_{t+1} (and so the probability of receiving the additional signal in the next period).

4) The firm profitability, Y_{t+1}, is realized and the board updates its estimate about the CEO's ability using the observed value of the firm's specific profitability, y_{t+1}. The updated belief distribution is $\mathcal{N}(\mu_{n,t+1}, \sigma^2_{n,t+1})$.

5) With probability $1 - e^{-p_{t+1}}$ the board receives the additional signal, z_{t+1}, and it updates again its belief distribution to $\mathcal{N}(\mu_{s,t+1}, \sigma^2_{s,t+1})$ where $\sigma^2_{s,t+1} < \sigma^2_{n,t+1}$.

6) With exogenous probability $q(\tau)$, the CEO leaves the firm.

Optimal Policies and Bargaining Outcome

The ultimate goal is to understand how the board composition and the belief distribution affect the board policies as well as the bargaining outcome. Interpreting the estimation results and identifying the model parameters require understanding the economics behind the model. To this end, I first analyze the monitoring, independence, and firing policies and then the bargaining outcome.

The optimal monitoring policy of the board is

$$p' = \begin{cases} log(\hat{p}') & \text{if} \quad \hat{p}' > 1 \\ 0 & \quad \text{otherwise} \end{cases} \tag{18}$$

where

$$\hat{p}' = \frac{\beta(1 - q(\tau))}{\theta(1 - i)}(E[V(\mu'_s, \sigma'_s, i', \tau', w')|\mu] - E[V(\mu'_n, \sigma'_n, i', \tau', w')|\mu]).$$

The monitoring intensity of the board is increasing in the discount factor, β, and in the probability that the CEO doesn't voluntary leave the company, $1 - q(\tau)$. Given the monitoring cost functional form, $\hat{p}'$ is decreasing in the monitoring cost parameter, θ, and the current fraction of dependent directors, $1 - i$. Moreover, the monitoring intensity $\hat{p}'$ is increasing in the difference between the expected value of the board continuation values if it receives the signal and if it doesn't. This term represents the board's expected marginal gain from receiving the additional signal at time $t+1$. This expected gain is not affected by the mean of the belief distribution about the CEO's ability, μ, if this is too high or too low. Indeed, if the CEO is perceived to be too bad (i.e. if μ is low) or sufficiently good (i.e. if μ is high), the board doesn't have much interest in understanding her exact level of ability. In the middle region the monitoring probability is hump-shaped in μ. Figure 1 illustrates the relationship between monitoring and the mean of the belief distribution for two different levels of independence based on parameters estimated in Section 3.5. In particular, it plots the average board's monitoring policy function on the level of the mean of the belief distribution, μ, and conditional on a high or a low independence, i. Given the nature of learning models, the difference between the two expected continuation values is increasing in σ as long as the variance of the belief distribution is not too high. In the latter case, the board does not monitor since the estimate of the CEO's ability is too imprecise and so the information gain is low. This is an important aspect that affects the identification of the parameters which are related to the monitoring, like the variance of the additional signal σ_z, and the standard deviation of the prior distribution σ_0: if the additional signal is too noisy (i.e. if σ_z is too high), or if the initial belief about the CEO's ability is too imprecise (i.e. if σ_0 is too high), the board does not have any incentive to monitor the CEO. Figure 2 illustrates the relationship between monitoring and the standard deviation of the prior belief distribution, σ_0. In particular, I plot the average monitoring intensity that is obtained by solving and simulating the model 20 times[16], each time corresponding to a different value of σ_0. For small values of this parameter the average board's monitoring intensity increases, while the mean starts decreasing when σ_0 gets too large. The monitoring intensity is positively affected by the future board independence, i', since the higher the board independence, the lower the board's firing cost in next period; if the board has a lower firing cost, it prefers a more precise estimate of the CEO's ability. Finally, the monitoring intensity, $\hat{p}'$, is increasing in the board's bargaining power, λ, since the greater the fraction of the surplus that the

[16] In all the figures in this section, the values of the parameters that do not vary are from Table 5.

board obtains from the bargaining, the higher its returns from monitoring, and hence its incentives to find out the CEO's ability level. Figure 3 illustrates the relationship between $\hat{p}'$ and λ. Specifically, I plot the average monitoring intensity that is obtained by solving and simulating the model 20 times, each time corresponding to a different value of the board's bargaining power, λ.

I analyze the firing policy by examining the expression of the surplus if the CEO stays, S_{stay}, provided by equation (16). The probability that the CEO is fired is negatively related to the mean of the belief distribution, μ, and to the standard deviation of the belief distribution, σ, since all agents are risk neutral. Figure 4 documents this relationship. Specifically, this graphs shows the average board firing decision on the level of the mean of the belief distribution, μ, conditional on a high and a low level of the standard deviation of the belief distribution, σ. It illustrates that for some levels of μ, some CEOs are retained by the board if the signal about their ability is noisy and they are fired otherwise. The probability that the CEO is dismissed is increasing in the level of board independence, i. Figure 5 depicts the relationship between the average firing decision and the mean of the belief distribution, μ, conditional on a high and a low level of independence, i. The logic behind this graph is similar to the one of the previous figure: for some level of the mean of the belief distribution about the CEO's ability, some CEOs would be fired if the board had a high fraction of independent directors and would be retained if the board had a low level of independence. Signing the effect of the monitoring disutility parameter on the CEO firing probability is difficult because the monitoring intensity is chosen endogenously by the board. Hence, a high θ hurts the CEO's survival chances insofar as the board has high incentives to monitor. The CEO's distaste of board independence, γ, increases the probability of a forced turnover. The firing probability is decreasing in the cost to the shareholders and the fixed cost to the board of firing, c_{firm} and $c_{board,0}$. A high board variable cost increases the CEO's survival chances as well, unless the board dependence increases and the current fraction of dependent directors is below the threshold $\delta\sqrt{\beta(1-q(\tau))}/(1-\sqrt{\beta(1-q(\tau))})$. This result means that if the board is made up of a high fraction of independent directors, adding a dependent director to the board increases the probability that the CEO is fired, since the board takes into account that firing her tomorrow will be more costly.

I now turn to analyzing the board's independence decision. The board chooses i' considering the marginal benefit and cost from board independence. As we have already discussed, the CEO has a disutility from board independence while, a higher i' lowers the board's monitoring cost and its future cost of firing the CEO. For example, when the current CEO is perceived of high (low) ability, the board is not interested (interested) in monitoring and in firing and so it will decrease (increase) its independence level since the only aspect that matters is the CEO's disutility from independence (CEO's disutility doesn't matter).

Finally, I turn to the bargaining over the future CEO's pay, w'. Since the compensation is set through a Nash bargaining problem, the standard results about this bargaining protocol apply. That is, the board receives a fraction λ of the surplus

$$\lambda S_{stay}(\mu, \sigma, i, \tau) = S_{board} \tag{19}$$

while the CEO receives the complementary fraction

$$(1-\lambda) S_{stay}(\mu, \sigma, i, \tau) = S_{CEO}. \tag{20}$$

Using the second condition, the bargained wage can be written as

$$w' = \frac{1}{\beta}\left((1-\lambda) S_{stay}(\mu, \sigma, i, \tau) + \gamma i^2 - \beta E[\bar{U}(\mu', \sigma', i', \tau+1)|p']\right) \tag{21}$$

where $\bar{U}(\mu', \sigma', i', \tau+1) \equiv U(\mu', \sigma', i', \tau+1) - w'$. This implies that the CEO's pay decreases in the bargaining power of the board, λ, and it is increasing in all variables that increase the total surplus.

Figure 6 depicts the correlation between the level of independence and CEO's pay chosen at time t, i' and w', on the mean of the belief distribution about the CEO's ability, μ, for a high and a low level of current board independence, i. I consider only levels of μ above the mean of the prior belief distribution, μ_0, to abstract from the effects of firing decisions. This figure shows that the model is able to accommodate several patterns that can be observed in the data according to the CEO's perceived ability, and the current composition of the board. If a board has a high level of independence, there is a positive correlation between the future CEO's pay and independence whether the CEO is perceived to be good or bad. Hence, in this case, the CEO obtains a higher (lower) wage if the board increases (decreases) its independence level. On the other hand, if a board has an initial low level of independence, the sign of the correlation between the future CEO's compensation and board independence depends on how good the CEO is perceived to be, that is on μ. Specifically, if the firm is led by a bad CEO (i.e. one with a low μ), the correlation is positive as in the previous case, while it is negative if the firm has a good CEO (i.e. one with a high μ). Thus, a good CEO obtains a less independent board as well as a higher wage in this scenario.

Table 1 summarizes the model variables, parameters, and the empirical measures that will appear more frequently in the paper.

[Insert Table 1 here]

3 Data and Estimation

This section first describes the data. Then it discusses how the model is solved computationally, the estimation procedure, and the identification strategy. Finally, it presents the estimation results.

3.1 Sample and Descriptive Statistics

The data come from several sources. I collect firm accounting data from the CRSP / Compustat merged annual database, and data about CEO compensation and CEO tenure from ExecuComp. Data about boards and directors come from the ISS Directors Legacy database up to 2006, and from the ISS Directors database from 2007 onwards. In order to link the two ISS databases with CRSP / Compustat, and to fix issues due to the change in methodology used to collect the data from Riskmetrics (the former provider of the ISS databases), I use the database constructed in Coles et al. (2014)[17]. In addition, I use data

[17] I thank Lalitha Naveen for making these data available on her website.

about forced CEO turnovers for the ExecuComp sample from Peters and Wagner (2014) and Jenter and Kanaan (2015). The Appendix provides further details on the sample construction.

The firm's total profitability (Y_t) is set equal to the firm's return on assets (ROA)[18]. The industry profitability (x_t) is computed as an industry equally-weighted average of the firms' ROA, using the Fama and French twelve industry classification. To compute the industry profitability I consider only firms that are larger than the median Compustat firm in terms of lagged assets in a given year[19]. The firm's excess profitability, defined as $y_t^* \equiv Y_t - x_t = y_t - \mathbb{1}\{tunover_t\}c_{firm}$, equals the firm's ROA minus the industry ROA.

To construct the moments of the baseline estimation, I consider a director as dependent if that individual is either not classified as independent in the ISS database (using the ISS item classification), which definition is consistent with the requirements of the major US stock exchanges, or if that individual is classified as an independent director in the database but has spent at least nine years on the board of the company[20]. Later I justify this choice in the identification subsection. Hence, Independent directors with a tenure greater than nine years, which are classified by the ISS as independent, are considered for the purpose of my estimation as dependent directors. I define independent directors with a tenure greater than nine years as tenured independent directors.

The CEO's wage rate (w_t) equals the total compensation including option grants (ExecuComp item TDC1) in year t divided by the midpoint value of total assets. Both this variable and the firm's ROA are winsorized at the 1st and 99th percentiles. Data about CEO forced turnovers is from Peters and Wagner (2014) and Jenter and Kanaan (2015), who classify a CEO succession as forced or voluntary following a similar procedure to that in

[18]The annual ROA equals operating income before depreciation (Compustat item OIBDP) divided by the midpoint of total assets (Compustat item AT) in the current and previous fiscal year

[19]I choose this measure for two reasons: to avoid bias from changes in the coverage of Compustat, and to compare the profitability of the firms in my sample with companies of similar size.

[20]The director's tenure is computed subtracting the year in which the director service began (ISS item DIRSINCE) from the fiscal year. I manually correct the starting year in case of misreports or re-elections after periods greater than one year.

Parrino (1997)[21]. All the financial variables are reported in 2014 US constant dollars. Table 2 provides descriptive statistics.

[Insert Table 2 here]

The final panel has 18,248 firm-year observations for 2,181 unique firms and 3,838 CEO spells from 1998 to 2014. The mean (median) firm assets in the sample is \$21.8 (\$2.9) billion. This number is quite large since the sample is biased toward large firms due to the availability of compensation and governance data. Board size is measured by the number of directors sitting on the board and it appears to be relatively constant in the sample: the mean (median) number of directors is 9.4 (9), and the difference between the 75th and 25th percentile is just 3 directors (11 versus 8). The fractions of dependent and tenured independent directors have similar distributions. The mean percentage of dependent and tenured independent directors according to ISS is 0.51 with a standard deviation of 0.2. The mean CEO compensation relative to the firm assets is considerably larger than the median and fluctuates considerably given the size of its standard deviation. On average, 2.3% of the firms in the sample dismiss their CEOs in a given year. The CEO spell length is computed as the difference between the dismissal or the voluntary leave year and the hiring year. The difference between the mean (median) tenure if fired and if voluntary leave equals 3.36 (2) years.

3.2 Model Solution

I solve the model computationally using backward induction starting from the last period a CEO could potentially work for the firm (i.e. $\bar{\tau}$). In each period, I first solve for the independence i_{t+1} that maximizes the total surplus subject to the constraint on the maximum fraction of directors that can change given in equation (17), taking into account the optimal monitoring from equation (18). If there is no i_{t+1} such that the total surplus is positive, the CEO is fired. If a forced turnover does not occur, I compute the optimal wage using

[21] Parrino (1997) classifies a CEO turnover as "forced" using press reports along with an age criterion and further refinements.

equation (21), and check that the it is non-negative, otherwise I set it to zero. Taking into account these results, I compute the expected values for the surplus, as well as the board and CEO continuation values (using a Gauss-Hermite quadrature with seven nodes), and I iterate this algorithm backward until the first period. Following Taylor (2010), I rewrite the value functions and the total surplus in terms of CEO's relative ability. This is measured as the difference between the CEO's ability, μ_t, and the mean of the prior belief distribution, μ_0. This implies that what affects the board's decisions and the bargaining outcome is the relative ability of the CEO rather than the absolute ability[22].

I estimate some of the parameters outside the SMM procedure. In particular, I set the discount factor, β, equal to 0.918 consistent with the yearly average equity return for shareholders in my sample (which goes from 1998 to 2014) of 8.9%. I set the maximum number of periods a CEO works for the firm equal to 15, i.e. $\bar{\tau} = 14$. In my sample, 14 years correspond to the 83rd percentile of the distribution of CEO's tenure. I consider that only one seat on the board can change every period by setting $\delta = 1/9$[23]. I use this parameter value since, in my sample, the average fraction of new directors is 7% for the whole sample, and 16% in the case of firms whose boards have appointed new directors. Moreover, consistent with the evidence in Matveyev (2016), I find that over two thirds of the companies, which appoint new directors, hire one independent director per year. Finally, I estimate the probability that the CEO leaves, $q(\tau)$, as the frequency of a voluntary CEO turnover after τ years conditional on the CEO surviving $\tau - 1$ years. These parameters set, there are eleven parameters to estimate.

3.3 Simulated Method of Moments Estimator

I estimate the remaining parameters $b = \{\mu_0, \sigma_0, \phi, \sigma_\epsilon, \sigma_z, \gamma, \theta, c_{firm}, c_{board,0}, c_{board,1}, \lambda\}$ via SMM (Lee and Ingram (1991)). The basic intuition of the SMM procedure is to find a set

[22] This change of variable doesn't affect any of the insights and interpretations discussed in Section 2.

[23] The grid of the percentage of independent directors, i, is built using the evidence of Table 2: since both the median and the mean number of directors on the board are close to nine and since the CEO sit on the board in the overwhelming majority of firms in my sample, I let i have nine grid points in the interval from zero to eight-ninth in my baseline exercise.

of parameter values so as to minimize the distance between a set of moments simulated from the model and their counterparts in the data[24]. So, for a given set of parameters, I solve the model and use the solution to generate simulated data. Next, I calculate a set of moments using the simulated panel. Based on the distance between the model generated statistics and their empirical counterparts, the values of the structural parameters are updated until convergence. Formally, define $\tilde{h}$ as the vector of empirical moments, $\hat{b}$ as the set of parameter estimates, $h_k(b)$ as the vector of simulated moments for parameter vector b and simulation k. The SMM estimator of b, $\hat{b}$, is then defined as the solution to the minimization of

$$\hat{b} = \underset{b}{argmin}\left(\tilde{h} - \frac{1}{K}\sum_{k=1}^{K} h_k(b)\right)' \hat{W}\left(\tilde{h} - \frac{1}{K}\sum_{k=1}^{K} h_k(b)\right) \tag{22}$$

where $\hat{W}$ is a positive definite weighting matrix that converges in probability to a deterministic positive definite matrix W, and K is the total number of simulations. I compute the weighting matrix $\hat{W}$ as the inverse of the covariance matrix of the empirical moments constructed using the influence function approach of Erickson and Whited (2002). I set $K = 10$ since Michaelides and Ng (2000) show that a simulated sample ten times as large as the empirical sample has good finite sample properties. To minimize the econometric objective function, I use a simulated annealing optimization algorithm to avoid local minima.

3.4 Identification Strategy

Before talking about the identification of the structural parameters, I discuss the classification of dependent and independent directors in the computation of the empirical moments. As already mentioned, I consider as dependent directors not only those not classified as independent according to the major US stock exchanges rules, but also those classified as independent that have served on the same board for over nine years. The reason for this assumption is twofold. First, outside the United States, over the last 20 years several countries have adopted tenure-related regulations regarding the definition of independent directors. For instance, the United Kingdom corporate governance code states that a board of direc-

[24]See Strebulaev and Whited (2012) for a review of the structural estimation literature in corporate finance.

tors should explain the reasons why a director who has served for more than nine years qualifies as independent. In France, a director loses the independence status after having served twelve years. The European Commission recommends that an independent director does not serve more than twelve years on the board of the same company. In Hong Kong, an independent director cannot work for the same company more than nine years, unless approved by the shareholders. Second, since 2014 the ISS has included director tenure in its company governance ratings. In particular, the threshold of nine years that I use is the same considered by the ISS in its statement "A tenure of more than nine years is considered to potentially compromise a director's independence" [25]. In support of this position, during the years of my sample, the percentage of dependent directors decreased from 39% to 19% because of the new requirements, while the percentage of independent directors with more than nine years of tenure went from 18% to 31%[26].

I now describe and explain the choice of moments that I match. The success of SMM relies on model identification, which in turn requires that the econometrician chooses moments that are sensitive to variations in the underlying structural parameters. In choosing the moments, it is important not to "cherry-pick" since the goal is to understand what features of the data the model can and cannot reconcile. Further, it is essential to keep in mind that, even though most of the decisions in the model are intertwined and all the parameters should affect the moments in some way, some of these moments are particularly useful for the identification of certain parameters. In total I use sixteen moments to estimate the eleven model parameters. The first four moments are related to the firm excess profitability. The first moment is the serial correlation of the firm's excess profitability, y_t^*, which I calculate using the method in Han and Phillips (2010) to account for firm and time fixed effects. This statistic helps to identify the persistence parameter of the firm's specific profitability, ϕ. The second and third moments are the coefficients of a regression of the firm's excess profitability on a dummy that equals one if the CEO was fired from period $t-2$ to $t+2$ and a constant. The first coefficient

[25] Huang and Hilary (2018) discuss this debate and find evidence that excessive average board tenure has negative effects on the firm value.

[26] See the Appendix for empirical evidence.

identifies the firm cost of a turnover, c_{firm}, since the higher the firm cost of a turnover, the lower the profitability of the firm in that window of time. The regression intercept is related to the mean of the prior belief distribution about the CEO's ability, μ_0. The higher the average quality of the pool of CEOs, the higher is the average firm profitability. The fourth moment is the variance of the firm's excess profitability, which is computed after removing firm and year heterogeneity, following the structural corporate finance literature (see, for instance, Hennessy and Whited (2005) and Hennessy and Whited (2007)). This variance helps to identify the standard deviation of the shock to the firm's specific profitability, σ_ϵ.

The second group of moments is made of statistics related to the percentage of fired CEOs over their careers. I compute four forced turnover hazard rates. Specifically, I compute the percentage of fired CEOs in four different tenure categories, which are one to three years, four to six years, seven to nine years and over ten years. These four moments identify the cost of a turnover to the shareholders, c_{firm}, the board fixed cost of firing the CEO, $c_{board,0}$, and the board variable cost of firing the CEO, $c_{board,1}$. To improve the accuracy of the identification of these three parameters I also compute other three moments. The first one is the coefficient of the regression to identify c_{firm}. The second one is the correlation between a dummy that equals one if a forced CEO turnover occurs, and zero otherwise, and the percentage of dependent directors. This provides information about the variable cost of firing the CEO, since the higher $c_{board,1}$ is, the more negative is this correlation. This moment is computed after removing industry and year heterogeneity rather than firm and year fixed effects since, as documented by Table 2, the percentage of forced turnovers in my sample is substantially low[27]. Moreover, I compute the frequency with which the percentage of dependent directors decreases since this moment is useful to identify the personal fixed cost of the board of firing, $c_{board,0}$. The intuition behind this relation is that the more costly to the board is to fire the CEO, the less the board is going to monitor the CEO. Consequently, the higher the value of $c_{board,0}$, the less frequently the board decreases the number of dependent directors, since

[27] The correlation computed after removing firm and year fixed effects has the same sign and it is slightly smaller in magnitude but it is not significant. Computing the same correlation including other regressors such as assets and age, I obtain a really close value to the one I am using.

monitoring becomes less relevant.

The third set of moments are statistics related to the composition of the board. Besides the correlation and the frequency moments, I compute the average percentage of dependent directors in the four tenure categories considered for the forced turnover hazard rates. These statistics are important to identify four parameters: the CEO's disutility from board independence, γ, the variance of the prior belief distribution about the CEO's ability, σ_0, the standard deviation of the additional signal, σ_z, and the board's disutility of monitoring, θ. Obviously, the higher the CEO's disutility from board independence, the larger the average fraction of dependent directors. The variance of the prior belief distribution is mainly identified by the mean fraction of dependent directors in the first 3 years of CEO's tenure: if the uncertainty about a new CEO is low, the directors are not going to care much about monitoring; conversely, if the uncertainty about a new CEO is high, the board will keep or increase its independence. The standard deviation of the additional signal, z_t, is identified by how board dependence evolves over time, especially during the first years that the CEO is in charge. For instance, suppose that σ_z is low, meaning that the additional signal is particularly informative, then, if the board receives a positive signal, it is highly likely that the CEO is skilled and, for this reason, the directors will be more prone to have a board with more dependent directors. If instead the initial signal is negative, the board will fire the CEO. Suppose instead that σ_z is particularly high. In this case, the board will need to monitor more the CEO if the firm profits are not particularly good, and for this reason the board will be less inclined to give seats to dependent directors. In sum, the standard deviation of the additional signal affects how board dependence changes during the CEO tenure, since board dependence affects the monitoring intensity and the probability of the CEO being fired. Finally, the board's disutility of monitoring the CEO, θ, is identified by the average fraction of dependent directors in the case of CEOs with a tenure above 10 years. This relationship is due to the fact that, if monitoring is costly, it's more likely that the board will retain independent directors during the first years of the CEO's tenure so to save on the monitoring cost. The more costly monitoring is, the more independent directors will be on the board when the CEO starts the last years of career.

There are two potential concerns about the identification strategy outlined so far. The first is about convincingly and separately identifying the standard deviation of the additional signal, σ_z, and the cost of monitoring, θ. To overcome this problem I use another moment that is sensitive to the changes in the first parameter, but not to the second one, that is the variance of the CEO wage rate in the first year [28]. This statistic is sensitive to changes in the noise of the additional signal, since the board receives an additional signal before bargaining the initial wage, and therefore the board's initial beliefs about the CEO's ability are directly affected by the noise of the additional signal. More specifically, there is a negative relation between the variance of the initial wage and the variance of the additional signal given that, the noisier is this signal, the less the board's beliefs will change, which causes the variance of the initial CEO's wage to drop.

The second concern about the identification strategy is the possibility that the standard deviations of the prior belief distribution and of the additional signal have opposite effects to the ones just described. That is, if the initial standard deviation is too high or the additional signal is too imprecise, the board may decide to have more dependent directors since it has no incentives to monitor the CEO. To address this potential concern, I show that this effect occurs only for particularly high values of σ_0 and σ_z, which are almost three and four times greater than the estimated values, respectively. Thus, given my data sample, there is a negative relationship between board dependence and noise.

The last moment I use is the average CEO wage rate which is included to capture the board bargaining power parameter, λ. In the Appendix, I illustrate how the moments relate to the parameters using comparative statics. In particular, I set all parameters to their SMM estimated values and document how moments vary when the value of the parameter of interest changes.

[28] As for the previous variance, also this moment is computed after removing firm and year heterogeneity.

3.5 Estimation Results

Table 3 and 5 presents the results from the SMM estimation of the model, with Table 3 reporting the actual versus the simulated moments, and Table 5 reporting the parameter estimates along with their standard errors. I begin by describing how the model fits the data and in particular by examining the sixteen moments individually to gauge where the model fails and where it is successful. The model matches the first four moments related to the firm's excess profitability fairly closely. The only profitability related moment with which it disagrees is the variance of the profitability, probably because of the magnitude of the impact of the shock to profits on the board's learning speed. The model performs well with the moments related to the percentage of CEOs fired, but it generates too few firings when a CEO has spent between four and six years at the firm resulting in the peak of firing occurring when the CEO has spent between 7 and 9 years in the simulated data. This result highlights that learning occurs slowly in my simulation with respect to what might happen in reality. The model fits remarkably well the five moments related to the fraction of dependent directors on the board, which are the average percentages of dependent directors over the CEO's career, and the frequency with which the percentage of dependent directors decreases, as well as the mean CEO pay scaled by firm assets. The model tends to produce a too negative of a correlation between CEO's firing and board dependence. This is likely to be due to the assumptions about board independence. In reality, a board could not have the possibility to change its composition at all for reasons such as firm bylaws and regulations that are not always observable. Finally, the variance of the CEO's wage rate in her first period in office is too low since the model assumes random matching between CEOs and firms when in reality there could be assortative matching based on synergies between executives and companies (see Tervio (2008) and Gabaix and Landier (2008)).

Moreover, Table 4 shows that the model does a good job of matching the distribution of the fraction of dependent directors, as well as moments about the average CEO's spell length, even though these moments are not directly targeted in the estimation.

The parameter estimates, and their standard errors, are in Table 5. To begin, it is worth emphasizing that most of my parameter estimates are within the range of those of Taylor (2010) even though I use a different sample and different moments to identify certain parameters, and the two models have certain striking differences. For instance, in Taylor (2010) model the additional signal is received in each period because monitoring occurs exogenously in every period whereas in my model monitoring doesn't always occur because it is endogenously determined. My estimates of the board cost are different from those of Taylor (2010) for two reasons. First, in Taylor (2010) the cost of firing doesn't depend on independence. Second, in my model the board cost is expressed in dollars while in Taylor (2010) the personal cost is expressed in utils (that is, the board cost doesn't corresponds to an actual loss for the directors).

Now, I present the parameter estimates. Since some of the parameters are expressed as a percentage of the firm assets, when possible I report also their dollar value using the median value of assets in my sample. The mean of the prior belief distribution, μ_0, is 1.8% or \$52 million. The shareholder cost of a turnover, c_{firm}, is 2.2%. This implies that firing the CEO costs to a firm in my sample is \$64 million, which is not unreasonable given the findings of the previous empirical studies (see, for instance, Yermack (2006)). The persistence parameter, ϕ, is 0.135 implying a highly persistent firm-specific profitability. The standard deviation of the firm profitability, σ_ϵ, is 1.92% and the standard deviation of the additional signal, σ_z, is 5.81%. Using the methodology adopted by Taylor (2010), this implies that the ratio $\sigma_\epsilon/(\phi\sigma_z)$, which measures by how much the additional signal influences the board's beliefs more than the earnings signal, equals 2.45 which is consistent with the findings of Cornelli et al. (2013). The prior standard deviation of the beliefs distribution is 1.07%, which implies that the difference in average profitability between a CEO at the 95th percentile of the distribution and one at the 5th percentile is equal to 3.5% of the firm assets. The value of the CEO's disutility from board independence equals $\gamma = 0.0006/1000$. This number implies that a CEO working for a firm with a board with five out of nine independent directors has a disutility of \$537, while a CEO who is the only dependent director in a firm with the same size has a disutility of \$1,375. While these numbers seem small, the distaste term of the CEO's utility captures perks and private benefits and does not reflect CEO's career concerns (i.e. impact on her pay or on her survival chances). The board cost of monitoring equals θ=0.0019/1000. The inverse of this estimate corresponds to 0.526 millions which roughly corresponds to the median equity holding value of an independent director in my sample. The board bargaining power, λ, is 5.9%. This value is smaller than what previous papers in the structural literature found (see, Page. (2018) and Frydman and Papanikolaou (2018)). However, as emphasized before, it is important to keep in mind that, not only I am targeting different moments with respect to the other studies, but also that λ has an indirect effect on the board composition: indeed, as the board bargaining power increases, the average fraction of dependent directors drops since the board has a higher return from monitoring. If one considers this effect, this value is consistent with the empirical literature documenting a strong influence of CEOs on the board composition (see Bebchuk and Fried (2004))[29]. Finally, the fixed and the variables costs to the board of firing the CEO, $c_{board,0}$ and $c_{board,1}$, are 0.12% and 0.14% of the firm assets, respectively. This implies that the board with five out of nine independent directors has a firing cost of $4.3 million. This means that, to rationalize the board firing rate, one could think that firing the CEO would cost the directors roughly three times what they earn in a year for their director service (see Fedaseyeu et al. (2018)). This result can be reconciled if one considers not only the uncompensated stress and effort around a CEO succession but also the fact that by firing a CEO directors might lose additional benefits, such as directorships on other boards as well as promotions to executive positions within and outside the firm.

4 Counterfactuals

In this section, I use the estimated model to quantify the extent to which the misalignment of incentives between the board and the shareholders destroys shareholder value and changes the CEO's average career earnings. Next, I evaluate how mean shareholder value and CEO's

[29]Taylor (2013) finds that the average CEO captures half of the surplus from good news and bears none of the negative surplus from bad news. Since I do not distinguish between good and bad news, my CEO's bargaining power estimate is not much different.

average career earnings change from removing the board monitoring and from increasing the board's bargaining power. Last, I use the model to evaluate the impact of a potential policy reform requiring boards to have a minimum percentage of independent directors with tenure below nine years.

4.1 Removing the Misalignment of Incentives

Table 6 reports, in percent, the effects of removing the board's cost of firing the CEO as well as the CEO's disutility from board independence. The variables that I consider for these exercises are the average value for the firm's excess profitability, y_t^*, the percentage of CEOs fired in a year, the average value for the fraction of dependent directors, $1 - i_t$, the CEO's pay scaled by assets, w_t, the average value for the shareholder value computed as the discounted sum of the firm's specific cash flows scaled by assets, $d_t = Y_t - x_t - w_t$, the CEO's average career earnings, which equals the product of the mean CEO's tenure by the mean CEO's pay scaled by assets, and average value for the board monitoring, which equals the probability of receiving the additional signal, $1 - e^{-p_{t+1}}$.

Before proceeding to describe the results, it is important to emphasize that the shareholder value computed as above represents a market-to-book ratio in excess of the average industry market-to-book ratio. Adding back the industry profits would force me to take a stand about what would happen to the firm's industry competitors which is problematic to forecast. Given that the firms in my sample are the most productive in the US economy, it is likely that the effects I find are downward biased.

Column (1) of Table 6 shows the results of the benchmark model and Column (2) the percent changes after removing the board's fixed cost of firing, setting $c_{board,0} = 0$. As expected from the discussion about identification, this parameter affects the board dependence, which drops by 34.74% causing an increase of 53.36% in the probability of arrival of the additional signal. This positively affects the CEO firing rate since the board discovers bad CEOs faster, and the profitability rises by 0.69%. Altogether, setting $c_{board,0} = 0$ has a positive effect on the shareholder value, which increases by 1.79% (corresponding to \$52 million using the median firm assets value), and a negative impact on the CEO's career earnings decrease, which decreases by 7.9% ($5.66 million). Column(3) shows what happens when the board's variable cost of firing is set to zero, $c_{board,1} = 0$. Perhaps counterintuitively, the monitoring intensity of the board decreases reducing the probability of arrival of the additional signal by 12.59%. This can be explained by the fact that the board has less incentive to monitor than in the benchmark case, where the board monitors to bargain over the CEO's pay as well as about its independence level. Despite the lower cost of firing the CEO, the reduction in monitoring results in a lower firing rate which declines by 5.94% and the profitability slightly drops by 0.05%. The smaller surplus results in the board paying less the CEO whose average career earnings drop by 3.12% ($2.23 million). Saving on the CEO's pay causes the shareholder value to increase by 0.65% ($19 million). Column (4) shows that removing the misalignment of incentives with regard to the firing decision increases the shareholder value by 2.08% ($60 million), and the CEO's earnings by 9.42% ($6.75 million).

Another interesting finding is the effect of removing the CEO's disutility from board independence. As discussed in the model section, this parameter negatively affects the CEO's survival chances. However, by setting γ to zero, the board does not have any interest in board independence besides reducing the monitoring intensity (as in Hermalin and Weisbach (1998)). This effect generates a 6.09% increase of the probability of arrival of the additional signal, which causes both the firing rate and the firm's profitability to increase by 1.85% ad 0.15%, respectively. Overall, the shareholder value goes up by 0.52% ($15 million) while the CEO's earnings fall by 2.07% ($1.48 million).

4.2 Changing the Monitoring Incentives

Table 7 shows, in percent, the effects of shutting off the board monitoring and setting the bargaining power between the two parties equal. Column (2) shows the percent changes to the model outcome from removing the board monitoring. Unsurprisingly, the firing rate drops by 8.7% making the profitability to drop by 1.26%. However, since the board cannot monitor and therefore cannot figure out quickly enough whether the firm is managed by a good CEO, the CEO's average earnings drops by 4.25% ($3.04 million), despite the higher survival chances. Ultimately, the negative effect on the profitability dominates the positive one due to the lower CEO's pay so that the shareholder value decreases by 0.69% ($20 million). Column (3) reports the changes to the model outcome when the board monitoring is removed and the board's variable cost of firing the CEO is set to zero. This case illustrates that, when the board cannot monitor, removing the dependent directors' additional cost of firing has a positive effects on the shareholder value from a higher profitability and a lower average CEO's pay. In this case, since the CEO's average earnings drop by 8.64% ($6.19 million) and the drop in profitability is smaller than the one in the previous case, the shareholder value increases by 0.16% ($5 million). Column (4) shows that removing both the board monitoring and the misalignment of incentives between the board and the shareholders about the firing decision has a net positive effect on the shareholder value. Similarly to what happens when the monitoring and the variable cost are removed, the profitability declines by 0.9%. However, since the fraction of fired CEOs increases by 23.31% in this case, the reduction is likely to be caused by the difficulty that the board has to distinguish good from bad CEOs. The fact that the board has no personal cost of firing the CEO leads to a drop of 15.84% ($11.35 million) of CEO's average career earnings, and the shareholder value rises by 1.58% ($46 million). Removing the personal board's cost of firing the CEO can partially offset the negative effect of the drop in productivity from not monitoring, but it reduces the CEO's ability of extracting rents from the shareholders, thus increasing the shareholder value.

Given the benefits of monitoring, I study what happens by increasing the board bargaining power. As mentioned before, increasing the board's bargaining power reduces the share that the CEO appropriates resulting in higher incentives for the board to monitor. In other words, the board captures a larger fraction of the return from its monitoring effort. Column (5) shows percent changes from increasing the board's bargaining power to 50%. As expected, the CEO's pay declines by 36.34%, and the CEO's average career earnings are cut by 43.94% ($31.47 million). The key result of this exercise is that the board is more reluctant to give up its independence, which increases by 55.9%, and monitoring increases by almost 3 times. As a result, the firing rate doubles, and the profitability grows by 2.79%. The combined effect increases the shareholder value by 10.42% ($302 million). Overall, the results of Table 7 show that monitoring is not just about reducing the CEO's rent extraction. It is also about the increase in firm's profitability.

4.3 Policy Counterfactuals

Next, I examine introducing a policy requiring US public companies to have a minimum number of non-tenured (less than or equal to nine years of tenure) independent directors sitting on the board, as it has been practice in many countries. I consider three thresholds of this policy: having one, three and five out of nine non-tenured independent directors sitting on the board. Table 8 shows, in percent, the effects of these reforms. Columns (2)-(4) show the percent changes under the three different independence requirements. Column (2) shows that having a minimum of one non tenured independent director out of nine directors has small positive effect on shareholder value, which increases by 0.2% ($6 million), and on the CEO's average career earnings, which drop by 0.43% ($0.31 million). This result is due to the fact that this constraint is rarely binding (for instance, in the baseline model less than 5% of the boards are made up by solely dependent and tenured independent directors). Column (3) shows the percent change of having a minimum of three non tenured independent directors out of nine directors. Such a rule has an effect more than three times greater than the previous one both on the shareholder value, which increases by 0.65% ($19 million), and on the CEO's average earnings that drop by 1.52% ($1.09 million). Such policy increases board monitoring by 2.75%, which is still modest relative to the baseline model but sufficient to increase the profitability by 0.34%. The change in the CEO's firing rate is lower than the one under the previous rule. The increment in CEO's firing rate is lower than in the previous case. This is due to the fact that, while today is cheap for a board to fire the CEO, the board understands that later firing will not become more expensive. Finally, column (4) illustrates the changes due to the introduction of a minimum fraction of non-tenured independent directors of five out of nine directors. Under this reform the shareholder value increases by 0.8% ($174 million) which is more than under the other two policy experiments. However, this increment is mainly due to the reduction of CEO's rent extraction, since the CEO's average earnings decrease by 2.4% ($13 million), than to a profitability increase, since profitability increases less than under the three out of nine directors case (0.29% versus 0.34%). Notice that this reform does not seem to incentivize the board to fire the CEO more often like in the previous cases despite the increase in monitoring of 8.81%. Altogether the results indicate that a minimum threshold number of non-tenured independent directors benefits shareholders both by an increase in firm profitability from more monitoring as well a reduction in CEO's rent extraction from a lower board cost of firing the CEO. However, the gains resulting from these reforms are relatively modest when compared with those in the previous subsection. Therefore, it seems more relevant to reduce the CEO's influence on the board decisions rather than introducing legislation requiring a minimum fraction of non-tenured independent directors.

5 Conclusion

I develop a dynamic model that includes passive learning about the CEO's ability, monitoring, costly CEO firing, negotiations between the board and the CEO over the manager's compensation, as well as endogenous board independence. Board independence enhances the board effectiveness by lowering the costs of monitoring and of firing the CEO. However, since board independence is chosen to maximize the total surplus of the match with a CEO's perceived ability, the independence choice takes into account the CEO's characteristics as well as the CEO's disutility from board independence. I estimate the model, and I quantify the importance of the different channels through which a board contributes to shareholder value. I show that removing the CEO's entrenchment by eliminating the cost to the board of firing the CEO increases shareholder value by decreasing the CEO's ability to extract rents. In contrast, board monitoring increases shareholder value by increasing the firm's profitability. I find that a CEO obtains a large portion of the bargaining surplus when wage negotiations take place. Setting the board's bargaining power equal to that of the CEO increases the firm's profitability by 3% and decreases the CEO's career earnings by 44%. Introducing regulation setting a minimum threshold number of non-tenured independent directors has a quantitatively smaller effect on the firm profitability, and so on the shareholder value. These results underscore the importance of giving the right incentives to the members of the board to be effective monitors from their higher utility from firm cash flows by reducing the CEO's rent extraction.

This paper represents a first attempt to understand the importance of board composition for shareholder value using a structural estimation approach. Many of the results obtained herein would have been hard to obtain with more conventional methods given that governance structure in companies is endogenous as well as the difficulty to observe how the board perceives the CEO's ability even resorting to instrumental variables.

This study opens up several avenues for future research. A fruitful extension of this analysis would be to incorporate assortative matching based on skills between firms and executives. In addition, embedding a richer set of directors' characteristics into this setting could achieve richer implications regarding boards' characteristics.

References

Adams, R. B., 2017. Boards, and the directors who sit on them. The Handbook of the Economics of Corporate Governance 1, 291–382.

Adams, R. B., Ferreira, D., 2007. A theory of friendly boards. Journal of Finance 62, 217–250.

Adams, R. B., Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. Journal of Financial Economics 94, 291–309.

Adams, R. B., Hermalin, B. E., Weisbach, M. S., 2010. The role of boards of directors in corporate governance: A conceptual framework and survey. Journal of Economic Literature 48, 58–107.

Ahern, K. R., Dittmar, A. K., 2012. The changing of the boards: The impact on firm valuation of mandated female board representation. Quarterly Journal of Economics 127, 137–197.

Amihud, Y., Stoyanov, S., 2017. Do staggered boards harm shareholders? Journal of Financial Economics 123, 432–439.

Baghat, S., Black, B., 1999. The uncertain relationship between board composition and firm performance. Business Lawyer 54, 921–963.

Baldenius, T., Melumad, N., Meng, X., 2014. Board composition and CEO power. Journal of Financial Economics 112, 53–68.

Bebchuk, L., Fried, J., 2004. Pay without performance: The unfulfilled promise of executive compensation. Harvard University Press Cambridge .

Chemmanur, T. J., Fedaseyeu, V., 2018. A theory of corporate boards and forced CEO turnover. Management Science 64, 4798–4817.

Coles, J. L., Daniel, N. D., Naveen, L., 2008. Boards: Does one size fit all? Journal of Financial Economics 87, 329–356.

Coles, J. L., Daniel, N. D., Naveen, L., 2014. Co-opted boards. Review of Financial Studies 27, 1751–1796.

Cornelli, F., Kominek, Z., Ljungqvist, A., 2013. Monitoring managers: Does it matter? Journal of Finance 68, 431–481.

Donaldson, J. R., Malenko, N., Piacentino, G., 2020. Deadlock on the board. Review of Financial Studies .

Erickson, T., Whited, T., 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. Econometric Theory 18, 776—799.

Fahlenbrach, R., Low, A., Stulz, R. M., 2010. Why do firms appoint CEOs as outside directors? Journal of Financial Economics 97, 12–32.

Fedaseyeu, V., Linck, J. S., Wagner, H. F., 2018. Do qualifications matter? new evidence on board functions and director compensation. Journal of Corporate Finance 48, 816–839.

Fisman, R. J., Khurana, R., Rhodes-Kropf, M., Yim, S., 2014. Governance and CEO turnover: Do something or do the right thing? Management Science 60, 319–337.

Frydman, C., Papanikolaou, D., 2018. In search of ideas: Technological innovation and executive pay inequality. Journal of Financial Economics 130, 1–24.

Gabaix, X., Landier, A., 2008. Why has CEO pay increased so much? Quarterly Journal of Economics 123, 49–100.

Garlappi, L., Giammarino, R., Lazrak, A., 2017. Ambiguity and the corporation: Group disagreement and underinvestment. Journal of Financial Economics 125, 417–433.

Graham, J. R., Kim, H., Leary, M., 2020. CEO-board dynamics. Journal of Financial Economics 137, 612–636.

Guthrie, K., Sokolowsky, J., Wan, K.-M., 2012. *ceo* compensation and board structure revisited. Journal of Finance 67, 1149–1168.

Han, C., Phillips, P. C. B., 2010. GMM estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory 26, 119–151.

Harris, M., Raviv, A., 2006. A theory of board control and size. Review of Financial Studies 21, 1797–1832.

Hennessy, C., Whited, T., 2005. Debt dynamics. Journal of Finance 60, 1129–1165.

Hennessy, C., Whited, T., 2007. How costly is external financing? evidence from a structural estimation. Journal of Finance 62, 1705–1745.

Hermalin, B. E., 2005. Trends in corporate governance. Journal of Finance 60, 2351–2384.

Hermalin, B. E., Weisbach, M. S., 1998. Endogenously chosen boards of directors and their monitoring of the CEO. American Economic Review 88, 96–118.

Hirshleifer, D., Thakor, A., 1994. Managerial performance, boards of directors and takeover bidding. Journal of Corporate Finance 1, 63–90.

Huang, S., Hilary, G., 2018. Zombie board: Board tenure and firm performance. Journal of Accounting Research 56, 1285–1329.

Jenter, D., Kanaan, F., 2015. CEO turnover and relative performance evaluation. Journal of Finance 70, 2155–2184.

Laux, V., 2008. Board independence and CEO turnover. Journal of Accounting Research 46, 137–171.

Lee, B.-S., Ingram, B. F., 1991. Simulation estimation of time-series models. Journal of Econometrics 47, 197–205.

Levit, D., Malenko, N., 2016. The labor market for directors and externalities in corporate governance. Journal of Finance 71, 775–808.

Matveyev, E., 2016. The labor market for corporate directors. Mimeo .

Michaelides, A., Ng, S., 2000. Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators. Journal of Econometrics 96, 231–266.

Page., T. B., 2018. CEO attributes, compensation, and firm value: Evidence from a structural estimation. Journal of Financial Economics 128, 378–401.

Parrino, R., 1997. CEO turnover and outside succession: A cross-sectional analysis. Journal of Financial Economics 46, 165–197.

Peters, F. S., Wagner, A. F., 2014. The executive turnover risk premium. Journal of Finance 69, 1529–1563.

Ravina, E., Sapienza, P., 2009. What do independent directors know? evidence from their trading. Review of Financial Studies 23, 962–1003.

Song, F., Thakor, a. V., 2006. Information control, career concerns, and corporate governance. Journal of Finance 61, 1845–1896.

Stern, L., 2019. Learning about directors. Mimeo .

Strebulaev, I. A., Whited, T. M., 2012. Dynamic models and structural estimation in corporate finance. Foundations and Trends in Finance 6, 1–163.

Taylor, L. A., 2010. Why are CEOs rarely fired? evidence from structural estimation. Journal of Finance 65, 2051–2087.

Taylor, L. A., 2013. Ceo wage dynamics: Estimates from a learning model. Journal of Financial Economics 108, 79–98.

Tervio, M., 2008. The difference that CEOs make: An assignment model approach. American Economic Review 98, 642–68.

Weisbach, M. S., 1988. Outside directors and *ceo* turnover. Journal of Financial Economics 20, 431–460.

Yermack, D., 1996. Higher market valuation of companies with a small board of directors. Journal of Financial Economics 40, 185–211.

Yermack, D., 2006. Golden handshakes: Separation pay for retired and dismissed CEOs. Journal of Financial Economics 41, 237–256.

Table 1: Variable Definitions
This tables presents definitions of the variables used throughout the paper.

Notation	Definition
Y_t	Total firm profitability in year t (% of assets)
x_t	Industry average profitability in year t (% of assets)
y_t^*	Firm excess profitability in year t (Y_t - x_t) (% of assets)
y_t	Firm specific profitability in year t (% of assets)
D_t	Firm cash flows in year t
d_t	Firm specific cash flows in year t ($Y_t - x_t - w_t$) (% of assets)
α	CEO ability (% of assets)
μ_0	Mean of board's prior beliefs about new CEOs ability before hiring (% of assets)
σ_0^2	Variance of board's prior beliefs about new CEOs ability before hiring (% of assets)
$\mu_{s,0}$	Mean of board's prior beliefs about new CEOs ability after hiring (% of assets)
$\sigma_{s,0}^2$	Variance of board's prior beliefs about new CEOs ability after hiring (% of assets)
μ_t	Mean of board's beliefs about new CEOs in year t (% of assets)
σ_t^2	Variance of board's beliefs about new CEOs in year t (% of assets)
σ_z^2	Variance of the the additional signal about the CEO ability (% of assets)
c_{firm}	CEO turnover cost to the shareholders (% of assets)
$c_{board,0}$	Fixed board personal cost of firing the CEO (% of assets)
$c_{board,1}$	Variable board personal cost of firing the CEO (% of assets)
θ	Board disutility from monitoring the CEO
γ	CEO disutility from board independence (% of assets)
τ	CEO tenure
$q(\tau)$	Probability that a CEO with tenure τ leaves the firm
ϕ	Persistence parameter of the firm profitability
σ_ϵ^2	Variance of the the firm specific profitability shock (% of assets)
λ	Board bargaining power
β	Discount Factor
δ	Maximum share of board seats that can change each period
i_t	Board independence in period t t negotiated in year $t-1$
p_{t+1}	Monitoring intensity set by the board in year t
$1 - e^{p_{t+1}}$	Probability set in year t that the board receives the additional signal in year $t+1$
w_t	CEO compensation in year t negotiated in year $t-1$ (% of assets)

Table 2: Descriptive Statistics
This table presents descriptive statistics for the main variables used in the estimation. The sample is based on Compustat Annual, ExecuComp, ISS Directors and Directors Legacy and Coles et al. (2014) data. The sample covers the period from 1998 to 2014 at an annual frequency. All the financial variables are reported in 2014 US dollars.

	Mean	SD	25%	50%	75%	N
Firm characteristics						
Book assets (in billions)	21.80	107.24	0.97	2.90	9.83	18248
Total firm ROA (%)	13.35	9.82	7.84	12.81	18.59	18248
Firm excess ROA (%)	1.42	9.14	-2.83	0.31	5.62	18248
Board characteristics						
Board size	9.40	2.65	8.00	9.00	11.00	18248
Fraction dependent directors (ISS)	0.28	0.16	0.14	0.25	0.38	18248
Fraction tenured independent directors	0.23	0.17	0.10	0.22	0.33	18248
Fraction dependent directors (ISS) plus tenured independent directors	0.51	0.20	0.36	0.50	0.67	18248
CEO characteristics						
CEO compensation over assets (x 1000)	2.61	3.70	0.54	1.40	3.16	18167
Forced turnover	0.02	0.15	0.00	0.00	0.00	18248
CEO spell length (years)	9.55	7.63	5.00	8.00	12.00	1534
CEO spell length if fired (years)	7.08	5.68	4.00	6.00	8.00	408
CEO spell length if left (years)	10.44	8.04	5.00	8.00	13.00	1126

Table 3: Simulated Method of Moments Baseline Estimation

This table presents the moments from the data sample and the model-simulated moments used for the SMM estimation of the model.

Moment	Data	Model
Serial correlation of firm excess profitability	0.8081	0.7131
Coefficient of firm excess profitability on dummy firing from $t-2$ to $t+2$	-0.0313	-0.0339
Coefficient of firm excess profitability on constant	0.0170	0.0196
Variance of firm excess profitability	0.0028	0.0016
Fraction fired CEOs, CEO tenure 1-3	0.0181	0.0173
Fraction fired CEOs, CEO tenure 4-6	0.0353	0.0190
Fraction fired CEOs, CEO tenure 7-9	0.0265	0.0223
Fraction fired CEOs, CEO tenure 10+	0.0146	0.0154
Fraction dependent directors, CEO tenure 1-3	0.4831	0.4828
Fraction dependent directors, CEO tenure 4-6	0.4786	0.4818
Fraction dependent directors, CEO tenure 7-9	0.4871	0.4845
Fraction dependent directors, CEO tenure 10+	0.5722	0.5767
Frequency fraction dependent directors decreases	0.4718	0.4445
Mean CEO pay scaled by assets	0.0026	0.0026
Correlation of firing rate and fraction dependent directors	-0.0244	-0.0634
Variance of first CEO pay scaled by assets (x100)	0.0009	0.0002

Table 4: Additional Statistics on Model Fit

This table presents additional moments from the data sample and model-simulated moments that are not used for the SMM estimation of the model.

Moment	Data	Model
Panel A: Additional moments about the fraction of dependent directors		
Mean fraction dependent directors	0.51	0.50
Standard deviation fraction dependent directors	0.20	0.26
25th Percentile fraction dependent directors	0.36	0.31
Median fraction dependent directors	0.50	0.48
75th Percentile fraction dependent directors	0.67	0.68
Panel B: Additional moments about the CEO's spell length		
Mean CEO spell length (years)	9.55	9.49
CEO spell length if fired (years)	7.08	6.24
CEO spell length if left (years)	10.44	10.16

Table 5: Baseline Parameter Estimates

This table presents the model parameter estimates from the SMM estimation of the model and their standard errors.

Parameter	Symbol	Estimate	Standard error
Prior mean about CEO's ability	μ_0	0.0178	(0.0017)
Prior standard deviation about CEO's ability	σ_0	0.0107	(0.0002)
Persistence of firm specific profitability	ϕ	0.1352	(0.0028)
Standard deviation of firm specific profitability	σ_ϵ	0.0192	(0.0006)
Standard deviation of the additional signal	σ_z	0.0581	(0.0011)
CEO's disutility from independence (x 1000)	γ	0.0006	(0.0001)
Board cost of monitoring (x 1000)	θ	0.0019	(0.0004)
Firm cost of a turnover	c_{firm}	0.0220	(0.0004)
Board fixed cost of firing (x 10)	$c_{board,0}$	0.0123	(0.0002)
Board variable cost of firing (x 10)	$c_{board,1}$	0.0140	(0.0003)
Board bargaining power	λ	0.0593	(0.0008)

Table 6: Removing Misalignment of Incentives

This table reports comparative statics on levels of variables of interest by varying the board cost of firing parameters and the CEO's disutility from board independence. Column (1) reports the values in the estimated model while Columns (2)-(5) report the percent changes under the specified incentive parameter relative to the benchmark model estimation. In column (2) I set $c_{board,0} = 0$, in column (3) I set $c_{board,1} = 0$, in column (4) I set $c_{board,0} = c_{board,1} = 0$, and in column (5) I set $\gamma = 0$.

	(1) Baseline	(2) No fixed cost	(3) No variable cost	(4) No cost	(5) No CEO disutility
Profitability	0.0160	0.69	-0.05	0.72	0.15
Firing rate	0.0178	41.07	-5.94	39.95	1.85
Dependence	0.5047	-34.74	-48.14	-61.19	-42.43
CEO's pay	0.0026	-4.94	-3.45	-6.51	-1.70
Shareholder value	0.1248	1.79	0.65	2.08	0.52
CEO's earnings	0.0247	-7.90	-3.12	-9.42	-2.07
Monitoring	0.2132	53.36	-12.59	52.76	6.09

Table 7: Changing Monitoring Incentives

This table reports comparative statics on levels of variables of interest by varying the board cost of monitoring parameter and the board bargaining power. Column (1) reports the values in the estimated model while Columns (2)-(5) report the percent changes under the specified incentive parameter relative to the benchmark model estimation. In column (2) I set $\theta \to \infty$, in column (3) I set $\theta \to \infty$ and $c_{board,1} = 0$, in column (4) I set $\theta \to \infty$ and $c_{board,0} = c_{board,1} = 0$, and in column (5) I set $\lambda = 0.5$.

	(1) Baseline	(2) No monitoring	(3) No monitoring and variable cost	(4) No monitoring and cost	(5) Equal bargaining power
Profitability	0.0160	-1.26	-1.21	-0.90	2.79
Firing rate	0.0178	-8.70	-5.79	23.31	105.46
Dependence	0.5047	14.06	98.14	98.14	-55.90
CEO's pay	0.0026	-4.07	-8.02	-13.22	-36.34
Shareholder value	0.1248	-0.69	0.16	1.58	10.42
CEO's earnings	0.0247	-4.25	-8.64	-15.84	-43.94
Monitoring	0.2132	-100	-100	-100	279.87

Table 8: Policy Reforms
This table reports the model-computed response on levels of variables of interest under three alternative policies. Column (1) reports the values in the estimated model while Columns (2)-(4) report the percent changes under the policy reforms.

	(1) Baseline	(2) Minimum 1/9	(3) Minimum 3/9	(4) Minimum 5/9
Profitability	0.0160	0.12	0.34	0.29
Firing rate	0.0178	1.45	1.21	1.09
Dependence	0.5047	-5.10	-26.79	-49.14
CEO's pay	0.0026	-0.24	-1.26	-2.06
Shareholder value	0.1248	0.20	0.65	0.80
CEO's earnings	0.0247	-0.43	-1.52	-2.40
Monitoring	0.2132	0.66	2.75	8.81

Figure 1: Monitoring Policy Function and Perceived Ability

This figure depicts the average board's monitoring intensity, p', for each level of the mean of the belief distribution about the CEO's ability, μ, for a high and a low level of board independence, i. The values of the parameters used are from Table 5.

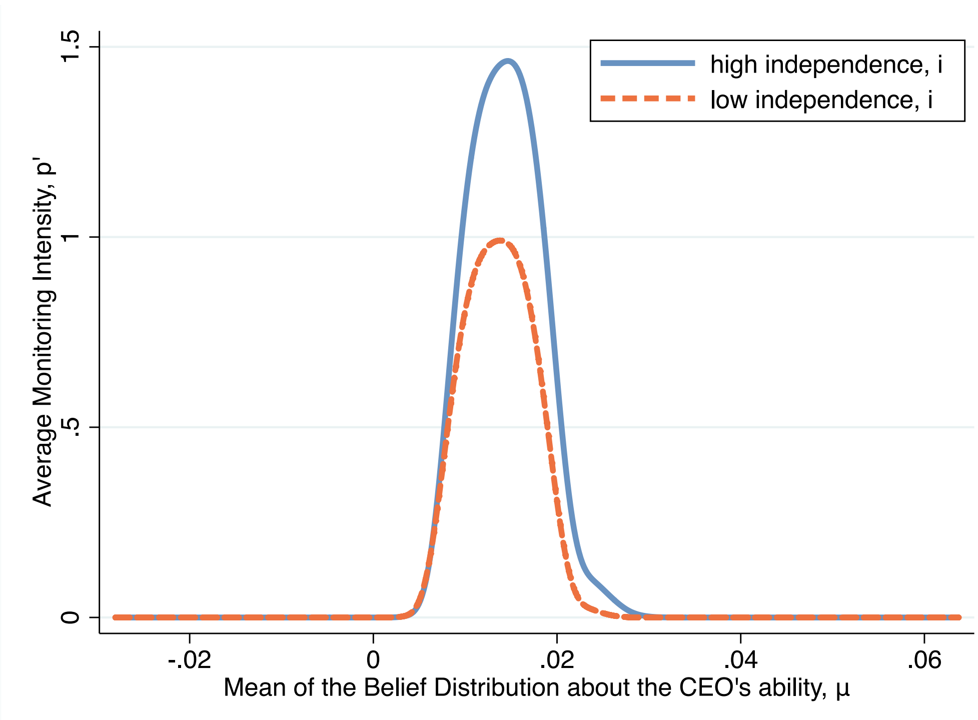


Figure 2: Average Simulated Monitoring and Initial Noise

This figure depicts the average simulated board's monitoring intensity, p', for 20 different levels of the standard deviation of the prior distribution about the CEO's ability parameter, σ_0, keeping constant the values of the other parameters.

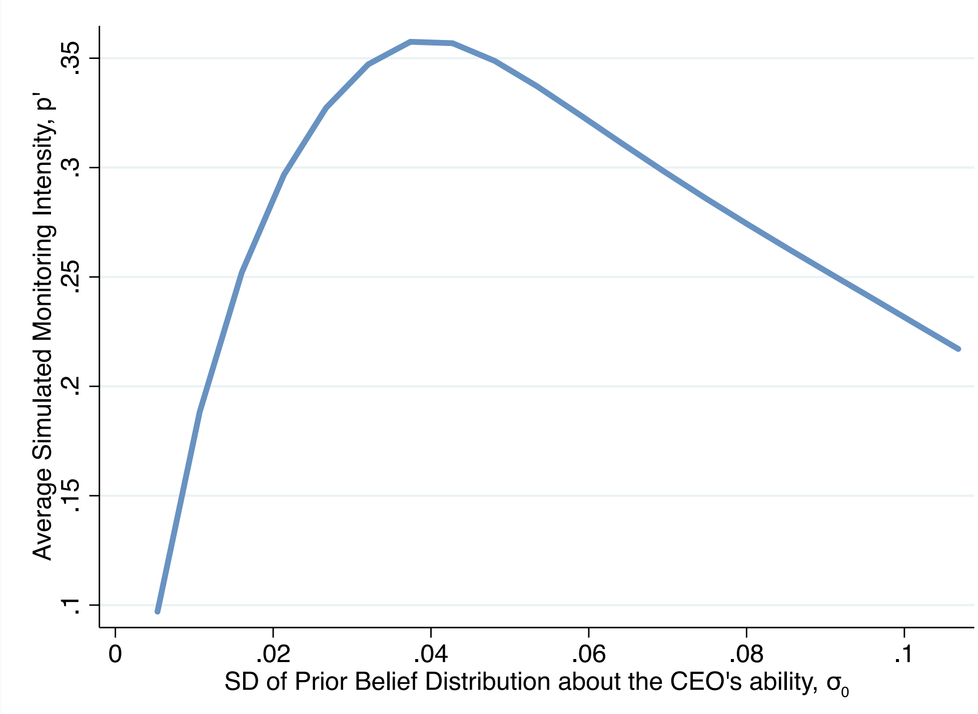


Figure 3: Average Simulated Monitoring and Board's Bargaining Power
This figure depicts the average simulated board's monitoring intensity, p', for 20 different levels of the board's bargaining power parameter, λ, keeping constant the values of the other parameters.

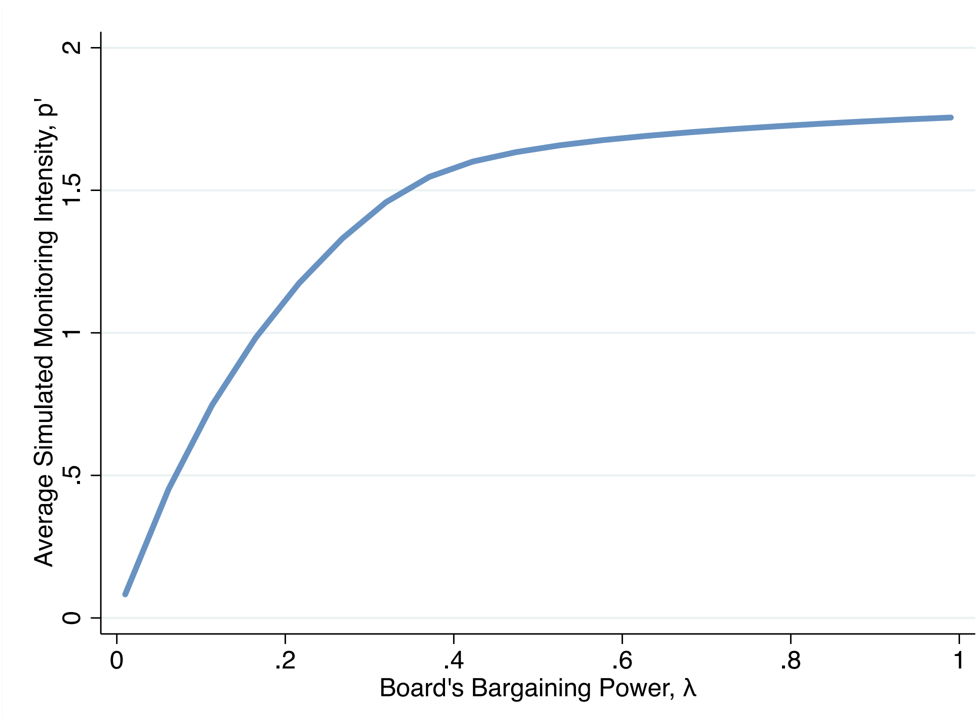


Figure 4: Firing Policy Function and Perceived Ability
This figure depicts the average board's policy function, p', for each level of the mean of the belief distribution about the CEO's ability, μ, for a high and a low level of standard deviation of the belief distribution about the CEO's ability, σ. The values of the parameters used are from Table 5.

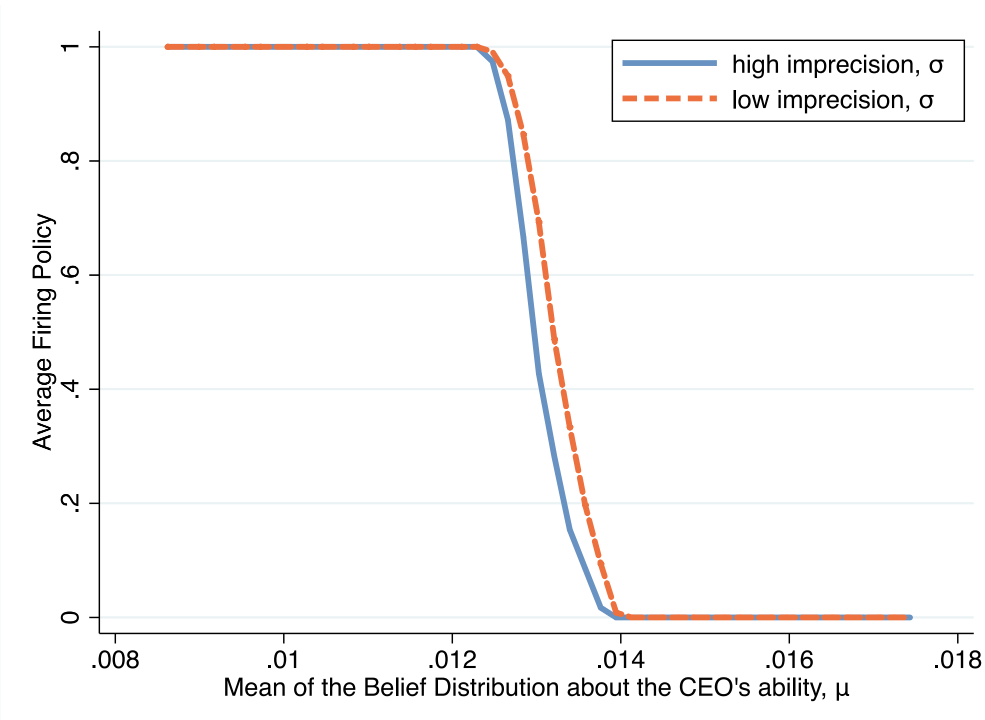


Figure 5: Firing Policy Function and Independence

This figure depicts the average board's policy function, p', for each level of the mean of the belief distribution about the CEO's ability, μ, for a high and a low level of board independence, i. The values of the parameters used are from Table 5.

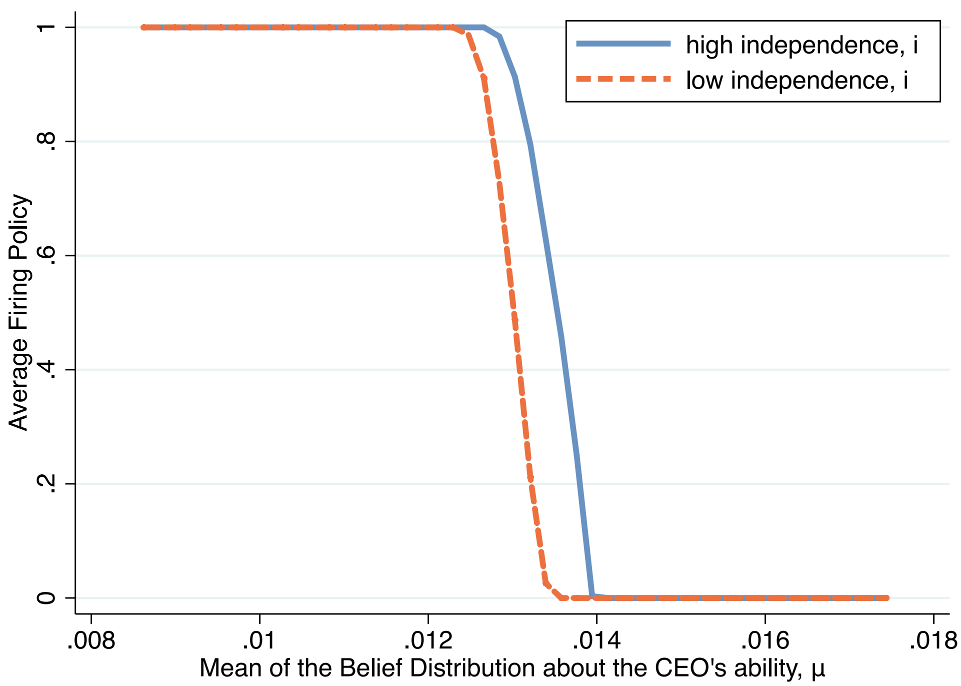
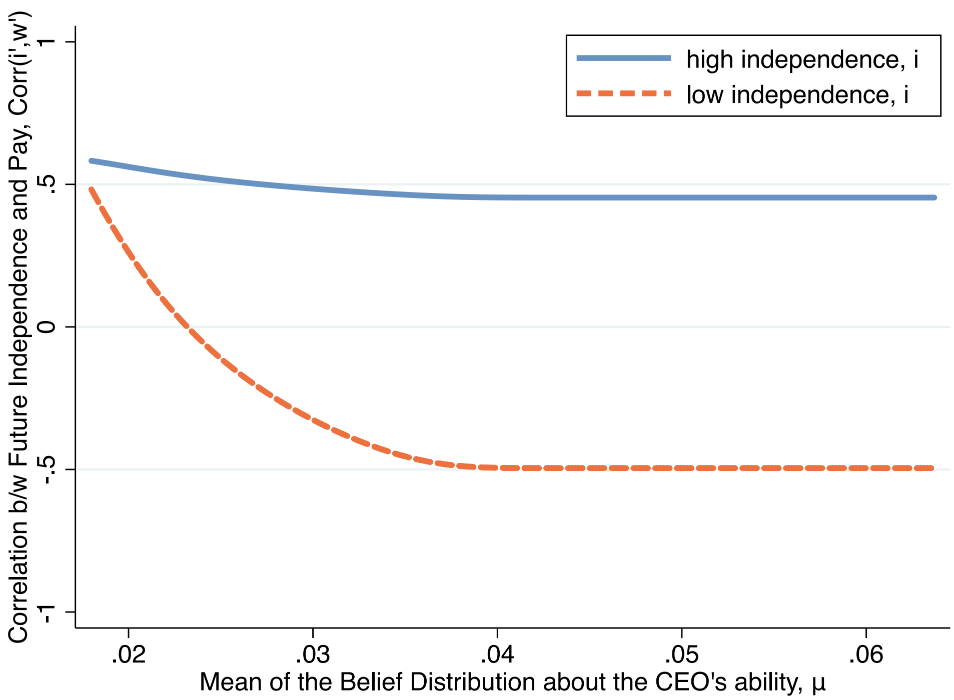


Figure 6: Correlation between CEO's Compensation and Board Independence

This figure depicts the correlation between CEO's pay, w', and board independence, i', for each level of the belief distribution about the CEO's ability,μ, for a high and a low level of board independence,i. The values of the parameters used are from Table 5.



Appendix A: Theory Appendix

A.1 Learning Problem

This subsection illustrates the learning problem, which is a Kalman filtering problem. Before the CEO starts her tenure, a skill level α is drawn from the initial prior distribution $N(\mu_0, \sigma_0^2)$ which is also the distribution of the pool of potential CEOs. The level of α is constant throughout the CEO's tenure at the firm. Before negotiating for the first time, the board receives an initial signal z_t which is normally distributed with mean equal to the drawn α and variance σ_z^2. So, the belief distribution when the CEO and the board negotiate for the first time is normally distributed with mean

$$\tilde{\mu}_0 = \mu_0 + \frac{1}{1 + \frac{\sigma_z^2}{\sigma_0^2}}(z_t - \mu_0) \tag{23}$$

and variance

$$\tilde{\sigma}_0^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma_z^2}} \tag{24}$$

Given the draw, the firm specific profitability evolves as

$$y_t = \phi\alpha + (1 - \phi)y_{t-1} + \epsilon_t \tag{25}$$

where ϕ is the persistence parameter and ϵ_t is an independent and identically and normally distributed shock with mean 0 and variance σ_ϵ^2. In each period t, the board observes the realization of the firm profitability, y_t, and with probability $1 - exp(-p_t)$ acquires an additional signal z_t which is distributed like in the first period as $\mathcal{N}(\alpha, \sigma_z^2)$. So, the distribution of the additional signal is constant over time. Before the period begins, that is before the time t profitability and extra signals are received, the board has a prior belief distribution $\mathcal{N}(\mu_{t-1}, \sigma_{t-1}^2)$. Define the time t surprise in the persistence adjusted profitability as

$$\delta_{y,t} \equiv \alpha + \frac{\epsilon_t}{\phi} - \mu_{t-1} = \frac{1}{\phi}y_t - \frac{1-\phi}{\phi}y_{t-1} - \mu_{t-1} \tag{26}$$

since y_t evolves as $y_t = \phi\alpha + (1-\phi)y_{t-1} + \epsilon_t$. Define also the time t surprise in the additional signal as

$$\delta_{z,t} \equiv z_t - \mu_{t-1} \tag{27}$$

Then, if the board acquires the additional signal, the variance of the time t posterior distribution equals

$$\sigma^2_{s,t} = \frac{1}{\frac{1}{\sigma^2_{t-1}} + \frac{\phi^2}{\sigma^2_\epsilon} + \frac{1}{\sigma^2_z}} \tag{28}$$

while the mean of the posterior distribution is equal to

$$\mu_{s,t} = \mu_{t-1} + \frac{1}{1 + \frac{\sigma^2_\epsilon}{\sigma^2_{t-1}\phi^2} + \frac{\sigma^2_\epsilon}{\sigma^2_z\phi^2}}\delta_{y,t} + \frac{1}{1 + \frac{\sigma^2_z}{\sigma^2_{t-1}} + \frac{\sigma^2_z\phi^2}{\sigma^2_\epsilon}}\delta_{z,t} \tag{29}$$

If the directors don't acquire the additional signal, the variance of the posterior distribution equals

$$\sigma^2_{n,t} = \frac{1}{\frac{1}{\sigma^2_{t-1}} + \frac{\phi^2}{\sigma^2_\epsilon}} \tag{30}$$

while the mean of the posterior distribution equals

$$\mu_{n,t} = \mu_{t-1} + \frac{1}{1 + \frac{\sigma^2_\epsilon}{\sigma^2_{t-1}\phi^2}}\delta_{y,t} \tag{31}$$

A.2 The Optimization Problem of the Board

This subsection derives the optimization problem of the board. Define f_t as an indicator function that takes value one if the CEO is fired and l_t as an indicator function that equals one if the CEO leaves the firm. Further, define the total cost of firing the CEO as $c_{fire}g(i_t)$ where $g(\cdot)$ is a decreasing and convex function in i_t and the total cost in the case the CEO leaves as c_{leave}. The optimization problem of the board can be written as follows

$$V_t = \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t\left[\sum_{s=0}^{\infty}\beta^s v_{t+s}\right] \tag{32}$$

$$= \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t\left[\sum_{s=0}^{\infty}\beta^s\Big(Y_{t+s}B - w_{t+s}B - \theta p_{t+1+s}(1-i_{t+s})B - f_{t+s}c_{fire}g(i_{t+s})B - l_{t+s}c_{leave}B\Big)\right] \tag{33}$$

$$= \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t\left[\sum_{s=0}^{\infty}\beta^s B\Big(y_{t+s} + x_{t+s} - w_{t+s} - \theta p_{t+1+s}(1-i_{t+s}) - f_{t+s}c_{fire}g(i_{t+s}) - l_{t+s}c_{leave}\Big)\right] \tag{34}$$

where in the second line I use the definition of the per-period utility of the board v_t and of D_t and in the third line the definition of Y_t. The last expression can be rearranged as

$$\frac{V_t}{B} = \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t\left[\sum_{s=0}^{\infty}\beta^s\Big(y_{t+s} + x_{t+s} - w_{t+s} - \theta p_{t+1+s}(1-i_{t+s}) - f_{t+s}c_{fire}g(i_{t+s}) - l_{t+s}c_{leave}\Big)\right] \tag{35}$$

$$= E_t\left[\sum_{s=0}^{\infty}\beta^s x_{t+s}\right]$$

$$+ \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t\left[\sum_{s=0}^{\infty}\beta^s\Big(y_{t+s} - w_{t+s} - \theta p_{t+1+s}(1-i_{t+s}) - f_{t+s}c_{fire}g(i_{t+s}) - l_{t+s}c_{leave}\Big)\right] \tag{36}$$

Using the definition $\hat{\delta}_{y,t} \equiv \alpha + \phi^{-1}\epsilon_t - \mu_t$, it is possible to rewrite the firm-specific profitability at time $t+s$ as a function of y_{t-1}, shocks and posterior means since

$$y_t = \phi\alpha + (1-\phi)y_{t-1} + \epsilon_t = (1-\phi)y_{t-1} + \phi(\hat{\delta}_t + \mu_t) \Rightarrow \tag{37}$$

$$y_{t+1} = \phi\alpha + (1-\phi)y_t + \epsilon_{t+1} = (1-\phi)((1-\phi)y_{t-1} + \phi(\hat{\delta}_t + \mu_t)) + \phi(\hat{\delta}_{t+1} + \mu_{t+1}) \Rightarrow \tag{38}$$

$$... \Rightarrow y_{t+s} = (1-\phi)^{s+1} y_{t-1} + \phi \sum_{\tau=0}^{s} (1-\phi)^{s-\tau} \hat{\delta}_{t+\tau} + \phi \sum_{\tau=0}^{s} (1-\phi)^{s-\tau} \mu_{t+\tau} \Rightarrow \tag{39}$$

$$E_t \left[\sum_{s=0}^{\infty} \beta^s y_{t+s} \right] = \frac{(1-\phi) y_{t-1}}{1-\beta(1-\phi)} + \phi \sum_{s=0}^{\infty} \sum_{\tau=0}^{s} \beta^s (1-\phi)^{s-\tau} E_t[\mu_{t+\tau}] \tag{40}$$

$$= \frac{(1-\phi) y_{t-1}}{1-\beta(1-\phi)} + \frac{\phi}{1-\beta(1-\phi)} \sum_{s=0}^{\infty} \beta^s E_t[\mu_{t+\tau}] \tag{41}$$

$$= \frac{(1-\phi) y_{t-1}}{1-\beta(1-\phi)} + \frac{\phi}{1-\beta(1-\phi)} \frac{1}{1-\beta} \mu_0 + \frac{\phi}{1-\beta(1-\phi)} \sum_{s=0}^{\infty} \beta^s E_t[\eta_{t+\tau}] \tag{42}$$

where in the second line I use the fact that $E_t[\hat{\delta}_{t+\tau}] = 0, \ \forall \tau$ and in the last line the relation $\mu_t = \mu_0 + \eta_t$. Plugging this expression into the one of $V_t/(B)$, I get

$$\frac{V_t}{B} = E_t \left[\sum_{s=0}^{\infty} \beta^s x_{t+s} \right] + \frac{(1-\phi) y_{t-1}}{1-\beta(1-\phi)} + \frac{\phi}{1-\beta(1-\phi)} \frac{1}{1-\beta} \mu_0$$

$$+ \max_{\{p_{t+1+s}\}_{s=0}^{+\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s \left(\frac{\phi \eta_{t+s}}{1-\beta(1-\phi)} - w_{t+s} - \theta p_{t+1+s}(1-i_{t+s}) - f_{t+s} c_{fire} g(i_{t+s}) - l_{t+s} c_{leave} \right) \right] \tag{43}$$

The last term on the right-hand side of the last expression is the value function of interest. The value function of the board can be written as

$$V(\eta_t, \sigma_t, i_t, \tau_t, w_t) = \max_{\{f_t, p_{t+1}\}} \Bigg\{ \frac{\phi \eta_t}{1-\beta(1-\phi)} - \theta p_{t+1+s}(1-i_{t+s}) - f_t c_{fire} g(i_{t+s}) - l_t c_{leave}$$

$$- w_t + \beta E_t[V(\eta_{t+1}, \sigma_{t+1}, i_{t+1}, \tau_{t+1}, w_{t+1})] \Bigg\} \tag{44}$$

where w_{t+1} and i_{t+1} are function of the states at t and $t-1$, respectively. Given the timing of the model, the value function of the board after the current CEO has been fired but before appointing a new CEO is

$$V_{leave}(\eta_t, \sigma_t, i_t, \tau_t, w_t) = E[V(\eta_{s,0}, \sigma_{s,0}, i_t, 0, w_t)] - c_{firm} \tag{45}$$

Given the assumption about the minimum offered wage, it can be rewritten

$$V_{leave}(\eta_t, \sigma_t, i_t, \tau_t, w_t) = -c_{firm}. \tag{46}$$

Similarly, the value function of the board if the directors have fired the CEO but before they have hired a new manager is,

$$V_{fire}(\eta_t, \sigma_t, i_t, \tau_t, w_t) = E[V(\eta_{s,0}, \sigma_{s,0}, i_t, 0, w_t)] - c_{firm} - c_{board,0} - c_{board,1}g(i_t) \tag{47}$$

that, given the above mentioned assumption about the minimum wage, gets

$$V_{fire}(\eta_t, \sigma_t, i_t, \tau_t, w_t) = -c_{firm} - c_{board,0} - c_{board,1}g(i_t). \tag{48}$$

If the CEO stays at the firm, (using the minimum wage assumption) the board value function is

$$V_{stay}(\eta_t, \sigma_t, i_t, \tau_t, w_t) = \max_{p_{t+1}} \Bigg\{ \frac{\phi\eta_t}{1-\beta(1-\phi)} - \theta(1-i_t)p_{t+1} - w_t - \beta q(\tau_t)c_{firm}$$

$$+\beta(1-q(\tau_t))\Bigg((1-e^{-p_{t+1}})E[V(\eta_{s,t+1}, \sigma_{s,t+1}, i_{t+1}, \tau_{t+1}, w_{t+1})|\eta_t]$$

$$+ e^{-p_{t+1}}E[V(\eta_{n,t+1}, \sigma_{n,t+1}, i_{t+1}, \tau_{t+1}, w_{t+1})|\eta_t]\Bigg)\Bigg\} \tag{49}$$

A.3 The CEO Utility

Before deriving the expression of the total surplus, let's derive the value function of the CEO as follows

$$U_t = E_t\left[\sum_{s=0}^{\bar{\tau}} \beta^s u_{t+s}\right] \tag{50}$$

$$= E_t\left[\sum_{s=0}^{\bar{\tau}} \beta^s (w_{t+s}B - \gamma i_{t+s}^2 B)\right] \tag{51}$$

$$\frac{U_t}{B} = E_t\left[\sum_{s=0}^{\bar{\tau}} \beta^s (w_{t+s}B - \gamma i_{t+s}^2)\right] \tag{52}$$

Then the manager's value function is given by

$$U(\eta_t, \sigma_t, i_t, \tau_t, w_t) = w_t - \gamma i_t^2$$

$$+\beta(1-q(\tau))\Bigg((1-e^{-p_{t+1}})E[U(\eta_{s,t+1}, \sigma_{s,t+1}, i_{t+1}, \tau_{t+1}, w_{t+1})|\eta_t]$$

$$+ e^{-p_{t+1}}E[U(\eta_{n,t+1}, \sigma_{n,t+1}, i_{t+1}, \tau_{t+1}, w_{t+1})|\eta_t]\Bigg)\Bigg\} \tag{53}$$

Appendix B: Data Appendix

B.1 Sample Construction

I download the CRSP / Compustat merged annual data file from 1996 to 2014[30]. I keep firms that are incorporated in the United States (using the Compustat item FIC). I keep only firms with reliable primary links to CRSP (that is, item LINKTYPE equal to LU or LC and item LINKPRIM equal to P or C). I drop firms with missing or non positive assets (item AT) or sales (item SALE).

I download the data to link the ISS databases with the CRSP / Compustat merged one from the Lalitha Naveen's website and I follow step-by-step the instructions on her website

[30]The reason why I use the CRSP / Compustat merged database rather than the Compustat one is simply because the database constructed in Coles et al. (2014) was created to merge the ISS Directors and the ISS Directors Legacy databases with the first rather than with the second database.

for the merge[31]. I classify a director as dependent if it is not classified as independent (that is, if the ISS item classification is different from I). To compute the tenure of the independent directors I use the ISS item DIRSINCE. I manually correct misreports of this variable either using information about that same director or using the first year a director participates to a meeting (if that's not the first year a firm appears in the sample). If an individual is not reported as a director of a firm for more than a year, I assume that the director's tenure at that firm restarted. I drop observations related to firms with independent directors whose tenure can't be computed. Because of this restriction, I consider only observations from 1998 onward (since for most of the firms before that year this information is missing). I compute the equity holding of directors multiplying the number of shares held by the directors (ISS item NUM_OF_SHARES) by the fiscal price at the end of the fiscal year (Compustat item PRCC_F).

I download the ExecuComp data file from 1998 to 2014. I manually fill in missing CEO indicators (ExecuComp item CEOANN) mainly following the instruction of the Internet Appendix of Taylor (2013). In particular, I classify an executive to be the CEO of a given firm-year observation if (i) that individual is the only executive listed in ExecuComp for that observation, and (ii) either (a) that individual is listed as the CEO in the previous and in the following year, or (b) that individual is listed as the CEO in the previous year, or (c) that individual is listed as the CEO in the following year. I drop observations with missing date in which they became a CEO (ExecuComp item BECAMECEO). I exclude observations related to CEOs for which there is a gap above a year between the reported date in which they became CEOs and the first year they are recorded as CEOs in ExecuComp. I consider the first year of an individual as a CEO as the first year in which he spent 6 months in the office. I exclude CEOs who stayed in the office less than 6 months. Finally, if two individuals are reported as the CEO for the same firm in the same year, I consider the one who has spent at least 6 months in office in that year.

I use the forced CEO turnover data provided to me by Florian Peters and Alexander

[31] See `https://sites.temple.edu/lnaveen/data/` for more information.

Wagner[32].

The variables are normalized in 2014 US dollars using the CPI data downloaded from the Federal Reserve Bank of St. Louis FRED database.

The total ROA is computed as operating income before depreciation (Compustat item OIBDP) at time t divided by the average of total assets (Compustat item AT) at time t and $t-1$. I drop observations with missing ROA or whose ROA is greater than one in absolute value. The industry ROA is computed as an equal-weighted average using the Fama and French 12 industry definition. The firm ROA is computed as total minus industry ROA. The wage rate is computed as CEO total compensation (ExecuComp item TDC1) divided by the average of total assets at time t and $t-1$. I don't drop observations for which this variable is missing. Both the total ROA and the wage rate are both winsorized at 1% and 99% yearly.

B.2 Leaving Probabilities

Table A1 presents the estimated frequencies with whom a non-forced CEO turnover occurs in my sample during the first 14 years a CEO has worked for a company.

B.3 Board Composition Over Time

Figure A1 documents how over the years the percentage of dependent directors has shrunk (mainly because of regulation) while on the other hand the percentage of independent directors with tenure above 9 years has increased all years except one.

B.4 Identification of Estimated Parameters

Figures A2 and A3 plot the moments as a function of the parameters that influence their value the most. All the other parameters are held constant at the values I find in the baseline estimation results. Figure A4 shows how the patterns of the fraction of dependent directors for three different values of σ_z keeping all the other values constant.

[32]For a complete description of this data see Peters and Wagner (2014) and Jenter and Kanaan (2015).

B.5 Details about the Numerical Solution

This subsection describes some of the details of the numerical solution of the model. I let the mean of the belief distribution, μ, have 501 grid points of support in $[\mu_0 - 2\sigma_0 - c_{firm} - c_{board,0} - c_{board,1}, \mu_0 + 2\sigma_0 + c_{firm} + c_{board,0} + c_{board,1}]$. The grid of the variance of the belief distribution, σ, depends on the CEO's tenure τ since, in each period except the first one, the board receives the earnings signal, and on the number of additional signals that the board has received up to that period. So, I build the grid of the variance taking into account the number of periods (and so the number of earnings signals) and the potential number of additional shocks. Since I set the maximum number of periods a CEO works for the firm equal to 15, this grid ends up having 120 grid points in total. According to the period and the number of additional signals received, the variance can be updated either only with the earnings signal or with the earnings and the additional signals.

Table A1: Probabilities of Non Forced CEO Turnover

This table presents the estimated probabilities that the CEO leaves the company during the first 14 years she spent working for a company. The sample is based on ExecuComp data from 1998 to 2018 at an annual frequency.

CEO Tenure	Estimated Value
1	0.0330
2	0.0313
3	0.0310
4	0.0340
5	0.0622
6	0.0645
7	0.0751
8	0.0968
9	0.0939
10	0.0949
11	0.0688
12	0.1135
13	0.1096
14	0.0933
15	0.0871

Figure A1: Percentage of Dependent and Tenured Independent Directors over Time
This figure plots the mean percentage of dependent directors and the mean percentage of independent directors with tenure above 9 years. The sample is based on the ISS Directors and ISS Legacy data from 1998 to 2014.

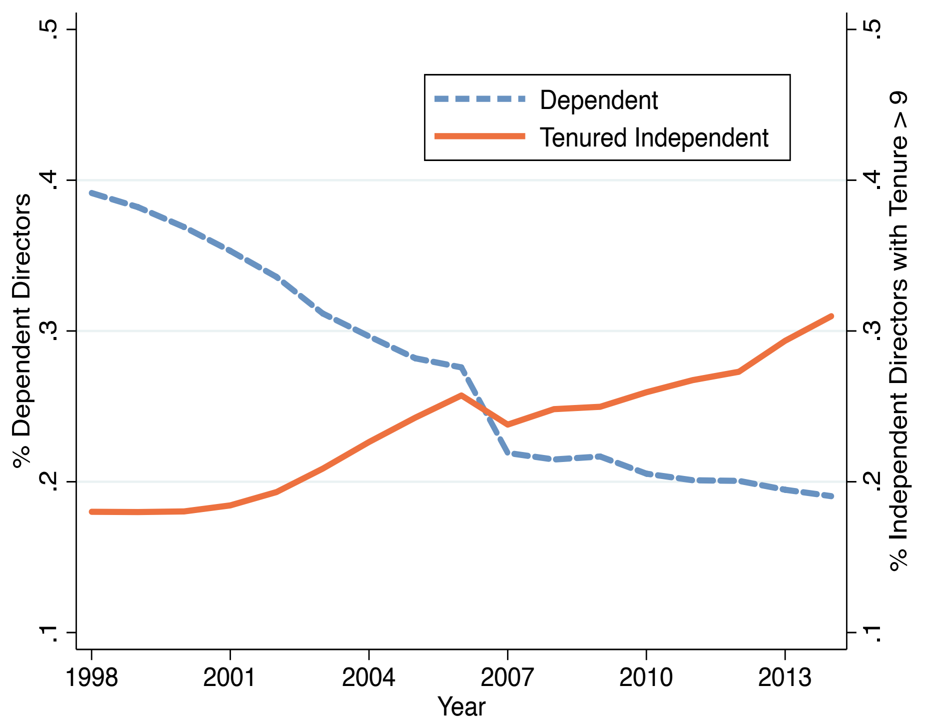


Figure A2: Elements of Identification

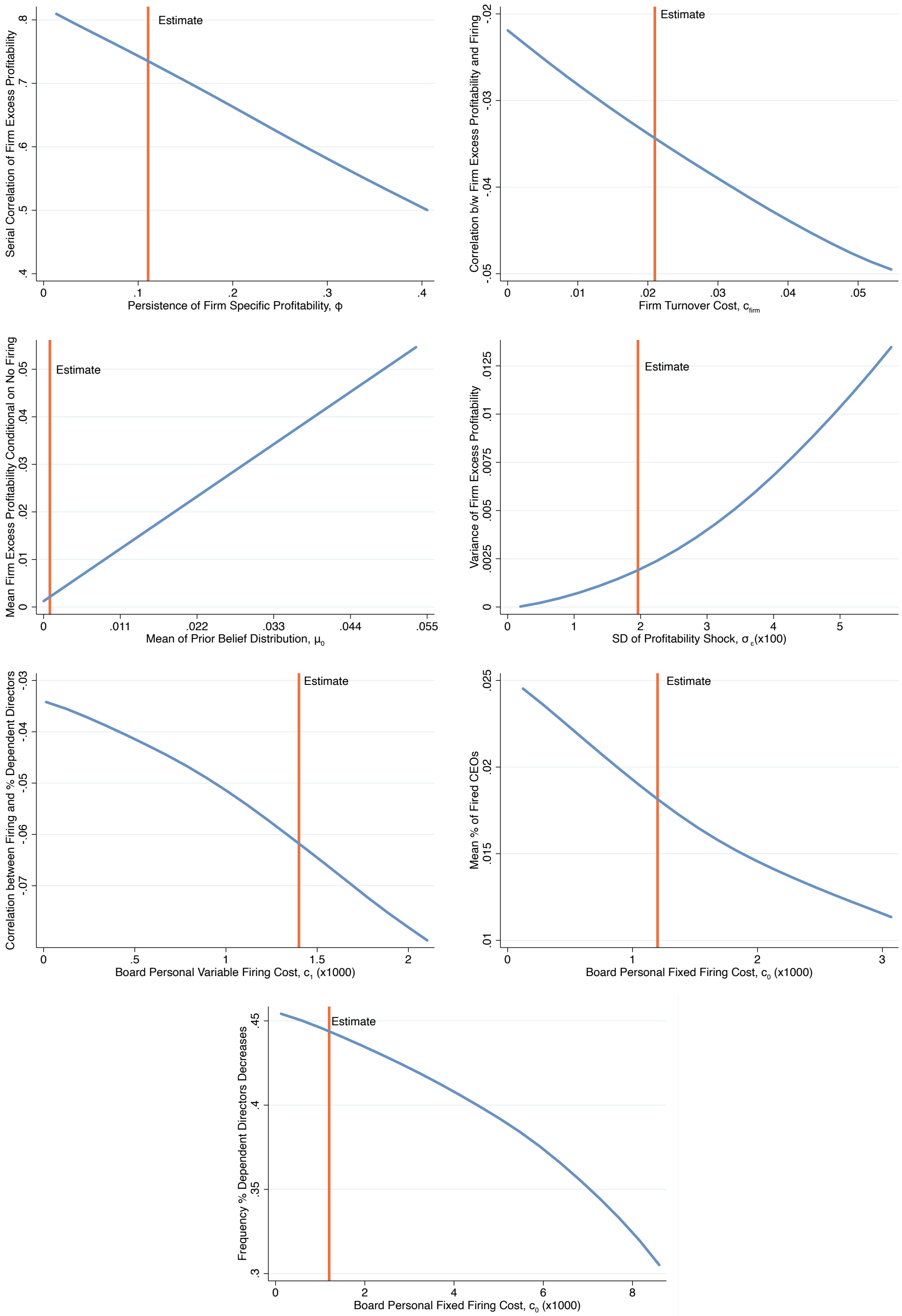


Figure A3: Elements of Identification

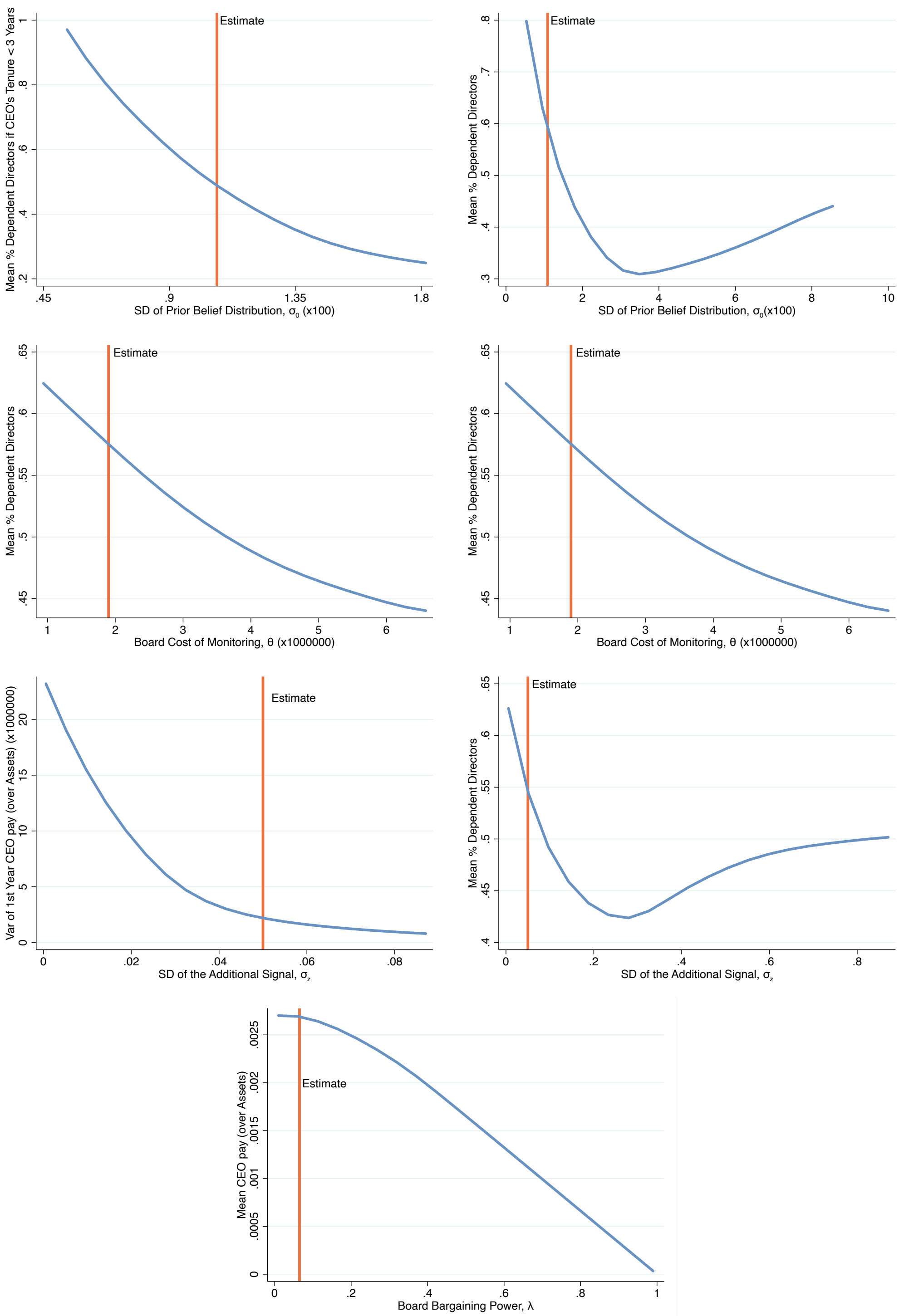


Figure A4: Identification of σ_z