working paper
1809

# How Important are Dismissals in CEO Incentives? Evidence from a Dynamic Agency Model

Alvaro Remesal

September 2018

CENTRO DE ESTUDIOS
MONETARIOS Y FINANCIEROS
Casado del Alisal 5, 28014 Madrid, Spain
www.cemfi.es

CEMFI Working Paper No. 1809
September 2018

# How Important are Dismissals in CEO Incentives? Evidence from a Dynamic Agency Model

**Abstract**

I estimate a dynamic agency model to quantify the importance of dismissals in CEO incentives vis-à-vis pecuniary compensation. The model features endogenous dynamics in deferred and ow compensation, as well as exogenous departures, and endogenous dismissals after poor firm performance. Thus, the model functions as a classification device for CEO turnover events that exploits information from all the departures in the data. I estimate the model via the Simulated Method of Moments, using data for CEOs in U.S. public firms appointed from 1993 to 2013. The estimated CEO dismissal rate is 1.2 percent, and the CEO replacement cost represents 3.4 percent of firm assets, 64 million in 2015 U.S. dollars for the median firm. Poor governance, proxied by director independence, increases the replacement costs in big firms. The relationship reverses in small firms, so board independence must also capture better hiring policies or career concerns of directors. The results confirm that CEO dismissals are infrequent. However, changes in the cost of replacements that generate small increases in the underlying dismissal rate lead to substantial reductions in the size of incentive compensation.

JEL Codes: G34, J33, J63.

Keywords: Executives, CEO turnover, CEO compensation, governance, dismissal, SMM.

Alvaro Remesal
CUNEF
alvaro.remesal@cunef.edu

**Acknowledgement**

I thank Manuel Arellano, Guillermo Caruana, Jonathan Cohn, Xavier Gabaix, Juan Pedro Gomez, Maria Gutierrez, Dirk Jenter, Simas Kucinskas, Daniel Paravisini, Javier Suarez as well as seminar audiences at the 3rd CEPR Annual Spring Symposium in Financial Economics, the 24th AEFIN Finance Forum, the 40th Symposium of the Spanish Economic Association and CEMFI for helpful comments. I especially thank Andres Almazan and David Martinez-Miera for their suggestions and for providing the data at different stages of this project. I gratefully acknowledge financial support from Banco Sabadell, the Banco Santander-CEMFI research chair, and the Spanish Ministry of Economics and Competitiveness (ECO2014-59262-P). All remaining errors are my own.

# 1 Introduction

The CEO replacement decision is one of the most important tasks of a board of directors. Incumbent CEOs must take into account that replacements represent a form of performance evaluation policy at hand for directors. Thus, CEO dismissals after poor performance may allow shareholders to provide incentives as an alternative to compensation packages. The empirical evidence suggests that CEO compensation succeeds in tying the wealth of managers to firm performance (Frydman and Jenter, 2010). In contrast, Jensen and Murphy (1990) and Murphy (1999) conclude that the turnover-performance sensitivity is economically irrelevant.[1] Moreover, estimates from news-based search algorithms suggest that CEO dismissals are infrequent: About 2 percent of CEOs in public firms are dismissed every year (Taylor, 2010, Huson et al., 2001).

What reasons underlie the coexistence of a tight relationship between firm performance and CEO compensation with a small CEO turnover-performance sensitivity? Do news-based search algorithms provide an accurate estimate of the share of CEO dismissals? This paper sheds light on these questions by estimating a structural model in which investors use both dismissals and compensation as tools of incentive provision for CEOs. Specifically, I estimate a dynamic principal-agent model adapted from Biais et al. (2007) and DeMarzo and Fishman (2007). In the theoretical framework, deferring CEO compensation allows investors to reward a good history of performance, thereby alleviating the costs of the repeated agency problem. The limited liability of the CEO restricts deferrals to be nonnegative. Hence, dismissal threats arise as an optimal incentive device after a poor history of performance. The empirical scope of dynamic agency models remains largely unexplored in the literature, and this paper represents one of the first attempts at exploiting the repeated agency framework in a structural estimation exercise.

The infrequent CEO dismissals may indicate that firms face severe frictions associated with the CEO replacement decision. I capture these frictions in the model through a replacement cost parameter that comprises the net present value of all implicit and explicit losses incurred by investors when replacing the CEO. This cost affects the investors' willingness to dismiss CEOs and, hence, influences the design of incentive compensation schemes. CEO replacement frictions may be purely technological, related to matching frictions and the firm-specific human capital contributed by CEOs. Alternatively, replacement costs can arise from CEO entrenchment or weak corporate governance that

[1] See, also, Huson et al. (2001), Denis and Denis (1995), Warner et al. (1988) and Weisbach (1988).

generates suboptimal replacement policies (Hermalin and Weisbach, 2012, Taylor, 2010, Weisbach, 1988). The structural estimation in this paper quantifies the extent of these frictions and, through estimations across subsamples of governance quality, provides evidence to disentangle between both interpretations.

Besides replacement costs, misclassification may also explain the low estimated CEO dismissal rate. News-based search algorithms yield a low CEO dismissal rate, understood as the share of forced departures. However, CEO departures may seem forced but be unrelated to poor performance. Alternatively, departures can be related to poor performance without being forced. Hence, misclassification can jeopardize the estimates of CEO dismissals and their relationship with firm performance (Fee et al., 2015, Jenter and Lewellen, 2014).

In this paper, I overcome the misclassification issues by providing a measurement of the CEO dismissal rate based on a structural model of incentives, which is absent in the literature. I make no *ex-ante* decisions regarding the nature of each departure event. Instead, all departures in the data are considered as potential dismissals — departures related to poor performance— while the structural model becomes a classification device. Specifically, I disentangle CEO dismissals from other departure reasons by jointly estimating an exogenous departure process. Thus, the model generates as a by-product an estimate of the share of CEOs that are dismissed.

I estimate the model via the Simulated Method of Moments (SMM) using information for CEOs in U.S. public firms appointed from 1993 to 2013. The identification strategy exploits the parameters' differential impact on the optimal contract and the joint variation in the empirical distribution of CEO turnover, compensation, and firm cash flows. An important step of identification is to disentangle the parameters that drive the exogenous departures from those that drive the endogenous dismissals. In order to achieve identification, I exploit that in the optimal contract investors reward good performance by increasing deferred compensation, up to an upper bound that sets a maximum distance from the dismissal decision. Therefore, changes in replacement costs or the severity of the agency problem change the slope of the CEO departure hazard function at the early years of tenure, while they change the level of the hazard function at later tenure-years.

One concern for identification is that the CEO's deferred compensation variable generated by the model is unobservable. In the optimal contract, after a good history of performance, CEOs reach the upper bound of deferred compensation, which vests in the form of flow compensation. I achieve identification by constructing a CEO flow compensation measure that captures the value obtained by the CEO from salaries, bonuses, vesting stock awards and vesting stock option awards (Taylor, 2013).

The results confirm that CEO dismissals are infrequent and costly for investors. In line with existing research, I estimate that on average 1.2 percent of CEOs are dismissed due to a poor history of firm performance (Huson et al., 2001, Taylor, 2010). The estimated CEO replacement cost accounts for around 3.4 percent of assets, equivalent to 64 million in 2015 U.S. dollars for the median firm in the sample. This estimate is around one half of that found by Taylor (2010) in an estimated learning model with pre-classified CEO departures with a news-based algorithm. The difference in the estimates can be explained by the different databases, but also by the alternative modeling and identification strategies that I exploit in this paper. The difference in the estimates also suggests that news-based classification algorithms suffer from some degree of misclassification.

I perform estimations across subsamples of the firm size and governance quality distributions to disentangle and quantify the frictions that may explain the CEO replacement costs. The estimates show that small firms face lower replacement costs and represent the most significant contributors to the CEO dismissal rate. Big firms with fewer independent directors feature a CEO replacement costs that are 4 percent higher than their counterparts with greater director independence. In contrast, small firms with more independent boards face higher replacement costs than firms with seemingly worse governance structures. These results suggest that frictions other than CEO entrenchment— e.g., the lack of suitable substitutes for the departing CEO— are the source of the CEO replacement costs in small firms.

The estimation provides a quantification of the importance of the agency problems between investors and managers. In particular, the results show that deferred compensation policies alleviate by 60 percent of the static agency costs. In counterfactual exercises, I show that incentive compensation policies are very sensitive to changes in CEO replacement frictions. Specifically, big firms with poor governance can reduce by 30 percent their CEO incentive compensation if their replacement costs match those of their counterparts with better governance. Such response contrasts with the relatively small response of the CEO dismissal rate. The results provide evidence that the threat of departure can still be an operative source of incentives without the realized CEO dismissal rate being high.

To the best of my knowledge, this paper represents the first attempt to use a dynamic principal-agent model to study CEO compensation dynamics and dismissals jointly. The structural estimation exercise provides a measurement of unobservable objects, which is unattainable with a reduced-form approach. The unobservable objects include, for instance, the reason for a CEO's departure, the degree of agency problems, the value of a CEO's expected compensation, or the CEO replacement costs.

This paper contributes to three strands of the literature. First, I estimate the CEO dismissal rate considering all departure events as a whole (Fee et al., 2015, Jenter and Lewellen, 2014). Second, I provide estimates of the size of frictions faced by shareholders when replacing their managers (Taylor, 2010). Third, I measure the importance of the agency problems between managers and investors in large corporations (Nikolov and Whited, 2014, Morellec et al., 2012, Gayle and Miller, 2009). Different from other approaches, I explicitly consider dynamic adjustments in CEO deferred compensation to avoid an otherwise underestimation of the actual extent of agency problems.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces the dynamic agency model. Section 4 presents the formulation of the optimal contract and the main implications of the model. Section 5 presents the data, the target moments and the estimation procedure. Section 6 presents the results and Section 7 concludes.

## 2 Related literature

The theoretical framework used in this paper originates in the literature on dynamic agency models, whose roots are on the seminal contribution of Abreu et al. (1990), on the analysis of repeated games with recursive techniques, and Spear and Srivastava (1987), on the use of a recursive formulation in a repeated agency framework. DeMarzo and Fishman (2007) use the repeated agency framework to study the problem of a risk-neutral entrepreneur with limited resources and the ability to undertake a long-term investment project with a known terminal date. Outside financiers can provide the necessary initial investment, but the entrepreneur can privately deviate cash flows in any period. In the optimal contract, promises of future payments as well as liquidation threats can be used to provide incentives to the entrepreneur.[2]

DeMarzo et al. (2012) study the empirical predictions of the dynamic agency theory regarding firm investment and financial structure dynamics. Financial slack— understood as cash reserves or available lines of credit— is the crucial variable determining firm investment decisions. In the managerial setting of the current paper, this slack represents the value of all unrealized and unvested compensation in the form of equity or options granted to the manager. The analysis in DeMarzo et al. (2012) is limited to provide

[2] Biais et al. (2007) study the infinite-horizon version of the same model and its convergence to the continuous-time limit. The model in its continuous-time version is also explored by He (2009) and Sannikov (2008).

Additionally, Spear and Wang (2005) propose an equilibrium model where firms can return to the labor market to select a new agent after the dissolution of a contract. The principal optimally terminates the contract either when the agent has performed poorly or when it becomes too costly to provide incentives with deferred compensation. However, the authors only provide theoretical implications.

comparative statics based on calibrations taken from other studies. In contrast, in this paper, I estimate the parameters of the model, which allows testing the ability of the model in matching the data and derive predictions on the structure of CEO's incentives. Gayle et al. (2015) estimate a moral hazard model within a matching equilibrium to jointly study the sources of executive compensation, promotion, and turnover. Their estimates suggest that firm performance is a noisier signal in larger firms. The result explains the firm size differential in executive compensation. Moreover, the authors abstract from analyzing termination threats in shaping managerial incentives.

Reduced-form methods are the most popular approach to study CEO replacements. The main findings are that CEO departures take place after poor firm performance while replacement announcements are associated with positive abnormal returns (Huson et al., 2004, 2001).[3] From a methodological point of view, most of the literature relies on news searches to classify CEO departures. Following the work of Parrino (1997) and Huson et al. (2001) researchers classify CEO turnover events as forced or unforced following some predetermined criteria. This procedure may leave room for a misclassification bias.[4]

Other papers study CEO turnover events without a previous classification. Fee et al. (2015) show that CEOs face a significant reduction in compensation and find worse jobs after their departure. The result indicates that turnover events represent adverse career outcomes and that CEO turnover is more sensitive to performance than suggested by prior research. In the same fashion, Jenter and Lewellen (2014) find that classifications of CEO turnover may underestimate the actual forced turnover figure. They suggest that the turnover rate of top performers should be the benchmark for the voluntary turnover rate. Using the departures of the five percent top-performing CEOs as a benchmark, they find that up to 50 percent of CEO replacements must be related to past firm performance.

This paper is also related to the structural estimation tradition in corporate finance started by Hennessy and Whited (2005, 2007). In this fashion, Taylor (2010) estimates a learning model to understand the reasons underlying the low rate (2 percent) of CEO dismissals. The estimation relies on an *ex-ante* classification of turnover events and yields a CEO replacement cost equivalent to 6 percent of the value of assets. A large part of the estimated cost arises from the personal cost for the board from replacing the CEO, which the author associates to CEO entrenchment. The author estimates that entrenchment reduces shareholder value by 3 percent.

[3] See also Jenter and Lewellen (2014), Kaplan and Minton (2012), Huson et al. (2001), Murphy (1999), Parrino (1997), Denis and Denis (1995), Jensen and Murphy (1990), Warner et al. (1988), Weisbach (1988), Coughlan and Schmidt (1985).

[4] For instance, in the Parrino (1997) algorithm all departures of CEOs above age 60 are classified as unforced. While some firms have a mandatory CEO retirement age, it is hard to believe that all CEO departures beyond age 60 are unrelated to performance.

In a separate paper, Taylor (2013) studies the dynamics of CEO compensation by estimating another learning model. The estimates show that CEOs enjoy downward rigidity in their compensation, but at the same time obtain 50 percent of a firm's surplus after good news. Besides, the results remain invariant when endogenous CEO turnover decisions appear as in Taylor (2010). However, CEO compensation in the model does not arise endogenously. This approach rules out the study of the endogenous feedback between CEO compensation and turnover policies.

Nikolov and Whited (2014) and Morellec et al. (2012) use structural estimation to assess the role of agency conflicts between managers and investors in shaping financial structure decisions. In both papers, the authors find important effects of agency frictions on, respectively, a firm's cash and leverage policies. However, they restrict their analysis to static contracts or reduced-form specifications. Finally, Li et al. (2016) propose and estimate a dynamic trade-off model of the capital structure. As in this paper, they use a dynamic contracting model in which collateral constraints arise endogenously. They find that taxes have little effect on the financial structure of firms, which prefer to reduce leverage to allow flexible adjustments in periods of stress.

# 3 The model

In this section, I introduce the theoretical framework, which is a version of the discrete-time dynamic principal-agent models of DeMarzo and Fishman (2007) and Biais et al. (2007). The model captures the repeated agency problem between investors and CEOs with endogenous, state-contingent replacement and compensation decisions.[5]

### Agents and preferences

Consider a firm owned by a continuum of risk-neutral infinitely-lived investors with discount factor $\beta < 1$. Investors can choose from a large pool of agents suitable to operate the firm as its CEO. Candidate CEOs are risk-neutral and discount payoffs with a lower discount factor $\beta_c < \beta$.[6] CEOs enjoy the protection of limited liability, and their outside option is zero. $t \in \{1, ..., T\}$ represent the tenure-year of the current CEO, where $T$ is an

[5]The baseline dynamic agency model needs some additional ingredients for the sake of realism, e.g., persistent cash flows shocks or exogenous CEO departures. The discrete-time formulation allows a more parsimonious representation of the additional ingredients than the continuous-time version (see, e.g., DeMarzo et al., 2012, He, 2009). The discrete-time version also allows for faster numerical computations of the optimal contract.

[6]The relatively small discount factor captures the investors' longer investment horizons, and their lower opportunity cost of funds.

exogenous limit on CEO tenure.

### Technology

The firm produces a random stream of cash flows per unit of assets, $\widetilde{x}_t$.[7] Firm cash flows have a persistent component $\widetilde{\eta}_t$ and a transitory component $\widetilde{\varepsilon}_t$, where $\widetilde{\eta}_t$ and $\widetilde{\varepsilon}_t$ are independent. Realized cash flows $x_t$ are

$$x_t = \eta_t + \varepsilon_t.$$

The persistent component $\widetilde{\eta}_t$ follows a stationary Markov chain with transition matrix $\Gamma$. The transitory component $\widetilde{\varepsilon}_t$ takes values in an equally-spaced discrete support $\{\varepsilon_1, ..., \varepsilon_n\}$ with $Pr(\widetilde{\varepsilon}_t = \varepsilon_j) = p_j$, $E(\widetilde{\varepsilon}) = 0$, $E(\widetilde{\varepsilon}^2) = \sigma_\varepsilon^2$ and $\varepsilon_j - \varepsilon_{j-1} = \Delta > 0$ for $j = 2, ..., n$. In particular, the grid for $\widetilde{\varepsilon}_t$ has bounds $\varepsilon_1 = -b\sigma_\varepsilon$ and $\varepsilon_n = b\sigma_\varepsilon$, where I set $b > 0$ to match the variance $\sigma_\varepsilon^2$.

### Agency problem

At the beginning of a CEO's tenure-year $t$ the permanent component of cash flows $\eta_t$ realizes and becomes public information. In contrast, the realization of the transitory component $\varepsilon_t$, and hence the realization of total cash flows $x_t$, is privately learned by the CEO. The CEO can conceal the true realization $\varepsilon_j$, report $\varepsilon_{j-1}$ and divert $\lambda(\varepsilon_j - \varepsilon_{j-1}) = \lambda\Delta$ per unit of assets for private consumption or saving, with $\lambda \in (0, 1]$ and $j = 2, ..., n$.[8]

$\lambda\Delta > 0$ captures weak governance or monitoring structures that allow the CEO to extract rents at the expense of investors. In a static setting, investors must provide the CEO an unpledgeable expected income of $-\lambda\varepsilon_1 = -\lambda b\sigma_\varepsilon$ for a truthful revelation of cash flows because of the CEO's limited liability (Holmstrom and Tirole, 1997). In contrast, in a dynamic setting, the agency problem can be alleviated by dismissal and deferred compensation decisions. Structural estimation provides inference on the extent to which CEO dismissal and deferred compensation policies alleviate these costs for investors.

[7] It is assumed that investors evaluate the performance of the firm, and thus of their CEO, relative to an industry benchmark, so that $\widetilde{x}_t$ must be understood in industry-adjusted terms.

[8] The per-unit-of-assets formulation implies that the degree of cash flows diversion grows linearly with firm size, as in Gayle and Miller (2009). In extensions, I perform separate estimations by subgroups of firm size. The estimations uncover a decreasing relationship between firm size and the degree of cash flows diversion per unit of assets (Edmans et al., 2009).

### CEO departures

Investors can dismiss the CEO at the beginning of each tenure-year $t$. As I explain below, dismissals arise endogenously and are regulated by an *ex-ante* contract. Additionally, I assume that, before a dismissal decision takes place as stipulated by the contract, the CEO can exogenously leave the firm with probability $h_t \in [0, 1]$. This exogenous hazard rate captures any departure event unrelated to performance, such as the CEO departing voluntarily or retiring. I assume a logistic form

$$h_t = \frac{1}{1 + \exp(\alpha_0 + \alpha_1 \ln t)} \tag{1}$$

for $t = 2, ..., T$, with $h_1 = 0$, $h_{T+1} = 1$, and $\alpha_1 < 0$. In case of departure, investors must select a new CEO, which yields an endogenous value $\underline{v}$ to investors as explained below.

CEO replacements require investors to incur in a fixed replacement cost $\kappa > 0$ per unit of assets. This parameter captures the net present value of all the explicit and implicit costs associated with the departure, including expenses such as fees paid to executive head-hunters, severance payments or pension packages (Yermack, 2006). Besides, the replacement cost can also capture, in reduced-form, two types of frictions associated with CEO departures and that discourage the use of dismissals as an incentive device. First, the replacement cost can capture the difficulty of finding suitable substitutes. Such difficulty may arise from the CEO's accumulation of firm-specific human capital (Eisfeldt and Kuhnen, 2013, Gabaix and Landier, 2008), or the CEO-specific organizational features that generate frictions from CEO replacements. Second, CEO entrenchment can work as an obstacle to replacements (Taylor, 2010, Weisbach, 1988).

### Contracts and timing

Investors and the CEO commit to the conditions of a long-term contract. The CEO observes cash flows privately. Therefore, contracts can only be contingent upon the history of cash flows reports made by the CEO. Relying on the revelation principle, DeMarzo and Fishman (2007) and Biais et al. (2007) show that there is no restriction in focusing on incentive-compatible direct revelation mechanisms.[9]

The timing of events, depicted in Figure 1, is as follows. Consider a CEO that has been in office for $t - 1$ periods, with a history of cash flows $x^h_{t-1} = \{x_1, ..., x_{t-1}\}$. At

[9]This is because both diversion and any CEO savings that can undo the incentives provided by investors are weakly inefficient, $\lambda \leq 1$ and $\beta_c < \beta$. In the direct revelation mechanism, compensation and termination policies are contingent on the actual history of cash flows and the CEO abstains from saving.

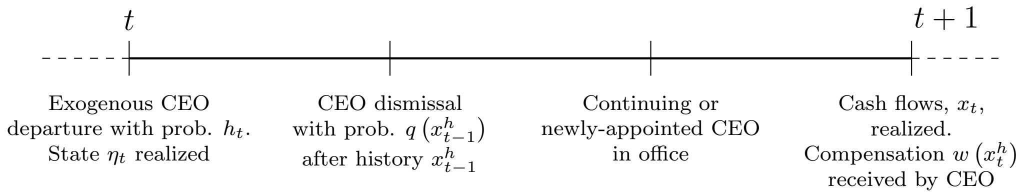


**Figure 1:** Timing of the model in a CEO's tenure-year $t$

the beginning of tenure-year $t$ the permanent cash flows $\eta_t$, and the CEO's exogenous departure become publicly known. In case of CEO retention, contracts specify, given $x_{t-1}^h$ and $\eta_t$, a dismissal decision. This decision is represented by a probability $q \in [0,1]$ with which investors dismiss the CEO. In case of exogenous departure or endogenous dismissal, investors hire a new CEO yielding a value $\underline{v}$ after incurring the cost $\kappa$. Then, transitory cash flows $\widetilde{\varepsilon}_t$ realize, and the CEO transfers them to investors in exchange for flow compensation $w$ per unit of assets. The timing repeats in period $t+1$.

Thus, contracts are represented by policies $(q, \{w_j\}_{j=1}^n)$ of CEO dismissal and compensation specified for all current realizations $\varepsilon$ and $\eta$, all possible histories of cash flows $x_t^h$ and all periods $t = 1, ..., T$, given the investors' outside option $\underline{v}$. Furthermore, I assume that the CEO's flow compensation $w$ consists of (i) an exogenous fixed salary $\underline{w} \geq 0$ and (ii) a non-negative performance-based component. The fixed salary captures, in reduced-form, rents associated with the bargaining power that the CEO may enjoy *vis-à-vis* investors.

# 4 The optimal contract and its implications

## 4.1 The optimal contract

Consider an ongoing CEO-firm relationship at the beginning of tenure-year $t$, with history of cash flows $x_{t-1}^h$. Consider also a contract specifying termination and compensation policies $(q, \{w_j\}_{j=1}^n)$ in period $t$ after $x_{t-1}^h$. Then, the expected utility of the CEO takes the following recursive expression

$$u = (1-q) \sum_{j=1}^{n} p_j \left[ w_j + (1-h_{t+1}) \beta_c u_j \right] \tag{2}$$

That is, the expected utility of the CEO is equal to the expected flow compensation in the current period, plus the expected discounted continuation utility in the following period, conditional on the CEO being retained. I denote $u$ as the value of the CEO's deferred compensation. $u$ is an endogenous object. It arises from the CEO's rational anticipation of future compensation and termination policies for each history of cash flows.

The optimal contract must be incentive-compatible, i.e. the CEO must reveal truthfully the realization of cash flows. Using the definition of $u$ we can write the incentive-compatibility constraints as follows

$$w_j + (1 - h_{t+1})\beta_c u_j \geq w_{j-1} + (1 - h_{t+1})\beta_c u_{j-1} + \lambda\Delta, \text{ for } j = 2, ..., n \tag{3}$$

Notice that incentive-compatibility requires both flow and deferred compensation to be non-decreasing in current performance, $u_j \geq u_{j-1}$ and $w_j \geq w_{j-1}$. Hence, $u$ is a sufficient statistic of the history of the CEO's performance. Specifically, the optimal contract solves a dynamic programming problem where $u$ is a state variable.[10]

Consider an ongoing relationship at the beginning of a CEO's tenure-year $t$, when the permanent component of cash flows is $\eta$ and the performance history of the CEO is summarized by $u$. Then, given the outside option $\underline{v}$, I denote firm value of the ongoing CEO-firm relationship by $V_t(u, \eta; \underline{v})$. The optimal contract solves the following problem:

$$V_t(u, \eta; \underline{v}) = \max_{\substack{q \in [0,1], \\ \{w_j, u_j\}_{j=1}^n}} q\left[\underline{v}(\eta) - \kappa\right] + (1 - q)\sum_{j=1}^n p_j \left[\eta + \varepsilon_j - w_j + \beta V_{t+1}^0(u_j, \eta; \underline{v})\right]$$

$$\text{s.t.} \quad (2),\ (3),\ (w_j, u_j) \in [\underline{w}, \infty)^n \times [0, \infty)^n$$

The first term after the maximization operator represents the expected value for investors from replacing the CEO. After termination, investors obtain an outside option $\underline{v}(\eta) - \kappa$ net of the CEO replacement costs.[11] The second term denotes the value from CEO retention, represented by the expected firm cash flows, $\eta + \varepsilon_j$, net of the CEO's flow compensation, $w_j$, plus the continuation value at $t + 1$, $V_{t+1}^0(u_j, \eta; \underline{v})$. At the beginning of tenure-year $t + 1$ permanent cash flows $\eta$ are realized and the CEO departs with probability $h_{t+1}$. Thus, $V_{t+1}^0$ takes the following expression:

$$V_{t+1}^0(u, \eta; \underline{v}) = E_{\widetilde{\eta}|\eta}\left\{h_{t+1}\left[\underline{v}(\widetilde{\eta}) - \kappa\right] + (1 - h_{t+1})V_{t+1}(u, \widetilde{\eta}; \underline{v})\right\}$$

where the superscript 0 indicates that this is the value prior to the potential exogenous departure of the CEO— while $V_{t+1}$ is the value conditional on the departure not happen-

[10]Formal arguments of this result appear in Abreu et al. (1986), Spear and Srivastava (1987), Abreu et al. (1990).

[11]I assume $\underline{v}$ to be a vector of $\eta$-specific outside options. That is, the replacement of the CEO does not affect the dynamics of firm cash flows. Notice however that the replacement cost $\kappa$ captures changes in expected cash flows that are specific to the replacement of the CEO.

ing. The maximization problem is subject to the individual rationality constraint (2), to the incentive compatibility constraints (3), to the manager's limited liability, to flow compensation being above or equal to $\underline{w}$ and the terminal condition $V_{T+1}(u, \eta; \underline{v}) = \underline{v}(\eta) - \kappa$ for all $u$ and $\eta$.

Finally, the vector of outside options $\underline{v}$ arises from investors hiring a new CEO from the pool of candidates. I assume that investors face a large pool of competing candidate CEOs. Hence, investors initialize the contract choosing at $t = 1$ a level of deferred compensation $u_1$ that maximizes their surplus.[12] Thus, the outside option $\underline{v}$ is the fixed point satisfying, for each $\eta$,

$$\underline{v}(\eta) = \max_{u \geq 0} \big\{ V_1(u, \eta; \underline{v}) \big\}.$$

Appendices A and B contain the details regarding the properties and numerical computation of the optimal contract.

## 4.2 Implications of the model

The optimal contract prescribes a mix of deferred compensation $u$, flow compensation $w$, and dismissals $q$, that minimize the cost of incentive provision. How do the optimal dismissal and compensation policies evolve with firm performance? In a dynamic setting, investors can alleviate the agency problem by providing incentives through dismissals and deferred compensation. Incentive compatibility requires upwards revisions in $u$ after good performance and downward revisions after poor performance.[13]

After surpassing a threshold performance history of cash flows the deferred compensation $u$ vests in the form of flow compensation $w$. The upper threshold in $u$ arises from the relative impatience of the CEO, which makes suboptimal to increase indefinitely the stock of deferred compensation for the CEO. On the downside, as cash flows worsen, both the values of deferred and flow compensation decrease. Once the manager's limited liability binds, pecuniary incentives become compromised. It is then when the dismissal decision appears naturally as a tool to alleviate the agency problem. The dismissal decision generates a discontinuous jump in CEO utility that contributes to incentive compatibility and represents a threat when the history of performance declines. Costly replacements make investors behave as risk-averse. Therefore, when the replacement cost increases investors

[12] The assumption of competition between candidate CEOs is made without loss of generality, since the estimation of $\underline{w}$ can capture the value of any rents related to the CEO's bargaining power.

[13] The dynamics of deferred compensation in the model are realistic, since much of the effect of performance on CEO wealth works through revaluations of stock and option holdings, rather than through changes in annual pay (Frydman and Jenter, 2010).

must rely more on deferred and flow compensation as tools to provide incentives.

Figure 2 illustrates the dynamics of the model. The figure depicts distributions from simulated spells of firm cash flows and CEO deferred and flow compensation. I compare the paths for endogenous dismissals (left panels) and exogenous departures (right panels) for CEO spells ending at the beginning of the CEO's tenure-year 11. The figure shows that the distributions of firm performance (top panels) for both types of departures are approximately similar until two periods before the departure event. The fact that performance is "flat" for exogenous departures, while it declines in the last periods before a CEO dismissal, highlights the endogenous link between the history of firm performance and CEO dismissals.

The middle and bottom panels of Figure 2 display the CEO compensation dynamics. In the initial years of CEO tenure, performance is good enough, and investors retain the CEO, who accumulates deferred compensation (middle panels). After a sufficiently long sequence of good cash flows, deferred compensation vests and investors unload flow compensation (bottom panels). The dispersion in the distribution of flow compensation arises from many CEOs failing to reach the performance threshold that allows unloading deferred compensation. When performance starts to decline (top left panel), both flow and deferred CEO compensation decrease drastically and CEOs are dismissed after two years of relatively poor performance.

Lastly, notice that flow and deferred compensation are bounded above— at the levels that entirely alleviate the moral hazard problem. Since the dismissals happen when the level of deferred compensation reaches zero, this means that CEOs located on the upper threshold of performance histories are equally likely to be dismissed in the future. This behavior is an essential characteristic of the model that allows identification.

# 5 Data, target moments and identification

## 5.1 Data

The estimation relies on yearly information from Execucomp and Compustat for the 1993-2013 period. Execucomp contains information on the compensation of the five best-paid executives of all firms in the S&P 500, S&P MidCap, and S&P SmallCap indexes. The database also includes firm-level variables. I include in the estimation sample those CEOs whose appointment takes place within the 21-year period of analysis. I do this to avoid the initial conditions problem that arises from the unavailability of CEO compensation

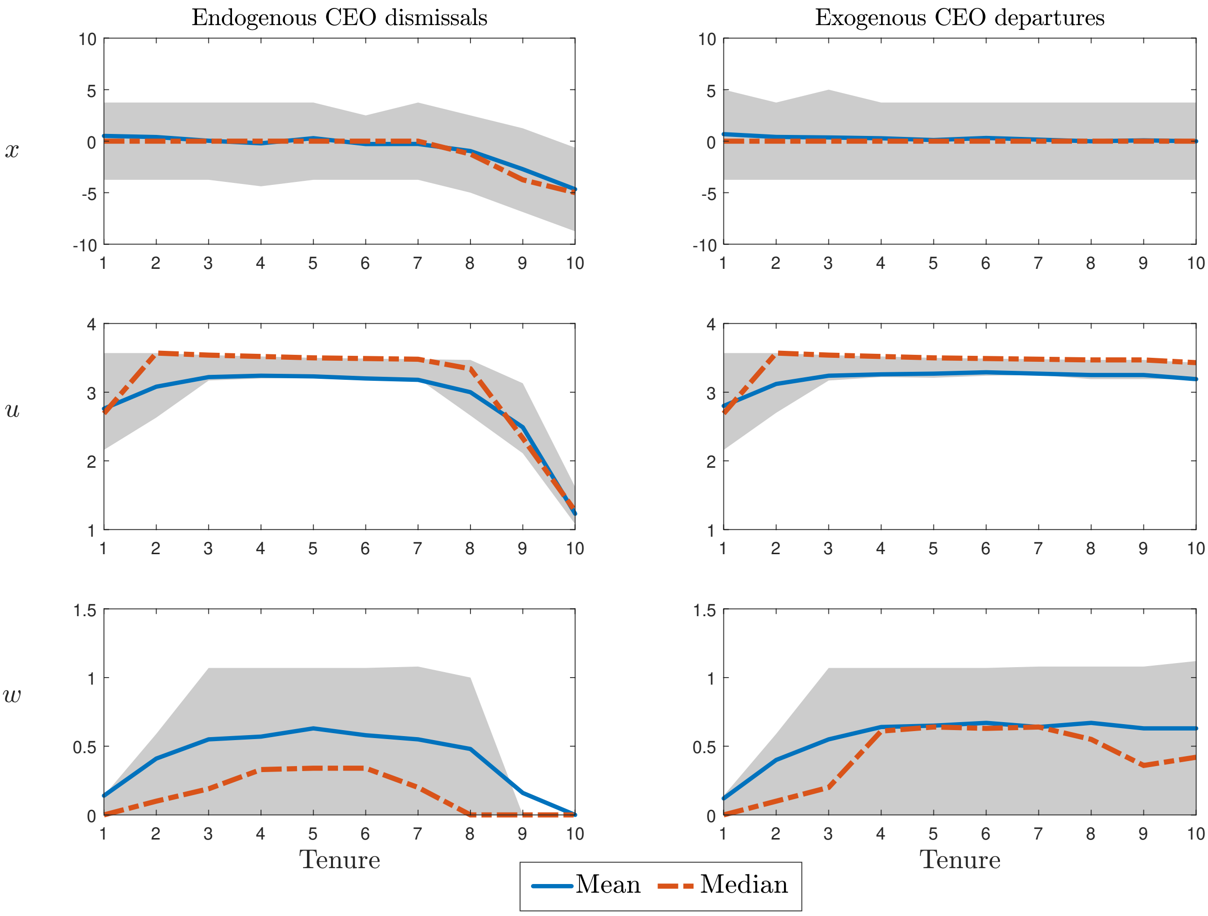


**Figure 2:** Distribution of simulated CEO spells. Firm cash flows and CEO compensation dynamics for CEOs departing at the beginning of tenure-year 11. The left panel depicts the evolution of firm cash flows (top), CEO deferred compensation (middle) and CEO flow compensation (bottom) at each tenure-year before the dismissal decision. Variables are expressed as percentage of the firm asset value. The right panel depicts the evolution of the same variables at each period for exogenous departures. Shaded areas represent the interquartile range of the variables at each tenure-year. The distributions are obtained from fully-observed CEO spells in an artificial sample of 15,000 firms simulated using the model solution across 100 years, where the last 50 are taken as the estimation period. In the simulations, the number of dismissals at tenure-year 11 is of 299, while exogenous departures amount to 1,125. Model parameters in this exercise are $\mu = 0.0$, $\rho = 0.0$, $\sigma_\eta = 0.05$, $\sigma_\varepsilon = 0.025$, $\beta = 0.96$, $\beta_c = 0.85$, $\lambda = 0.35$, $\kappa = 0.025$, $\underline{w} = 0.0$, $\alpha_0 = 4$, $\alpha_1 = -0.4$, $T = 50$.

information for earlier years.

The model generates two different forms of CEO compensation: (i) flow compensation $w$ and (ii) deferred compensation $u$. In practical terms, $u$ comprises the actual stock awards or stock option awards that have a vesting schedule. Besides, $u$ also includes the value of future unrealized compensation. Hence, no direct data counterpart can fully account for $u$. The identification strategy deals with this issue by matching other measures of compensation. In particular, I construct an empirical counterpart of flow compensation $w$. Following Taylor (2013), Edmans et al. (2009) and Coles et al. (2006), I construct a variable, $flcomp$, that is composed of yearly salaries, yearly bonuses and the value realized by the CEO from vesting stock and vesting stock option awards.[14]

The data counterpart of firm cash flows $x$ is the item "NIBEX" in Execucomp. I adjust it at the industry level by subtracting the industry median in each year, where I take industry definitions from the Fama-French 49 industry classification. Thus, the resulting variable can is an industry-adjusted ROA.[15] As in the model, I express compensation and cash flows as fractions of each firm's lagged value of assets. I trim observations at the 1 percent and 99 percent levels to avoid the influence of outliers or data mistakes on the computation of target moments. Besides, I drop any CEO spell with missing information about their compensation, tenure or firm cash flows in any of their tenure-years. I provide further details on the construction of the sample of CEOs in Appendix C.

To understand the role of CEO entrenchment in explaining the CEO replacement costs, I also use information from the ISS directors database. This database covers individual information about board members for the period 1996-2013. More specifically, it classifies directors depending on the lack of pecuniary interests of each director in the firm, i.e., their independence. Thus, I proxy the lack of CEO entrenchment by the number of independent directors in the board (Arena and Ferris, 2007, Gillette et al., 2003, Weisbach, 1988, Fama and Jensen, 1983).

Table 1 reports summary statistics. The full sample covers more than 19,000 firm-year observations from 4,098 CEOs in 2,235 firms. Panel A shows that, on average, 11.4 percent of the CEOs in the sample leave office in a given year. The CEO turnover rate decreases with firm size and with the degree of director independence. Panel B of Table 1 reports other statistics of interest. As expected, the sample includes large firms,

[14]The details on the construction of this variable are provided in Appendix C. Many of the studies on CEO compensation use the item "TDC1" as the measure of total CEO compensation (Gabaix and Landier, 2008, Gabaix et al., 2014). "TDC1" includes deferred compensation items in the form of restricted stock and option awards that are not directly available for the executive at the date of granting.

[15]In the model, cash flows $x$ are gross of CEO compensation. I add $flcomp$ to the empirical measure of cash flows for consistency. In practice, CEO compensation represents a little part of gross firm cash flows. Hence, the results would remain unchanged without considering CEO compensation.

**Table 1:** Descriptive statistics

| Panel A: Number of observations and CEO turnover events by size and board independence | | | | | | |
|---|---|---|---|---|---|---|
| | Firm-years | Turnover events | Turnover rate | CEOs | Firms | |
| All firms | 19,686 | 2,246 | 0.114 | 4,098 | 2,235 | |
| Firm size and board independence | | | | | | |
| Bottom size quartile | 4,810 | 612 | 0.127 | 1,025 | 684 | |
| <median board independence | 1,088 | 142 | 0.131 | 239 | 195 | |
| >median board independence | 629 | 72 | 0.114 | 137 | 111 | |
| 2nd size quartile | 4,975 | 572 | 0.115 | 1,024 | 753 | |
| <median board independence | 1,706 | 212 | 0.124 | 368 | 298 | |
| >median board independence | 1,098 | 98 | 0.089 | 210 | 175 | |
| 3rd size quartile | 5,109 | 548 | 0.107 | 1,025 | 749 | |
| <median board independence | 1,664 | 171 | 0.103 | 325 | 281 | |
| >median board independence | 1,285 | 124 | 0.096 | 253 | 205 | |
| Top size quartile | 4,792 | 514 | 0.107 | 1,024 | 606 | |
| <median board independence | 1,914 | 191 | 0.100 | 407 | 304 | |
| >median board independence | 992 | 95 | 0.096 | 217 | 160 | |
| Panel B: Other variables | | | | | | |
| | Firm-years | Mean | SD | Median | p5 | p95 |
| Assets (2015 US$ billion) | 19,686 | 9.814 | 35.600 | 1.879 | 0.140 | 37.562 |
| Cash flows (% assets) | 19,686 | -0.640 | 8.192 | -0.049 | -14.724 | 11.177 |
| CEO flow compensation (% assets) | 19,686 | 0.250 | 0.395 | 0.117 | 0.010 | 0.938 |
| CEO spell (years) | 2,246 | 4.887 | 3.497 | 4 | 1 | 12 |

Panel A reports the size of the estimation sample and CEO turnover statistics across several subsamples. The estimation sample consists of 4,098 CEOs in 2,235 firms in Execucomp appointed in the 1993-2013 period. Panel B reports summary statistics for firm and CEO variables. Firm cash flows are defined as items "NIBEX" in Execucomp divided by the lagged value of assets and gross of CEO flow compensation. Industry-adjusted cash flows are gross cash flows minus the industry median in the period. I use the Fama-French 49 industry classification. CEO flow compensation is computed for observations before the date of CEO departure. The quantiles of the distribution of firm size and board independence are measured at the beginning of a CEO spell. Subsamples of firm size and director independence are computed as of the beginning of a CEO spell.

with the median firm having 1.9 billion of 2015 U.S. dollars in assets. On average firms obtain negative industry-adjusted cash flows, with a dispersion of 8.2 percent. CEO flow compensation represents on average 0.25 percent of a firm's assets, with a median of 0.12 percent.

## 5.2 Target moments and estimation procedure

The model parameters are estimated using the Simulated Method of Moments (SMM). The identification of the model requires the choice of target moments that are sensitive to changes in the structural parameters. The parameters affect most of the target moments in some manner. However, I can provide intuitive arguments to understand how to identify each parameter. For brevity, I refer the reader to a more detailed identification discussion in Appendix D.

I must fix two parameters of the model *ex-ante*. First, the model is identified up to the expression $\beta - \beta_c$. Thus, I fix $\beta$ to 0.96, corresponding to a required real return for investors of roughly 4 percent per year, and estimate $\beta_c$ with the data. Second, I fix the maximum duration of the contract at tenure-year $T = 50$. This horizon is sufficiently large to let the parameters explain the variation observed in the data, making the arbitrary choice of $T$ irrelevant.

I select 15 target moments to capture the joint variation in firm cash flows, CEO departure frequencies and the patterns of CEO compensation and identify the model parameters. The model aims to capture the time-series variation in firm performance, turnover, and compensation within CEO spells and abstracts from cross-sectional heterogeneity. Thus, I filter out from the target statistics the fixed cross-sectional heterogeneity at the CEO-firm level.

Let $x_{it}$ denote the observed firm cash flows for CEO spell $i$ at tenure-year $t$. Cash flows in the model have a permanent-transitory structure. I construct the transitory component $\varepsilon_{it}$ as a discrete Gaussian distribution with zero mean and standard deviation $\sigma_\varepsilon$.[16] I characterize the permanent component $\eta_t$ with a drift term $\mu$, autoregressive parameter $\rho \in (-1, 1)$ and symmetric disturbances with standard deviation $\sigma_\eta$. I use four target moments related to firm cash flows to identify the four parameters of the cash flows process. First, the cash flows sample mean, $\widehat{E}(x_{it})$, allows to identify the parameter $\mu$. Second, I estimate the coefficient $\phi$ in the AR(1) panel model

$$x_{it} = \phi x_{it-1} + \nu_{it} \; . \tag{4}$$

I follow Han and Phillips (2010) to compute an estimate $\widehat{\phi}$ that deals with any fixed cross-sectional heterogeneity at the CEO-firm level in $\nu_{it}$. Third, I capture additional variation in cash flows by matching $\widehat{Var}(\Delta x_{it})$ and $\widehat{Cov}\,(\Delta x_{it}, \Delta x_{i,t-2})$, where I take first differences to remove the cross-sectional heterogeneity. These moments allow to identify the parameters $\rho$, $\sigma_\eta$, and $\sigma_\varepsilon$.

The next set of seven moments capture CEO departure frequencies. I do not make any *ex-ante* classification of CEO departures since the model generates its classification. First, I match the CEO turnover rate in the data. Second, I match the CEO turnover hazard rates at tenure years $t = \{1, 2, 3\text{-}4, 5\text{-}6, 7\text{-}8, 9\text{-}10\}$.

The remaining five moments convey information about the behavior of CEO flow compensation. I match the sample mean of flow compensation, $\widehat{E}(flcomp)$, and the variance of the change in flow compensation, $\widehat{Var}(\Delta flcomp_{it})$. I also match the sensitivity of CEO flow compensation to contemporaneous cash flows and CEO tenure. In particular,

[16] I build the grid for $\varepsilon_{it}$ with $n = 11$ equidistant points as I explain in Appendix B. I build the grid of transitory cash flows by setting upper and lower bounds $\pm b\sigma_\varepsilon$ so that the grid's variance matches $\sigma_\varepsilon$. Given the choice of $n = 11$ the grid bounds are given by $b \simeq 2.5$.

I match the estimated coefficients $\widehat{\delta}_1$ and $\widehat{\delta}_2$ in equation

$$flcomp_{it} = \delta_0 + \delta_1 x_{it} + \delta_2 \ln t + \zeta_{it} \tag{5}$$

where I estimate the first-differenced version of the equation to remove any fixed cross-sectional heterogeneity in $\zeta_{it}$.[17]

The combined set of moments on CEO turnover and compensation provides identification on the remaining parameters. Importantly, the size of compensation increases with replacement cost parameter $\kappa$, the moral hazard parameter $\lambda$ and the fixed salary $\underline{w}$. However, they have opposite effect on the dispersion of compensation and the frequency of CEO turnover.[18] The wage-tenure slope parameter $\delta_2$ provides identification of the CEO's discount factor $\beta_c$.

Lastly, I identify the parameters $\alpha_0$ and $\alpha_1$ in the exogenous departure process (1) through their impact on the dispersion of compensation and on the CEO turnover hazard rates. Importantly, while $\alpha_0$ and $\alpha_1$ affect, respectively the level and slope of the CEO departure hazard function, the parameters $\kappa$ and $\lambda$ only have an impact on the slope of the function at the early years of CEO tenure and on the level of the function at later years. This effects are due to the upper bound in the level of deferred compensation $u$ — or, equivalently, an upper bound in the history of performance— at which CEOs face the same probability of dismissal in the future. For further details on the identification of the model I refere to Appendix D.

The vector of ten parameters to estimate is given by

$$\theta = (\mu, \rho, \sigma_\varepsilon, \sigma_\eta, \beta_c, \lambda, \underline{w}, \kappa, \alpha_0, \alpha_1).$$

I develope the estimation as follows. First, given a vector $\theta$, I numerically compute the optimal contract. Using the numerical solution, I simulate a panel of as many firms as in the real sample over 100 years, and I pick the last 21 years as the counterpart of the sample period (1993-2013). More specifically, I keep those CEOs whose date of appointment takes place within the estimation period to match the sampling criteria in the data. Then, I compute the 15-moment counterparts from the simulated model.

[17] When computing moments related to CEO compensation I exclude observations in the departure year. Many CEOs receive severance and retirement packages that inflate their compensation in the last period in office (Edmans and Gabaix, 2015, Yermack, 2006). Such packages can be captured through the replacement cost $\kappa$. Moreover, I estimate equation (5) for CEOs beyond their tenure-year 3, which represents the usual vesting period of deferred compensation (Gopalan et al., 2014). I filter aggregate effects in compensation by removing yearly averages from the CEO turnover indicator and compensation measures, while I estimate equation (5) including year fixed effects.

[18] This is consistent with Gillan and Nguyen (2016) who find that firms that can forfeit compensation after a CEO departure tend to show lower pay-for-performance.

Hence, given the vector $\widehat{m}$ of 15 empirical moments, the objective of SMM is to choose a vector of model parameters $\theta$ that generates simulated moments $m(\theta)$ that are close to $\widehat{m}$. $m(\theta)$ corresponds to the average of the simulated moments over $S$ artificial panels.[19] I choose $S = 20$ as a conservative criterion. Then, the parameter estimates are given by the vector $\widehat{\theta}$ that solves:

$$\widehat{\theta} = \arg\min_{\theta} \left(\widehat{m} - m(\theta)\right)' W \left(\widehat{m} - m(\theta)\right) \ . \quad (6)$$

I construct the weight matrix $W$ and compute robust standard errors as in Li et al. (2016), Nikolov and Whited (2014), following the influence function approach of Erickson and Whited (2002). I provide further details in Appendix E.

# 6 Estimation results

## 6.1 Baseline results

Table 2 reports the parameter estimates for the full sample. The estimated CEO replacement cost is 3.42 percent of the value of firm assets, with a standard error of 0.95 percent. This estimated cost represents 64 million in 2015 U.S. dollars for the median firm in the sample and 336 million for the average firm. An alternative way to understand the magnitude of these costs is to compare them with the annual salary of CEOs. Specifically, the estimated cost of CEO replacements represents around 30 times the annual CEO pay in the median firm and around 80 times the annual CEO pay in the average firm.[20] While large, the estimate of the CEO replacement costs is smaller than the estimates in Taylor (2010), who finds that CEO replacement costs are around 6 percent of firm's assets. Nevertheless, the confidence bands of the estimates are large, the 95 percent confidence interval of $\kappa$ ranges from 1.4 to 5.4.

The disagreement with the estimates in Taylor (2010) may arise from three alternative explanations associated with the identification strategy, on top of the different theoretical framework. First, in this estimation exercise, I use a different database. Taylor (2010) exploits information from CEOs in the Forbes annual compensation surveys who left of-

[19]For larger samples, the model reaches a stationary distribution quite fast, so statistics are stable after few simulations (Biais et al., 2013).

[20]I consider the widely used variable "tdc1" in Execucomp as a measure of yearly pay. From another perspective, Yermack (2006) estimates that the mean separation package, i.e., severance pay and other related compensation, across Fortune 500 CEOs in 1996-2002 is worth 5.4 million of U.S. dollars. In 2015 terms, the explicit costs of departures represent around 7.5 million of U.S. dollars, 2.2 percent of the estimated costs of CEO departures.

**Table 2:** Parameters and estimation results: Full sample

| Parameter | Description | Estimated value | Standard error |
|---|---|---|---|
| $T$ | Maximum contract duration (years) | 50 | - |
| $\beta$ | Investors discount factor | 0.96 | - |
| $\beta_c$ | CEO discount factor | 0.800 | 0.009 |
| $\mu$ | Mean of cash flows | -0.635 | 0.042 |
| $\rho$ | Autocorrelation of cash flows | 0.534 | 0.035 |
| $\sigma_\eta$ | Std. deviation of permanent cash flows | 5.562 | 0.081 |
| $\sigma_\varepsilon$ | Std. deviation of transitory cash flows | 1.067 | 0.047 |
| $\lambda$ | Moral hazard parameter | 24.036 | 1.191 |
| $\underline{w}$ | Fixed CEO compensation | 0.000 | 0.009 |
| $\kappa$ | CEO replacement cost | 3.417 | 0.959 |
| $\alpha_0$ | Exogenous departure: Intercept | 2.543 | 0.089 |
| $\alpha_1$ | Exogenous departure: Tenure slope | -0.244 | 0.039 |
| | CEO turnover rate (%) | 11.132 | 0.027 |
| | CEO dismissal rate (%) | 1.185 | 0.004 |

The Table shows the parameter estimates and robust standard errors for the full sample. The estimation sample consists of 4,098 CEOs in 2,235 firms in Execucomp appointed in the 1993-2013 period. The bottom part of the table displays the estimated probabilities of CEO turnover and CEO dismissal with their corresponding standard errors, computed from 10,000 Monte Carlo simulations of the model using panels of 2,235 firms and parameter values in the table.

fice from 1971 to 2006.[21] Second, I exploit information from all CEO turnover events, regardless of any previous classification of CEO departures. News-based search algorithms may identify as forced departures only those with more intense media coverage, which may reflect a more complex succession process or the incumbent CEO displaying more resistance to the replacement. In turn, this Type-II errors— i.e., dismissals being classified as exogenous departures— may yield an upward bias in the estimated CEO replacement costs. Lastly, Taylor (2010) identifies $\kappa$ by matching the cash flows behavior surrounding the departure period. The identification strategy in this paper relies on the negative comovement between CEO compensation and CEO turnover due to changes in CEO replacement costs. Moreover, the estimate of $\kappa$ captures the net present value of all expected costs associated with a CEO replacement and is independent of any timing assumptions.[22]

The bottom of Table 2 reports that the estimated model yields a relatively low frequency of CEO dismissals: 1.2 percent of CEOs are dismissed on average in every period. This result is consistent with studies based on news-search algorithms, which find CEO forced departure rates in the order of 1-2.5 percent per year (Parrino, 1997, Taylor, 2010) for different time periods. Thus, the results go against more recent research that suggests that classification algorithms underestimate the number of forced CEO departures (Fee

[21] Using the Forbes dataset for the exercise in this paper is unfeasible since disaggregated data on CEO compensation is scarce in the period before 1992.

[22] The size of the estimated CEO replacement cost here is comparable with the estimates of Nickerson (2013) in the context of matching frictions. However, the selection of turnover events also follows the Parrino (1997) criterion.

et al., 2015, Jenter and Lewellen, 2014, Kaplan and Minton, 2012).

As stated above, the estimated value of $\kappa$ suggests that classification algorithms may yield some Type-II errors. However, the estimated CEO dismissal rate is of the same magnitude than the usual estimates from news-based classification algorithms. Hence, the estimates suggest that classification algorithms must also suffer from a Type-I error— i.e., algorithms classify exogenous dismissals as forced departures. The intense media coverage of some CEO departures may be related to the higher costs of succession rather than poor firm performance.[23]

Figure 3 depicts the cumulative share of CEOs that depart before or in each tenure-year $t$. The pace at which CEOs depart from their firms is decreasing with their tenure, with the majority of CEO departures taking place in the first six years of tenure. The estimated parameters $\alpha_0$ and $\alpha_1$ imply that, in the absence of endogenous dismissals, 40 percent of CEOs would depart after five years in office. Figure 3 also shows the corresponding share of CEO dismissals. By tenure-year 20 the model predicts that around 16 percent of all CEOs are dismissed. CEO dismissals are more frequent at the initial stages: 40 percent of all dismissals concentrate in the first five years of CEO tenure.

Exogenous departures may obscure— both in reality and in the model— the relevance of CEO dismissals. That is, exogenous departure shocks can happen during or just before a sequence of poor firm performance. In order to understand the importance of this feature, Figure 3 also depicts the counterfactual cumulative share of CEO departures that would arise from the estimated model in the absence of exogenous CEO dismissals. When exogenous dismissals are not present, the estimated model yields a cumulative CEO dismissal rate of 33 percent at tenure-year 21, which doubles the estimates including exogenous dismissals. The result points out that the CEO dismissal rate without exogenous departures is in the order of 2 percent. This exercise highlights that exogenous CEO departures prevent CEO dismissals from appearing more frequently, but the latter would still be infrequent.

The remaining parameter estimates provide a quantification of the agency problem between investors and CEOs and how deferred compensation and dismissal policies alleviate such problem. The estimate of $\lambda$, which measures the off-equilibrium rate of consumption of concealed cash flows, is 24 percent. The estimated standard deviation of idiosyncratic cash flows $\sigma_\varepsilon$ is 1.1 percent of firm assets. In a static agency framework, these parameters imply that agency costs represent 0.6 percent per year.[24] However, incentive CEO flow

[23]Besides, the news-based classification algorithms list as unforced those departures where the CEO is more than 60 years old, regardless of the media coverage.

[24]As explained in the model section, $-\lambda\varepsilon_1 = -\lambda b\sigma_\varepsilon$ captures the expected unpleadgeable income of the CEO in the static version of the model.

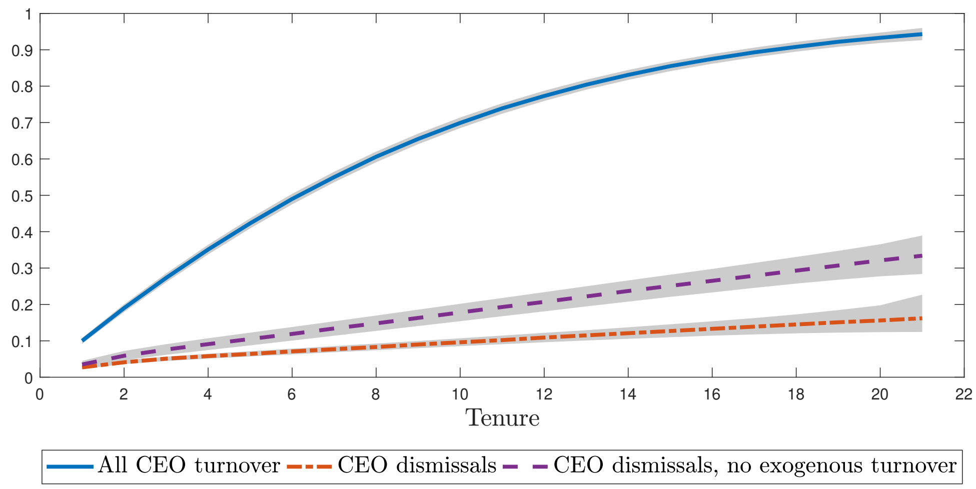


**Figure 3:** Predicted patterns of CEO departures. The Figure shows the cumulative percentage of CEOs departing at different tenure levels for all turnover and dismissals separately. The predicted probabilities come from averaging across 10,000 Monte Carlo simulations using the estimates in Table 2, each of which contains 2,235 firms. The shaded areas correspond to the 95 percent confidence region from the 10,000 simulated samples.

compensation represents on average 0.25 percent of firm assets, 40 percent of the static costs.[25] While the per-period average compensation costs are in line with those estimated by Morellec et al. (2012), the dynamic features of the optimal contract— deferred compensation and dismissals— reduce by 60 percent the static agency costs. Lastly, the estimation yields a CEO discount factor $\beta_c$ of 0.82. This value highlights the CEO's need for early compensation, which reduces the effectiveness of deferrals in alleviating the agency problem.

Table 3 reports the fit of the estimated model. Overall, the model shows a good fit. The null of all moments being matched is rejected by the test of overidentifying restrictions, as shown in the bottom part of the table. However, this is not problematic since the set of target moments captures many dimensions of the data. The estimation matches many target moments with high precision and, as I show later, the model fit improves in the estimations across subsamples. As I show in Figure 4, the estimated model also captures quite well the actual process of CEO departures in the data, even outside the range of matched hazard rates— 10 years.[26]

[25]Since the estimated fixed wage $\underline{w}$ is close to zero I treat all flow compensation as incentive compensation.

[26]The empirical CEO hazard rates are also likely to suffer from some measurement error issues from the assignment of each CEO departure event to a specific tenure-year. The error in the assignment may remain unexplained by the exogenous departure process specified in expression (1).

**Table 3:** Model fit.

| Fit of moments in SMM | | | | |
|---|---|---|---|---|
| Moment description | Target | Estimated value | 95% Conf. Interval | |
| Mean of cash flows | -0.640 | -0.634 | -0.763 | -0.505 |
| AR(1) coefficient of cash flows | 0.488 | 0.505 | 0.483 | 0.528 |
| Variance of cash flows | 43.515 | 42.555 | 41.757 | 43.345 |
| Covariance of cash flows | -4.309 | -5.005 | -5.653 | -4.362 |
| Mean CEO flow pay | 0.246 | 0.247 | 0.244 | 0.249 |
| Variance of first diff. CEO flow pay | 0.096 | 0.087 | 0.085 | 0.089 |
| PPS of CEO flow pay | 0.009 | 0.009 | 0.009 | 0.010 |
| Tenure slope of CEO flow pay | 0.051 | 0.054 | 0.044 | 0.065 |
| CEO turnover rate | 11.409 | 11.132 | 10.814 | 11.453 |
| Hazard rate tenure 1 | 9.956 | 9.994 | 9.188 | 10.799 |
| Hazard rate tenure 2 | 9.668 | 10.124 | 9.502 | 10.758 |
| Hazard rate tenure 3-4 | 11.576 | 10.887 | 10.124 | 11.656 |
| Hazard rate tenure 5-6 | 13.810 | 11.714 | 10.790 | 12.651 |
| Hazard rate tenure 7-8 | 13.765 | 12.395 | 11.226 | 13.600 |
| Hazard rate tenure 9-10 | 13.913 | 12.952 | 11.511 | 14.449 |
| Test of overidentifying restrictions | | | | |
| $\chi^2$ | | 24.982 | | |
| p-value | | 0.000 | | |

The table shows the 15 moments used in the SMM estimation. The estimation sample consists of 4,098 CEOs in 2,235 firms in Execucomp appointed in the 1993-2013 period. Moments' standard errors are computed from 10,000 Monte Carlo simulations of the model using panels of 2,235 firms and parameter values in Table 2. The table shows the $\chi^2$ statistic and corresponding p-value for the test of overidentifying restrictions, which tests the null hypothesis that the empirical and simulated moments are jointly equal.

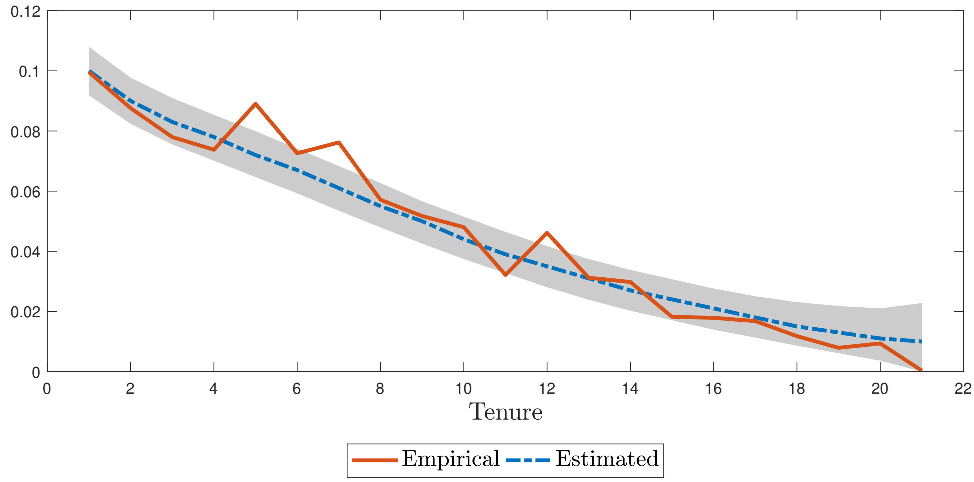


**Figure 4:** Empirical and predicted patterns in CEO departures. The Figure shows the unconditional percentage of CEOs departing at different tenure levels. I recover the empirical probabilities from the empirical hazards. The predicted probabilities come from averaging across 10,000 Monte Carlo simulations using the estimates in Table 2, each of which contains 2,235 firms. The shaded areas correspond to the 95 percent confidence region from the 10,000 simulated samples.

## 6.2 Results across the firm size distribution

Are dismissal policies more relevant in some types of firms than in others? For instance, smaller firms are less likely to invest in monitoring structures through corporate governance policies than bigger firms (Aggarwal et al., 2009, Bebchuk et al., 2009). In Table 4 I report results from estimations of the model across two subsamples of the firm size distribution.[27] I perform separate estimations for CEOs that manage firms in the bottom quartile and the top quartile of the firm size distribution. I select the subsets of largest and smallest firms to obtain sufficient variation in the target moments across subsamples, which may yield differences in the parameter estimates.[28] The bottom part of Table 4 shows that the match of the model, summarized by the test of overidentifying restrictions, is better for firms in the bottom size quartile. The result provides evidence that the standard principal-agent framework is better at capturing the behavior of CEO turnover and compensation in smaller firms.

The point estimates suggest that CEO replacement costs represent are relatively higher in bigger than in smaller firms, although the difference is not statistically significant. The estimated replacement cost in the top quartile of firm size is of 2.2 percent— standard error of 0.46 percent— and yields an estimated CEO dismissal rate of 0.37 percent, which evidences the uncommon nature of CEO dismissals in big firms. Relative to small firms, big firms face higher replacement costs and lower share of CEO dismissals, 0.08 percent. In contrast, the estimated CEO replacement cost in the bottom size quartile— 1.7 percent, standard error of 0.67— yields a share of CEO dismissals of almost 4 percent, which accounts for around one-third of all CEO departures in the subsample. Therefore, termination incentives seem to play a more important role for CEOs in small firms than their counterparts in large firms.

Two interpretations can explain that CEO dismissals are less prevalent and costlier in big than in small firms. The first is that bigger firms tend to attract more talented managers, who are scarce and hard to replace (Gabaix and Landier, 2008). Besides, CEOs may accumulate firm-specific human capital that is hard to substitute in bigger firms (Lustig et al., 2011). Thus, investors in big firms may be less willing to replace their CEOs due to the difficulty of finding suitable substitutes. Additionally, replace-

[27]In the main text I restrict the attention to relevant parameters in the estimations. Table F.1 in the Appendix shows set of target moments across subsamples, and Table F.2 reports the full set of estimation results. As the measure of size, I consider the beginning-of-spell value of firm assets. The median firm in the bottom quartile has 0.23 billion in 2015 U.S. dollars, whereas the median firm in the top quartile has 14.5 billion in 2015 U.S. dollars.

[28]For the sake of completeness, I show in Appendix G the target moments and parameter estimates for the remaining two subsamples of firm size. The results show that the CEO dismissal rate decreases in firm size, while CEO replacement costs are increasing in firm size. However, the confidence bands in the estimates are wide, which suggests that firms in both subsamples are similar in many dimensions.

**Table 4:** Parameters and estimation results: Firm size distribution

| Parameter | Full sample | Bottom size quartile | Top size quartile |
|---|---|---|---|
| $\beta_c$ | 0.800 | 0.776 | 0.827 |
| | (0.009) | (0.022) | (0.039) |
| $\sigma_\varepsilon$ | 1.067 | 1.656 | 0.827 |
| | (0.047) | (0.164) | (0.160) |
| $\lambda$ | 24.036 | 27.357 | 7.470 |
| | (1.191) | (1.598) | (2.322) |
| $\kappa$ | 3.417 | 1.686 | 2.240 |
| | (0.959) | (0.668) | (0.469) |
| Simulated firms | 2,235 | 684 | 606 |
| Estimated CEO turnover and CEO dismissal rates | | | |
| CEO turnover rate (%) | 11.132 | 12.328 | 10.372 |
| | (0.027) | (0.105) | (0.078) |
| CEO dismissal rate (%) | 1.185 | 3.964 | 0.367 |
| | (0.004) | (0.044) | (0.004) |
| Test of overidentifying restrictions | | | |
| $\chi^2$ | 24.982 | 7.635 | 42.655 |
| p-value | 0.000 | 0.178 | 0.000 |

The Table shows the parameter estimates for the full sample and firm size subsamples. Robust standard errors appear in parentheses. The table reports the $\chi^2$ statistic and corresponding p-value for the test of overidentifying restrictions, which tests the null hypothesis that the empirical and simulated moments are jointly equal. Moreover, the table reports the estimated CEO dismissal rate with the corresponding Monte Carlo standard errors in each subsample.

ment costs can represent adjustment costs arising from the reorganization of the firm after a CEO's departure. The second interpretation, meanwhile, is that CEOs in large firms may be entrenched and have more significant influence over the board's dismissal decision. Entrenchment can lead to higher estimates of the implicit costs of CEO replacement, capturing the directors' unwillingness to fire the CEO.[29] In the next section, I exploit heterogeneity in board independence within firm size quartiles to shed light on the governance interpretation of $\kappa$.

Further inspection of the parameter estimates in Table 4 suggests that heterogeneity in the extent of agency problems is also important in explaining the difference in CEO dismissal rates across firm sizes. Both the moral hazard parameter, $\lambda$, and the idiosyncratic component of firm cash flows, $\sigma_\varepsilon$, are decreasing with firm size. Thus, the estimates are consistent with smaller firms having weaker monitoring structures, which make CEO dismissals more important in the provision of incentives. In line with the estimates for the full sample, deferred compensation and dismissal policies allow firms to reduce by 55 and 68 percent of the static agency costs, respectively, in small and big firms.

[29]Firms may have a poor design of their governance structures, which enhance the rent-extraction ability of managers. Alternatively, weak governance structures may allow shareholders to reduce the direct burden of incentive compensation schemes (Hermalin, 2005, Almazan and Suarez, 2003).

## 6.3 Results across the governance quality distribution

Can CEO entrenchment— or poor governance in general— explain the size of CEO replacement costs? To answer this question, I split the two firm size subsamples according to their degree board independence, which is a popular proxy for the quality of governance (Arena and Ferris, 2007, Gillette et al., 2003, Weisbach, 1988, Fama and Jensen, 1983). I divide each size subsample into two separate groups of firms: those with an above-median number of independent directors and those with a below-median number of independent directors.[30]

I report the estimation results in Table 5. The bottom part of the table reports that the match of the model, summarized by the test of overidentifying restrictions, is worse for big firms with more independent directors. This result provides strong evidence that the principal-agent framework fits better those firms where managers face less strict monitoring, either in small firms or big firms with poor governance.

We would expect CEO entrenchment to be less of a concern with more independent boards. Therefore, firms with more independent boards should display smaller CEO replacement costs because CEOs are less likely to effectively resist against their replacement. The estimates confirm this hypothesis for big firms. The main difference between parameter estimates within big firms lies on the CEO replacement cost parameter $\kappa$— 4.6 percent for firms with less independent boards and 0.57 percent for firms with more independent boards, although the precision of the estimates is low. The estimates yield an estimated CEO dismissal rate in firms with more independent directors of 2.5 percent, while it is of 0.19 percent in firms with fewer independent directors. All in all, the results suggest that within big firms director independence is informative about less costly CEO replacements, but is uninformative about the degree of agency problems— point estimates of $\lambda$ are very close for both subsamples.

In contrast, the relationship between board independence and CEO replacement costs reverses for smaller firms. In the bottom size quartile, firms with more independent directors have higher replacement costs than firms with fewer independent directors. The difference between estimates is substantial, 3 percent of assets, and statistically significant. Moreover, small firms with less board independence feature lower CEO fixed pay— low $\underline{w}$— and earlier CEO flow compensation— lower $\beta_c$— than their counterparts

[30] I measure board independence as the number of independent directors in the board at the beginning of a CEO's spell. Big firms have bigger boards than small firms so, mechanically, big firms tend to have more independent directors than small firms. Thus, separate estimations across subsamples of board independence would automatically lead to similar results than estimations across firm sizes. The use of subsamples within size brackets limit the impact of size on the analysis of governance. This is at the cost of reduced precision in the estimates.

**Table 5:** Parameters and estimation results: Governance quality distribution

| Parameter | Bottom size quartile | | Top size quartile | |
|---|---|---|---|---|
| | Board independence | | Board independence | |
| | $\leq$ median | $>$ median | $\leq$ median | $>$ median |
| $\beta_c$ | 0.728 | 0.946 | 0.881 | 0.802 |
| | (0.027) | (0.030) | (0.036) | (0.031) |
| $\sigma_\varepsilon$ | 2.407 | 2.269 | 0.631 | 0.516 |
| | (1.009) | (0.416) | (0.085) | (0.113) |
| $\lambda$ | 19.804 | 30.998 | 10.716 | 9.746 |
| | (8.266) | (7.760) | (2.409) | (2.846) |
| $\underline{w}$ | 0.165 | 0.295 | 0.001 | 0.003 |
| | (0.034) | (0.052) | (0.010) | (0.008) |
| $\kappa$ | 0.491 | 3.575 | 4.624 | 0.567 |
| | (0.351) | (1.106) | (2.613) | (0.130) |
| Simulated firms | 195 | 111 | 304 | 160 |
| Estimated CEO turnover and CEO dismissal rates | | | | |
| CEO turnover rate (%) | 12.819 | 11.299 | 9.991 | 8.828 |
| | (0.657) | (0.736) | (0.148) | (0.243) |
| CEO dismissal rate (%) | 5.009 | 2.047 | 0.185 | 2.495 |
| | (0.209) | (0.145) | (0.004) | (0.102) |
| Test of overidentifying restrictions | | | | |
| $\chi^2$ | 7.560 | 7.594 | 8.671 | 15.999 |
| p-value | 0.182 | 0.180 | 0.123 | 0.007 |

The Table shows the parameter estimates for the board independence subsamples. Robust standard errors appear in parentheses. The table reports the $\chi^2$ statistic and corresponding p-value for the test of overidentifying restrictions, which tests the null hypothesis that the empirical and simulated moments are jointly equal. Moreover, the table reports the estimated CEO dismissal rate with the corresponding Monte Carlo standard errors in each sample.

with greater board independence.

Three interrelated reasons can explain the result for small firms. First, independent boards can make better CEO selection and replacement decisions. More efficient decisions translate into better firm-CEO matches that translate into higher implicit replacement costs. The higher estimated fixed wage $\underline{w}$ suggests that this is a possibility. That is, investors in firms with better governance must increase fixed pay to retain their CEOs. Second, independent directors in small firms may face reputation concerns that limit their willingness to dismiss a CEO. Lastly, more independent boards may also be optimal in firms that face large costs from CEO replacements— e.g., due to transition and reorganization costs— and want to reduce the turnover rate of their managers.[31]

[31] In Appendix G I report the parameter estimates for the remaining two subsamples of firm size. The patterns that arise are similar to the estimates for big firms, although the confidence bands of the parameters are wide.

## 6.4 Counterfactuals

Given the results above, some counterfactual exercises are of interest. What is the contribution of dismissal threats to reduce the moral hazard problem? How would incentive compensation change if investors can replace CEOs at almost zero cost? Would dismissal threats be a more important source of incentives for CEOs in such a world?

Using the estimates in Table 5, I obtain the CEO compensation and dismissal rates consistent with the optimal contract for values of the replacement cost parameter, $\kappa$, ranging from 0.25— equivalent to the average annual flow compensation received by a CEO in the full sample— to 7 percent. Figure 5 depicts the behavior of the average CEO flow compensation as a function of $\kappa$. The exercise keeps constant the extent of the agency problem and, hence, the change in incentive compensation captures the change in CEO incentives that are provided more or less through dismissal threats.

The top panels in Figure 5 shows that incentive compensation in small firms would remain almost unchanged if replacement costs dropped to values close to zero. Such conclusion changes once one takes into account the different parameter estimates across firms with different governance quality. Investors in firms with more independent directors could reduce their incentive compensation costs by 10 percent if replacement costs dropped to the value of replacement costs in small firms with fewer independent directors. In other words, small firms could provide the same level of incentives by using a 10 percent lower incentive pay and increasing the threat of dismissal.

The bottom panels in Figure 5 show that incentive compensation in big firms is more sensitive to changes in replacement costs than for smaller firms. Firms with fewer independent directors could reduce by close to 30 percent the size of incentive compensation if their replacement costs dropped to levels similar to those of firms with greater director independence. That is, big firms with worse governance could induce the same level of incentives with a 30 percent lower incentive pay and increasing the threat of dismissal.

How large is the change in the realized CEO dismissal rates that generate these changes in CEO incentive compensation? I show this in Figure 6. For the subset of smaller firms the CEO dismissal rate increases from the estimated 4 percent to 6 percent when $\kappa$ decreases up to 0.25 percent. The sensitivity of incentive compensation to changes in $\kappa$ in big firms contrasts with the little sensitivity of CEO dismissals. The impact on dismissals for bigger firms is limited and only relevant for values of the replacement costs close to zero, where CEO turnover in the reaches around 5 percent. This behavior is suggestive that small changes in the realized dismissal rate can be associated with sizable changes in CEO incentives.

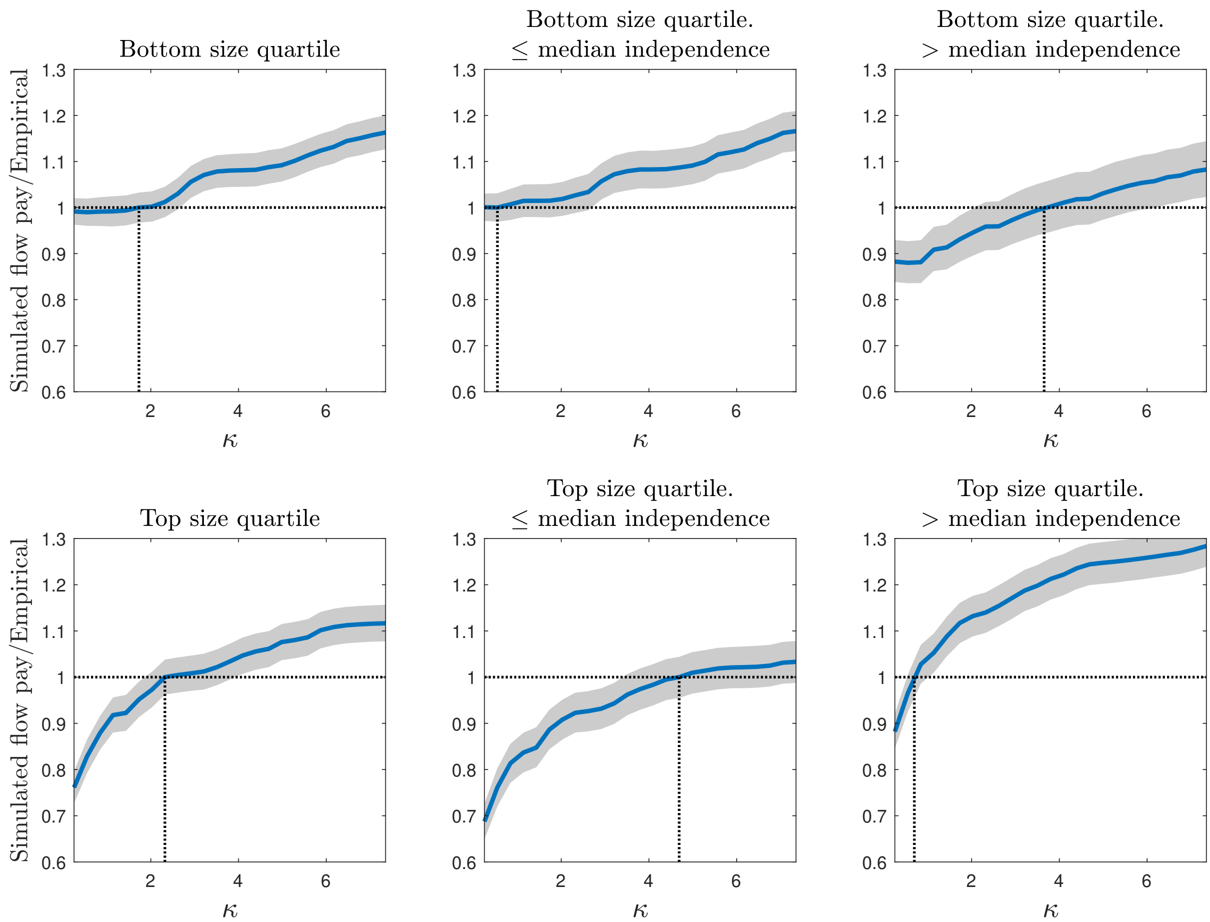


**Figure 5:** Impact of CEO replacement costs ($\kappa$) on average CEO flow compensation ($w$) relative to its empirical value. Results are normalized with respect to the empirical value of expected CEO flow compensation. The figure plots the average CEO flow compensation from 10,000 Monte Carlo simulations of panels of 200 firms, using parameter values from Table 5. The shaded areas represent the 95% confidence intervals from the 10,000 Monte Carlo simulations.

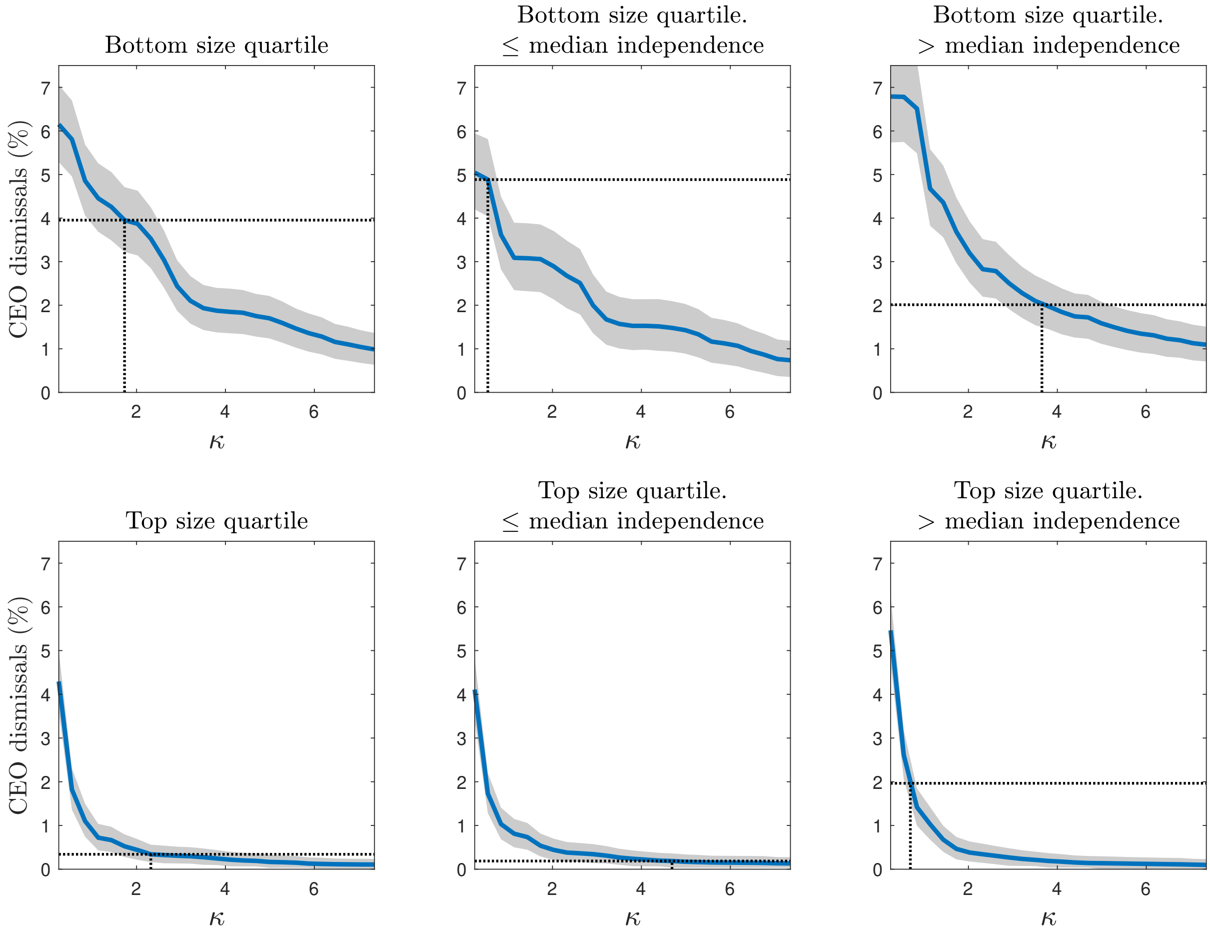


**Figure 6:** Impact of CEO replacement costs ($\kappa$) on the CEO dismissal rate. The figure plots the average CEO dismissal from 10,000 Monte Carlo simulations of panels of 200 firms, using parameter values from Table 5. The shaded areas represent the 95% confidence intervals from the 10,000 Monte Carlo simulations.

# 7 Conclusions

In this paper I estimate a dynamic moral hazard model to understand the role of dismissal policies *vis à vis* compensation in providing incentives to CEOs. The model allows for endogenous CEO compensation and dismissals that depend on the performance history of the CEO in the firm. Moreover, the model includes an exogenous departure process, which allows computing an endogenous share of CEO dismissals. I identify the model parameters by exploiting the joint variation in firm performance, CEO compensation and CEO turnover frequencies.

Using data from Execucomp for the 1993-2013 period and SMM estimations I find that firms face a cost equivalent to 3.4 percent of the value of assets when replacing a CEO. The estimated share of CEO dismissals is small: around 1.2 percent of CEOs are dismissed due to poor firm performance. This is in line with previous findings in the literature based on news-based search algorithms (Huson et al., 2001, Taylor, 2010). The estimated model suggests that 40 percent of all CEO dismissals are concentrated in the first five years of CEO tenure. The dynamic aspects of the CEO contract allow investors to reduce the static costs of the underlying agency problem by 60 percent.

I perform several estimations across size and board independence subsamples to obtain further insight about the frictions that generate the CEO replacement costs. The estimates show that the incentive role of dismissals is more prevalent in small firms, mostly due to the harsher agency frictions. I explore the governance interpretations of the CEO replacement costs by estimating the model separately across subsamples of board independence. CEO entrenchment explains a 4 percent difference in CEO replacement costs within bigger firms. In contrast, I find that small firms with more independent boards face greater replacement costs than those firms with apparently worse governance structures. Hence, the results suggest that CEO replacement costs in small firms talent-attraction or reputation concerns rather than governance problems.

All in all, the results imply that dismissal threats play a small role in CEO incentives. The bulk of CEO incentives comes from flow and deferred compensation. However, in a counterfactual analysis, I find that a sufficiently large reduction in CEO replacement costs would considerably modify the relative importance of both policies in the incentive packages of CEOs. This is particularly true in big firms with worse corporate governance. From a policy perspective, the results of this paper suggest that observing a higher frequency of CEO departures or a CEO dismissal should not be taken as an indication of better governance or monitoring over the actions of the CEOs. Conversely, in small firms, more frequent dismissals indicate lower governance quality and worse CEO hiring decisions.

The results and methodology in this paper provide new avenues for future research. For instance, mixed classification algorithms that use the structural approach but also incorporate information from news-based algorithms may allow to improve the knowledge about CEO departure events. Moreover, I can consider more flexible specifications of firm performance— e.g, multi-dimensional, or non-linear— to better capture heterogeneity across firms and provide more accurate measures of the CEO dismissal rate.

# References


Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986) 'Optimal cartel equilibria with imperfect monitoring.' *Journal of Economic Theory* 39, 251–269

_ (1990) 'Toward a theory of discounted repeated games with imperfect monitoring.' *Econometrica* 58(5), 1041–1063

Aggarwal, Reena, Isil Erel, René Stulz, and Rohan Williamson (2009) 'Differences in governance practices between U.S. and foreign firms: Measurement, causes, and consequences.' *Review of Financial Studies* 22(8), 3131–3169

Almazan, Andres, and Javier Suarez (2003) 'Entrenchment and severance pay in optimal governance structures.' *Journal of Finance* 58(2), 519–548

Arena, Matteo P., and Stephen P. Ferris (2007) 'When managers bypass shareholder approval of board appointments: Evidence from the private security market.' *Journal of Corporate Finance* 13(4), 485 – 510. Private Equity, Leveraged Buyouts and Corporate Governance

Bebchuk, Lucian, Alma Cohen, and Allen Ferrell (2009) 'What matters in corporate governance?' *Review of Financial Studies* 22(2), 783–827

Biais, Bruno, Thomas Mariotti, and Jean Charles Rochet (2013) 'Dynamic financial contracting.' In *Advances in economics and econometrics, Tenth world congress,* ed. Daron Acemoglu, Manuel Arellano, and Eddie Dekel, vol. I of *Econometric society monographs* Cambridge University Press pp. 125–171

Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet (2007) 'Dynamic security design: Convergence to continuous time and asset pricing implications.' *The Review of Economic Studies* 74(2), pp. 345–390

Coles, Jeffrey L., Naveen D. Daniel, and Lalitha Naveen (2006) 'Managerial incentives and risk-taking.' *Journal of Financial Economics* 79(2), 431–468

_ (2013) 'Calculation of compensation incentives and firm-related wealth using Execucomp: Data, program, and explanation.' Working Paper, July

Core, John, and Wayne Guay (2002) 'Estimating the value of employee stock option portfolios and their sensitivities to price and volatility.' *Journal of Accounting Research* 40(3), 613–630

Coughlan, Anne T., and Ronald M. Schmidt (1985) 'Executive compensation, management turnover, and firm performance: An empirical investigation.' *Journal of Accounting and Economics* pp. 43–66

DeMarzo, Peter M., and Michael J. Fishman (2007) 'Optimal long-term financial contracting.' *Review of Financial Studies* 20(6), 2079–2128

DeMarzo, Peter M., Michael J. Fishman, Zhiguo He, and Neng Wang (2012) 'Dynamic agency and the q theory of investment.' *Journal of Finance* 67(6), 2295–2340

Denis, David J., and Diane K. Denis (1995) 'Performance changes following top management dismissals.' *The Journal of Finance* 50(4), 1029–1057

Edmans, Alex, and Xavier Gabaix (2015) 'Executive compensation: A modern primer.' NBER Working Papers 21131, National Bureau of Economic Research, Inc, April

Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009) 'A multiplicative model of optimal CEO incentives in market equilibrium.' *Review of Financial Studies* 22(12), 4881–4917

Eisfeldt, Andrea L., and Camelia M. Kuhnen (2013) 'CEO turnover in a competitive assignment framework.' *Journal of Financial Economics* 109(2), 351–372

Erickson, Timothy, and Toni M. Whited (2002) 'Two-step GMM estimation of the errors-in-variables model using high-order moments.' *Econometric Theory* 18(03), 776–799

Fama, Eugene F., and Michael C. Jensen (1983) 'Separation of ownership and control.' *The Journal of Law & Economics* 26(2), 301–325

Fee, Edward C., Charles J. Hadlock, Jing Huang, and Joshua R. Pierce (2015) 'Robust models of CEO turnover: New evidence on relative performance evaluation.' Working Paper, May

Frydman, Carola, and Dirk Jenter (2010) 'CEO compensation.' *Annual Review of Financial Economics* 2(1), 75–102

Gabaix, Xavier, and Augustin Landier (2008) 'Why has CEO pay increased so much?' *The Quarterly Journal of Economics* 123(1), 49–100

Gabaix, Xavier, Augustin Landier, and Julien Sauvagnat (2014) ' pay and firm size: An update after the crisis.' *Economic Journal* 124(574), F40–F59

Gayle, George-Levi, and Robert A. Miller (2009) 'Has moral hazard become a more important factor in managerial compensation?' *American Economic Review* 99(5), 1740–69

Gayle, George-Levi, Limor Golan, and Robert A. Miller (2015) 'Promotion, turnover, and compensation in the executive labor market.' *Econometrica* 83(6), 2293–2369

Gillan, Stuart L., and Nga Q. Nguyen (2016) 'Incentives, termination payments, and CEO contracting.' *Journal of Corporate Finance* 41, 445 – 465

Gillette, Ann B., Thomas H. Noe, and Michael J. Rebello (2003) 'Corporate board composition, protocols, and voting behavior: Experimental evidence.' *Journal of Finance* 58(5), 1997–2032

Goffe, William L., Gary D. Ferrier, and John Rogers (1994) 'Global optimization of statistical functions with simulated annealing.' *Journal of Econometrics* 60(1-2), 65–99

Gopalan, Radhakrishnan, Todd Milbourn, Fenghua Song, and Anjan V. Thakor (2014) 'Duration of executive compensation.' *The Journal of Finance* 69(6), 2777–2817

Han, Chirok, and Peter C. B. Phillips (2010) 'GMM estimation for dynamic panels with fixed effects and strong instruments at unity.' *Econometric Theory* 26(01), 119–151

He, Zhiguo (2009) 'Optimal executive compensation when firm size follows geometric Brownian motion.' *Review of Financial Studies* 22(2), 859–892

Hennessy, Christopher A., and Toni M. Whited (2005) 'Debt dynamics.' *Journal of Finance* 60(3), 1129–1165

_ (2007) 'How costly is external financing? Evidence from a structural estimation.' *Journal of Finance* 62(4), 1705–1745

Hermalin, Benjamin E. (2005) 'Trends in corporate governance.' *Journal of Finance* 60(5), 2351–2384

Hermalin, Benjamin E., and Michael S. Weisbach (2012) 'Information disclosure and corporate governance.' *Journal of Finance* 67(1), 195–234

Holmstrom, Bengt, and Jean Tirole (1997) 'Financial intermediation, loanable funds, and the real sector.' *The Quarterly Journal of Economics* 112(3), 663–91

Huson, Mark R., Paul H. Malatesta, and Robert Parrino (2004) 'Managerial succession and firm performance.' *Journal of Financial Economics* 74(2), 237–275

Huson, Mark R., Robert Parrino, and Laura T. Starks (2001) 'Internal monitoring mechanisms and CEO turnover: A long-term perspective.' *Journal of Finance* 56(6), 2265–2297

Jensen, Michael, and Kevin J. Murphy (1990) 'CEO incentives — it's not how much you pay, but how.' *Journal of Applied Corporate Finance* 3(3), 36–49

Jenter, Dirk, and Katharina Lewellen (2014) 'Performance-induced CEO turnover.' Research Papers 1992, Stanford University, Graduate School of Business

Kaplan, Steven N., and Bernadette A. Minton (2012) 'How has CEO turnover changed?' *International Review of Finance* 12(1), 57–87

Kopecky, Karen, and Richard Suen (2010) 'Finite state markov-chain approximations to highly persistent processes.' *Review of Economic Dynamics* 13(3), 701–714

Li, Shaojin, Toni M. Whited, and Yufeng Wu (2016) 'Collateral, taxes, and leverage.' *The Review of Financial Studies* 29(6), 1453

Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh (2011) 'Technological change and the growing inequality in managerial compensation.' *Journal of Financial Economics* 99(3), 601 – 627

Morellec, Erwan, Boris Nikolov, and Norman Schurhoff (2012) 'Corporate governance and capital structure dynamics.' *The Journal of Finance* 67(3), 803–848

Murphy, Kevin J. (1999) 'Executive compensation.' In *Handbook of Labor Economics,* ed. O. Ashenfelter and D. Card, vol. 3 of *Handbook of Labor Economics* (Elsevier) chapter 38, pp. 2485–2563

Nickerson, Jordan (2013) 'A structural estimation of the cost of suboptimal matching in the CEO labor market.' Available at SSRN: https://ssrn.com/abstract=2356680 or http://dx.doi.org/10.2139/ssrn.2356680

Nikolov, Boris, and Toni M. Whited (2014) 'Agency conflicts and cash: Estimates from a dynamic model.' *Journal of Finance* 69(5), 1883–1921

Pakes, Ariel, and David Pollard (1989) 'Simulation and the asymptotics of optimization estimators.' *Econometrica* 57(5), 1027–57

Parrino, Robert (1997) 'CEO turnover and outside succession a cross-sectional analysis.' *Journal of Financial Economics* 46(2), 165–197

Peters, Florian S., and Alexander F. Wagner (2014) 'The executive turnover risk premium.' *Journal of Finance* 69(4), 1529–1563

Press, William H., Saul A. Teukolsky, William T. Vetterling, and Biran P. Flannery (1996) *Numerical recipes in Fortran 90: The art of parallel scientific computing,* second edition ed. (Cambridge University Press)

Rouwenhorst, K. Geert (1995) 'Asset pricing implications of equilibrium business cycle models.' In *Frontiers of Business Cycle Research,* ed. T.F Cooley (Princeton University Press) pp. 294–330

Sannikov, Yuliy (2008) 'A continuous-time version of the principal-agent problem.' *Review of Economic Studies* 75(3), 957–984

Spear, Stephen E., and Cheng Wang (2005) 'When to fire a CEO: Optimal termination in dynamic contracts.' *Journal of Economic Theory* 120(2), 239–256

Spear, Stephen E, and Sanjay Srivastava (1987) 'On repeated moral hazard with discounting.' *Review of Economic Studies* 54(4), 599–617

Taylor, Lucian A. (2010) 'Why are CEOs rarely fired? Evidence from structural estimation.' *Journal of Finance* 65(6), 2051–2087

Taylor, Lucian A. (2013) 'CEO wage dynamics: Estimates from a learning model.' *Journal of Financial Economics* 108(1), 79–98

Warner, Jerold B., Ross Watts, and Karen H. Wruck (1988) 'Stock prices and top management changes.' *Journal of Financial Economics* 20(1-2), 461–492

Weisbach, Michael (1988) 'Outside directors and CEO turnover.' *Journal of Financial Economics* 20(1-2), 431–460

Yermack, David (2006) 'Golden handshakes: Separation pay for retired and dismissed CEOs.' *Journal of Accounting and Economics* 41(3), 237–256

# A The optimal contract

In this appendix, I solve for the optimal termination, flow, and deferred compensation policies in the model introduced in Section 3. The derivations and discussion summarize the those of DeMarzo and Fishman (2007) and Biais et al. (2007). For clarity in the exposition, I focus on the case where cash flows are i.i.d., there are no exogenous CEO departures, and the investors' outside option is $\underline{v} = 0$. Straightforward modifications would apply otherwise. Proceeding by backward induction, I characterize the optimal policies in the next two propositions.

**Proposition A.1.** *The optimal contract at the CEO's tenure-year $T$ is characterized by two regions in the value of the CEO's deferred compensation $u$:*

- *For $u \geq u^l \equiv \underline{w} - \lambda\varepsilon_1$, the CEO receives compensation after shock $\varepsilon_j$ equal to $w_j = (u - u^l) + \underline{w} + (j-1)\lambda\Delta$*
- *For $u < u^l$ continuation must be randomized. With probability $q = \max\left\{\frac{u^l - u}{u^l}, 0\right\}$ the CEO is dismissed. Otherwise, the CEO begins the period with deferred compensation $u^l$.*

*Proof.* At $T-1$ investors solve

$$V_{T-1}(u) = \max_{\substack{q \in [0,1] \\ \{w_j\}_{j=1}^n}} (1-q)\left[\mu - \sum_{j=1}^n p_j w_j\right]$$

$$u = (1-q)\sum_{j=1}^n p_j w_j$$

$$w_j \geq w_{j-1} + \lambda\Delta, \text{ for } j = 2, ..., n$$

$$(w_j)_{i=1}^n \in [\underline{w}, \infty)^n$$

This is the Holmstrom and Tirole (1997) static setting. The optimal policies imply that the incentive compatibility conditions bind, $w_j = w_{j-1} + \lambda\Delta = w_1 + (j-1)\lambda\Delta = w_1 + \lambda(\varepsilon_j - \varepsilon_1)$. The last step comes from $(j-1)\lambda\Delta = \sum_{k=2}^j (\varepsilon_k - \varepsilon_{k-1})$. It would be optimal to minimize the cost of the contract by setting $w_1 = \underline{w}$. This is only feasible, from the promise-keeping constraint and $E(\varepsilon) = 0$, for $u = u^l \equiv \underline{w} - \lambda\varepsilon_1$. Thus, for $u < u^l$ continuation must be randomized, setting $q = \max\left\{\frac{u^l - u}{u^l}, 0\right\}$. For $u \geq u^l$ we have that $w_1 = \underline{w} + (u - u^l)$. □

**Proposition A.2.** *The optimal contract at the CEO's tenure-years $t = 1, ..., T-1$ is characterized by two thresholds in the CEO's deferred compensation $u^l \equiv \underline{w} - \lambda\varepsilon_1$ and $u^*_t$, with $u^l \leq u^*_t$, such that:*

- *For $u \geq u_t^* + \underline{w}$ the CEO receives flow compensation under all realizations of $\varepsilon_j$, $w_j = u - u_t^* + (j-1)\lambda\Delta$. There exists some $\widehat{u}_t$ such that $u_j = \widehat{u}_t$ for all $\varepsilon_j$.*

- *For $u \in [u^l, u_t^* + \underline{w})$, the CEO receives compensation in some shocks. More specifically:*

$$w_j = \max\{\underline{w}, u - u_t^* + (k-1)\lambda\Delta\} = \max\{\underline{w}, u + \lambda\varepsilon_j - \beta_c \widehat{u}_t\}$$
$$u_j = \min\left\{\frac{1}{\beta_c}\left(u - \underline{w} + \lambda\varepsilon_j\right), \widehat{u}_t\right\}$$

- *For $u < u^l$ continuation must be randomized. With probability $q = \max\left\{\frac{u^l - u}{u^l}, 0\right\}$ the CEO is dismissed. Otherwise, the CEO begins the period with deferred compensation $u^l$.*

*Proof.* For $t < T$, the optimal contract arises from the solution to the following problem:

$$V_t(u) = \max_{q,\{w_j,u_j\}_{j=1}^n} q\left[\mu - \sum_{j=1}^n p_j w_j + \beta \sum_{j=1}^n p_j V_{t+1}(u_j)\right]$$
$$u = q\sum_{j=1}^n p_j\left(w_j + \beta_c u_j\right)$$
$$w_j + \beta_c u_j \geq w_{j-1} + \beta_c u_{j-1} + \lambda\Delta, \text{ for } j = 2, ..., n$$
$$(w_j, u_j)_{i=1}^n \in [\underline{w}, \infty)^n \times [0, \infty)^n$$

First, I solve for the optimal policies $\{w_{j,}, u_j\}$ assuming that $q = 0$.

$$V_t(u) = \max_{\{w_j,u_j\}_{j=1}^n} \left\{\mu - u + \sum_{j=1}^n p_j\left[\beta_c u_j + \beta V_{t+1}(u_j)\right]\right\}$$

Suppose first that, given $u$, the optimal deferred compensation policies are equal to $u_j = \widehat{u}_t$ for all $j = 1, ..., n$ and given by the solution to $\max_u\{\beta_c u + \beta V_{t+1}(u)\}$, which is a concave function (Biais et al., 2007, DeMarzo and Fishman, 2007). Then, incentive compatibility requires that $w_j = w_1 + \lambda\Delta(j-1)$. Using this result in the promise-keeping constraint we find that $w_1 = u - \beta_c\widehat{u}_t + \lambda\varepsilon_1$.

Then, for $u - \underline{w} > u_t^* \equiv \beta_c\widehat{u} - \lambda\varepsilon_1$ the CEO receives incentive flow compensation ($w_1 > \underline{w}$) for all realizations of $\varepsilon_j$. The policies are given, for all $j = 1, ..., n$, by

$$u_j = \widehat{u}_t$$
$$w_j = u - u_t^* + (j-1)\lambda\Delta$$

This is because $u_t^*$ is the largest value of deferred compensation that enhances incentive compatibility. Therefore, it is optimal to provide incentives only through flow compensation for deferred compensation levels above $u_t^*$.

Now suppose that $u_i < \widehat{u}_t$ for some $j \leq n$. From the incentive-compatibility constraints we get that $w_j + \beta_c u_j = w_1 + \beta_c u_1 + (j-1)\lambda\Delta$. Using this result in the promise-keeping constraint we get

$$u_1 = \frac{1}{\beta_c}\left(u - w_1 + \lambda\varepsilon_1\right)$$
$$w_1 = \underline{w}$$

Where the second result must arise from investors minimizing the cost of the contract. Now, we can find a shock $\varepsilon_k$, $k > 1$, such that the CEO receives some incentives through flow compensation, $w_k > \underline{w}$. In particular we have that

$$u_j = u_{j-1} + \frac{\lambda\Delta}{\beta_c} \text{ if } j < k$$
$$w_{j+1} = w_{j-1} + \lambda\Delta \text{ if } j > k$$
$$w_k + \beta_c\widehat{u}_t = \underline{w} + \beta_c u_{k-1} + \lambda\Delta$$

Using the first line we have that $u_j = u_1 + (j-1)\frac{\lambda\Delta}{\beta_c} = \frac{1}{\beta_c}(u - \underline{w} + \lambda\varepsilon_j)$. Therefore, we get $w_k + \beta_c\widehat{u}_t = u + \lambda\varepsilon_k$. This means that the CEO receives incentive flow compensation in shock $k$ if and only if $u - \underline{w} > u_t^* - (k-1)\lambda\Delta = \beta_c\widehat{u}_t - \lambda\varepsilon_k$. Then, compensation policies are given by

$$w_j = \underline{w} \text{ if } j < k$$
$$u_j = \frac{1}{\beta_c}\left(u - \underline{w} + \lambda\varepsilon_j\right) \text{ if } j < k$$
$$w_j = u - u^* + (k-1)\lambda\Delta \text{ if } j \geq k$$
$$u_j = \widehat{u}_t \text{ if } j \geq k$$

The second policy follows from the incentive-compatibility constraints up to shock $k$, with $w_j = \underline{w}$, $j = 1, ..., k-1$. The third follows from setting $u_j = \widehat{u}_t$ for all $j \geq k$ and using the policies for $j < k$.

If there is no $j$ such that $u - \underline{w} > \beta_c\widehat{u}_t - \lambda\varepsilon_j$, then $w_j = \underline{w}$ for all $j = 1, ..., n$. Hence, $u_j = \frac{1}{\beta_c}\left(u - \underline{w} + \lambda\varepsilon_j\right)$ for all $j$, as it arises from $u_1$ and the incentive-compatibility constraint.

Lastly, the optimal policy $q$ implies a randomization for all values of $u \in [0, u^l]$ where $u^l$ is the minimum continuation value below which the CEO's limited liability would be satisfied. From the expression for $u_j$ we have that $u^l = \underline{w} - \lambda\varepsilon_1$. Thus, $q = \max\left\{\frac{u^l - u}{u^l}, 0\right\}$.
□

# B Numerical solution to the optimal contract

In this appendix, I describe the numerical algorithm to solve for the value function $V_t(u; \underline{v})$ and the equilibrium outside option $\underline{v}$. First, I construct equally-spaced grids for each $u_t$, $t = 1, ..., T$, in the interval $[0, \overline{u}_t]$, where $\overline{u}_t$ is taken as the present value of the CEO's unpledgeable income

$$\overline{u}_t = (\underline{w} - \lambda\varepsilon_1) \sum_{s=t}^{T} \beta^{t-s}(1 - h_s)$$

I choose a grid size of 5000 points in $u_t$ for each $t$.

The permanent component of cash flows, $\eta$, is constructed as a discrete approximation to a stationary AR(1) process with unconditional mean $\mu$, autocorrelation parameter $\rho$ and symmetric disturbances with variance $\sigma_\eta^2$. I follow the Rouwenhorst (1995) procedure to construct the process, using $s = 5$ states. With this method, I can match the conditional and unconditional mean and variance, and the first-order autocorrelation of any stationary AR(1) process. The method is more reliable than others in approximating highly persistent processes and generating accurate model solutions (Kopecky and Suen, 2010).

The transitory component of cash flows is constructed with $n = 11$ points in an equally-spaced grid, so that $\{\varepsilon_1, ..., \varepsilon_n\}$ has a discrete approximation to a normal distribution with zero mean and variance $\sigma_\varepsilon^2$. The first and last points in the grid are set to $\varepsilon_1 = -b\sigma_\varepsilon$ and $\varepsilon_n = b\sigma_\varepsilon$, where $b$ is a free parameter to be determined. The rest of the grid is constructed by setting $\varepsilon_j = -b\sigma_\varepsilon + \frac{2b\sigma_\varepsilon}{n-1}(j-1)$. Denote by $m_j = \frac{1}{2}(\varepsilon_{j+1} + \varepsilon_j)$ the midpoint between grid points $\varepsilon_{j+1}$ and $\varepsilon_j$. Let $\Phi$ denote the cumulative density function (CDF) of the standard normal distribution. For $j = 2, ..., n-1$, $p_j$ (the probability assigned to $\varepsilon_j$) can be computed as follows:

$$p_j = \Phi\left(\frac{m_j}{\sigma_\varepsilon}\right) - \Phi\left(\frac{m_{j-1}}{\sigma_\varepsilon}\right)$$

In other words, $p_j$ is the probability that a draw from the normal distribution falls into the interval $[m_{j-1}, m_j]$. For the end points we have

$$p_1 = \Phi\left(\frac{m_1}{\sigma_\varepsilon}\right)$$
$$p_n = 1 - \Phi\left(\frac{m_{n-1}}{\sigma_\varepsilon}\right)$$

The parameter $b$ is then chosen so that $\sigma_\varepsilon^2 = \sum_{j=1}^{n} p_j \varepsilon_j^2$. It can be shown that the choice of $b$ is independent of $\sigma_\varepsilon^2$, and only depends on the number of nodes in the grid, $n$.

The optimal contract is solved for each state $\eta$ and $u$ by backwards induction from $t = T$ to $t = 1$, setting $V_{T+1}(\eta) = \underline{v}(\eta) - \kappa$, following the solution in Appendix A. Starting with an initial guess $\underline{v}_{(0)}$, I follow a fixed-point iteration algorithm to find the vector $\underline{v}$ of outside options that is consistent with the contract. That is, after $k-1$ iterations, I compute the new guess for the equilibrium outside options as:

$$\underline{v}_{(k)}(\eta) = \max_u \left\{V_1(u, \eta; \underline{v}_{(k-1)})\right\} \text{ for all } \eta.$$

I stop iterating when

$$\frac{\|\underline{v}_{(k)} - \underline{v}_{(k-1)}\|_\infty}{\|\underline{v}_{(k-1)}\|_0 + \tau_1} < \tau_2$$

where $\tau_1 = 10^{-6}$ and $\tau_2 = 10^{-8}$. $\|.\|_\infty$ denotes the maximum norm, $\|x\|_\infty = \max\{|x_1|, ..., |x_n|\}$, and $\|x\|_0 = \min\{|x_1|, ..., |x_n|\}$.

# C Data appendix

This appendix provides details on the construction of the estimation sample and the variables used in the estimation. In the following, variable names within quotation marks (“ ”) refer to variables appearing in Execucomp or Compustat. I merge the data with ISS using “cusip” identifiers.

## Identifying CEOs in the database

I construct the information regarding a CEO’s tenure duration as follows. First, following Taylor (2013), I fill in CEO indicators in Execucomp for those firm-years in which the variable “ceoann” identifies the executive as the CEO in the firm. In addition, I label an executive as a CEO in a firm/year if (i) Execucomp lists no one as CEO in the given firm/year, and (ii) either (a) the individual was CEO of the firm in the previous and following year; (b) the individual was CEO in the previous year, and we do not know who was CEO in the following year; or (c) the individual was CEO in the following year, and we do not know who was CEO in the previous year.

## Dates of CEO appointment and departure

The beginning of a CEO's tenure is taken as the minimum of (i) the year given by item "becameceo" and (ii) the first year in which the item "ceoann" highlights the executive as a CEO. Equivalently, a CEO is assumed to leave office in the maximum of (i) the year given by item "leftofc" and (ii) the last year when item "ceoann" identifies the executive as the CEO.

I further adjust the appointment and replacement dates in the following way. I assume the CEO's first fiscal year is the one when she/he completes at least six full months in office. Similarly, I assume that the CEO's last fiscal year is the last in which she/he completes six full months in office. Then, a CEO's tenure is the difference between the current year and the year of appointment. I drop those CEOs who disappear from the database due to reasons different from mergers, acquisitions, bankruptcy or liquidation, according to Compustat variables "DLDTE" and "DLRSN". I also remove CEOs with incomplete information, missing years and multi-firm CEOs.

## Computation of CEO flow compensation, $flcomp$

The construction of the flow compensation variable, $flcomp$, is as follows. Most of the studies on CEO compensation use the item "TDC1" as the measure of total CEO compensation. This variable includes the value of stock awards, stock option awards and other deferred compensation items that are granted in the current period but vest, i.e., become a flow, in future periods. To construct $flcomp$ I need to subtract those items from "TDC1" and add the value of stock and option awards that vest in a period.

Due to changes in the reporting rules of executive compensation, $flcomp$ is constructed separately for the periods 1993-2006 and 2006-2013. In 2006 the Securities and Exchange Commission (SEC) adopted new disclosure requirements concerning, among other items, CEO compensation. Firms had to comply with the new rules if their fiscal year ended on or after December 15th, 2006. This is why the periods overlap. Some firms' executive compensation reports in 2006 appear in the previous reporting format according to each firm's fiscal year-end. In the following, I separately discuss the construction of $flcomp$

for both reporting formats.

### Computation of $flcomp$ in the new reporting format 2006-2013

For firms in 2006-2013 using the new reporting format, $flcomp$ is defined as:

$$flcomp_{it} = \text{“salary”}_{it} + \text{“bonus”}_{it} + \text{“noneq\_incent”}_{it} + \text{“othcomp”}_{it} + \\ + \text{“defer\_rpt\_as\_comp\_tot”}_{it} + opt\_vest\_val_{it} + \text{“shrs\_vest\_val”}_{it}$$

This definition contains the usual cash payments in the form of salaries and bonuses plus other payments that are included in the definition of “TDC1”. On top of that, I add the value that the CEO realizes from vesting stock awards and vesting option awards in the period. That is, I include the realized value of the past promises of CEO compensation.

The value realized on vesting option awards, $opt\_vest\_val_{it}$, is not provided in Execucomp. Notice that the value at $t$ of the change in a CEO's unexercised exercisable (already vested) options is given by

$$\Delta option\_unex\_exer\_val_{it} = opt\_vest\_val_{it} - \text{“opt\_exer\_val”}$$

where “opt_exer_val” denotes the value realized by the CEO from exercised options. Hence, the value of the vested options is given by

$$opt\_vest\_val_{it} = \Delta option\_unex\_exer\_val_{it} + \text{“opt\_exer\_val”}$$

I follow Coles et al. (2006, 2013) in order to compute $\Delta option\_unex\_exer\_val_{it}$ for firms using the new reporting format in 2006-2013.

Post-2006, Execucomp provides a separate record for each outstanding option tranche (denoted by a different value of “outawdnum”), indicating the number of vested, unvested, and unearned options of each tranche, and their corresponding exercise price and expiration date. For each option tranche $j$ for CEO $i$, I compute the number of options vesting at time $t$ as follows:

$$\Delta option\_unexer\_exer_{jit} = \text{“opts\_unex\_exer”}_{ji,t} \frac{\text{“ajex”}_{it}}{\text{“ajex”}_{i,t-1}} - \text{“opts\_unex\_exer”}_{ji,t-1}$$

Where “ajex” is an adjustment factor for stock splits. If $\Delta option\_unexer\_exer_{jit}$ is negative I set it to zero and assume that all the value from option-vesting arises from exercised options. If the option award has been granted in the current fiscal year I set $\Delta option\_unexer\_exer_{jit} = \text{“opts\_unex\_exer”}_{ji,t}$. This same adjustment is made for observations in 2006, when I do not observe tranch-specific observations. An additional adjustment is made for these observations as explained below.

I use the Black-Scholes formulae to compute the value of the vested options, taking into account the stream of dividends. To compute the Black-Scholes value of options I need the exercise price and expiration date of the option tranche and estimates of the dividend yield, the volatility of the firm stock and the risk-free rate of return.

Execucomp stopped providing the estimate of the stock return volatility, through the variable "bs_volatility", as of 2006. I follow the Execucomp methodology as closely as possible. Accordingly, I (i) use the annualized standard deviation of (log) stock returns estimated over the 60 months prior to the beginning of each fiscal year; (ii) require at least 12 months of returns data; (iii) use mean volatility (across all firms) for that year if 12 months of data are not available; and (iv) winsorize the volatility estimates at the 5th and 95th levels. The Black-Scholes volatility is denoted by $bs\_vol_{it}$.

I also compute estimates of the dividend yield because Execucomp stopped providing this variable, "bs_yield", as of 2006. Following their methodology as closely as possible: I (i) use the average of "divyield" provided by Compustat/CRSP over the current year and the two prior years and (ii) winsorize the values at the 5th and 95th levels. The "divyield" is expressed as a percentage in Execucomp and I divide by 100 to use it in the Black-Scholes formula. I denote the dividend yield by $bs\_divy_{it}$.

I impute the risk-free rate of return as that corresponding to the (rounded) maturity of the options as of fiscal year-end. I obtain the risk-free rate from historical data provided by the Federal Reserve on their website for "Treasury constant maturities" using the "annual" series: (`https://www.federalreserve.gov/datadownload/Build.aspx?rel=H15`). The website provides data for 1, 2, 3, 5, 7, and 10 year Treasury securities. I interpolate the rates to obtain the risk-free rates for 4, 6, 8, and 9 years. If the option maturity is more than 10 years, I use the 10-year rate. The rates are expressed as a percentage and divided by 100 to use them in the Black-Scholes formula. I denote this variable with the name $r_{jit}$.

Thus, using this information, the value of unexercised exercisable option awards in tranche $j$ for CEO $i$ and fiscal year $t$ is given by

$$\Delta option\_unex\_exer\_val_{jit} = \Delta option\_vest_{jit} \times BS(\text{“prccf”}_{it}, bs\_vol_{it}, bs\_divy_{it}, r_{jit})$$

Where $BS()$ denotes the Black-Scholes formula and "prccf" is the stock price at the end of the fiscal year. I set to zero any resulting negative value. The total value of unexercised exercisable options is given by the sum of $\Delta option\_unex\_exer\_val_{jit}$ across all tranches $j$

in each period.

I adjust those observations in 2006 with previously-awarded tranches by subtracting the aggregate value of the unexercised exercisable options in the previous year

$$\begin{aligned} opt\_vest\_val_{it} =& \Delta option\_unex\_exer\_val_{it} + \text{“opt\_exer\_val”} - \\ & - \text{“opt\_unex\_exer\_num”}_{i,t-1} \times BS(\text{“prccf”}_{it}, bs\_vol_{it}, bs\_divy_{it}, r_{jit}) \end{aligned}$$

This variable is not available for executives at their first tenure-year who were not previously an executive of the firm. For these observations I assume $opt\_vest\_val_{it}$ is equal to 0.

### Computation of $flcomp$ in the old reporting format 1993-2006

For the period 1993-2006 the item “shrs_vest_val” is not provided by the database. Thus, following Taylor (2013) I define a variable $shrs\_vest\_val\_approx$ as follows. First, I define $\Delta stock\_unvest\_num_{it}$ as the year-on-year change in the number of unvested shares held by the CEO:

$$\Delta stock\_unvest\_num_{it} = \text{“stock\_unvest\_num”}_{i,t-1} - \text{“stock\_unvest\_num”}_{i,t} \frac{\text{“ajex”}_{i,t}}{\text{“ajex”}_{i,t-1}}$$

where “stock_unvest_num” are the number of unvested shares held by the CEO. This variable is not available for CEOs at their first tenure-year who were not previously an executive of the firm. For these observations I assume “stock_unvest_num”$_{i,t-1}$ is equal to 0. Similarly, this variable cannot be computed for CEOs in their tenure-year 1 in 1992. I drop these CEOs from the sample. The value of shares that vest in each year is approximated by the following variable $shrs\_vest\_val\_approx_{it}$:

$$shrs\_vest\_val\_approx_{it} = \text{“rstkgrnt”}_{it} + \text{“prccf”}_{it} \Delta stock\_unvest\_num_{it}$$

Where “rstkgrnt” is the value of stock awards received by the CEO in the current year. I set to 0 all negative values that arise for $shrs\_vest\_val\_approx_{it}$.

In order to compute the value of vesting options, $opt\_vest\_val_{it}$, I follow the discussions in Taylor (2013), Core and Guay (2002), Edmans et al. (2009), Coles et al. (2006, 2013) to compute the value of unexercised exercisable options. The number of options vesting during the year is given by

$$opt\_vest\_num_{it} = \text{“opt\_unex\_exer\_num”}_{i,t} \frac{\text{“ajex”}_{i,t}}{\text{“ajex”}_{i,t-1}} - \text{“opt\_unex\_exer\_num”}_{i,t-1} + \\ + opt\_exer\_num_t$$

If $opt\_vest\_num_{it}$ is negative I set it to zero. Regarding the old reporting format, firms were required to report tranche level details only for the current year's option grants. That is, I have the number of options granted "numsecur", the exercise price "expric", and the maturity of each tranche of options awarded in the current year "exdate". Firms were not required to report tranche-level details on previously granted options. Instead, they only had to report the intrinsic value and number separately for the portfolio of vested options and the portfolio of unvested options.

For vested options, I calculate the average exercise price based on the realizable value "opt_unex_exer_est_val" and number of vested options "opt_unex_exer_num" in each period $t$:

$$strike\_ex = \text{“prccf”} - \frac{\text{“}opt\_unex\_exer\_est\_val\text{”}}{\text{“opt\_unex\_exer\_num”}}$$

Edmans et al. (2009) recommend the following transformation:

$$strike\_ex = \text{“prccf”} - \frac{(\text{“opt\_unex\_exer\_est\_val”} - (ivnew - \text{“opt\_unex\_unexer\_est\_val”})^+)^+}{\text{“opt\_unex\_exer\_num”} - (numnewop - \text{“opt\_unex\_unexer\_num”})}$$

Where $ivnew$ is the intrinsic value of the newly granted options, $(P - expric)^+ \times numsecur$, and $numnewop$ is the total number of newly granted options. For the option maturities, Core and Guay (2002) recommend assuming a maturity for existing unexercisable options of one year less than the maturity of newly granted options. (Where there are multiple new grants, I take the longest maturity option.) If there were no new grants, I use 8.5 years. The maturity of exercisable options is assumed to be three years less than for unexercisable options. I multiply the maturities of all options by 70 percent to capture the fact that CEOs typically exercise options prior to maturity. If the estimated maturity is negative, I assume a maturity of one day.

Using the estimated maturities and exercise prices of vested options then I can compute their Black-Scholes prices. For that means, I use the estimated volatilities and dividend yields, together with the risk-free returns, as explained above for the new reporting framework.

Finally, the variable $flcomp$ for observations in 1993-2005 and firms using the old reporting format in 2006 is given by

$$flcomp_{it} = \text{“salary”}_{it} + \text{“bonus”}_{it} + \text{“ltip”}_{it} + \text{“othann”}_{it} + \\ + \text{“allothpd”}_{it} + opt\_vest\_val_{it} + shrs\_vest\_val\_approx_{it}.$$

# D Identification

In this section I describe the intuition behind the identification of the individual parameters with the set of 15 target moments.

The first four moments correspond to the distribution of firms' cash flows, $x_{it}$ in the model. First, the empirical mean of cash flows, $\widehat{E}(x_{it})$, allows to identify the parameter $\mu$. The second and third moments are the variance of the change in cash flows, $\widehat{Var}(\Delta x_{it})$ and the covariance $\widehat{Cov}(\Delta x_{it}, \Delta x_{i,t-2})$. Cash flows are first-differenced to remove cross-sectional heterogeneity that is not captured in the model.

The fourth moment is the estimated parameter $\widehat{\phi}$ in the AR(1) panel model $x_{it} = \phi x_{it-1} + \nu_{it}$. In order to estimate $\phi$ consistently I use the Han and Phillips (2010) estimator that deals with the individual cross-sectional heterogeneity. $\widehat{E}(x_{it})$, $\widehat{Var}(\Delta x_{it})$, $\widehat{Cov}(\Delta x_{it}, \Delta x_{i,t-2})$ and $\widehat{\phi}$ allow to identify the parameters $\rho$, $\sigma_\eta$, and $\sigma_\varepsilon$. In particular, for a sufficiently long time series of cash flows, the four moments would be consistently estimated and given by

$$\begin{aligned} &E(x_{it}) = \mu \\ &Var(\Delta x_{it}) = 2\left(\sigma_\eta^2/(1+\rho) + \sigma_\varepsilon^2\right) \\ &Cov(\Delta x_{it}, \Delta x_{it-2}) = -\rho(1-\rho)\sigma_\eta^2/(1+\rho) \\ &\phi = \rho \frac{\sigma_\eta^2/(1+\rho)}{\sigma_\eta^2/(1+\rho) + \sigma_\varepsilon^2} \end{aligned}$$

The next set of moments capture the features of CEO departure frequencies. The moments are computed without considering any *ex-ante* classification criteria. First, I match the turnover rate in the data, i.e. the fraction of CEO-firm observations that feature a turnover event in the sample. This moment can be interpreted as the unconditional probability of CEO departure. Second, I match the CEO departure hazard rates for tenure years $t = \{1, 2, 3\text{-}4, 5\text{-}6, 7\text{-}8, 9\text{-}10\}$.

Figure D.1 illustrates the separate identification of $\kappa$ and $\lambda$. An increase in $\kappa$ and reduction in $\lambda$ both would lead to lower use of dismissals as an incentive device (right panels). This behavior happens because dismissals become more costly (increase in $\kappa$) or

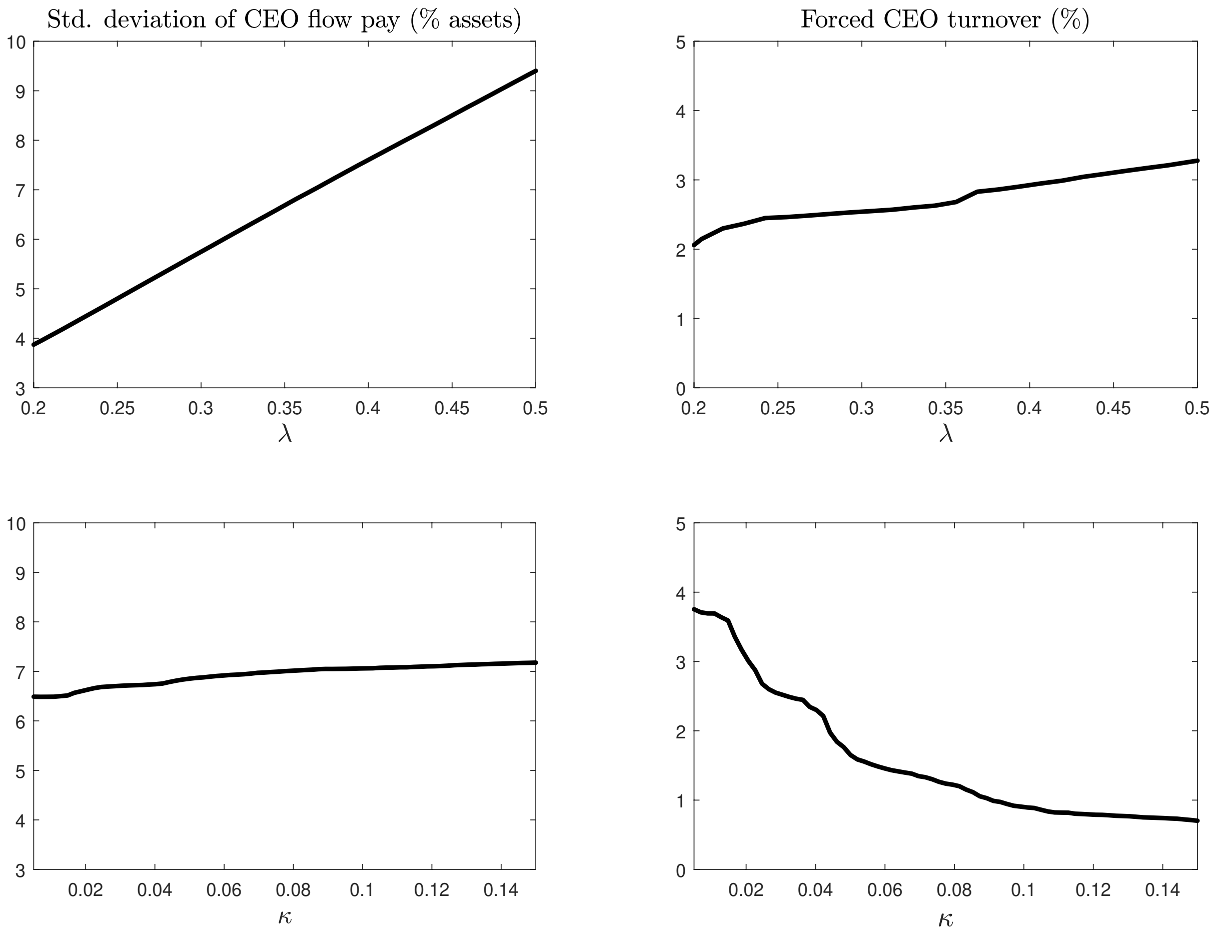


**Figure D.1:** Comparative statics. Identification of $\lambda$ and $\kappa$. Model parameters in this exercise are $\mu = 0.0$, $\rho = 0.0$, $\sigma_\eta = 0.05$, $\sigma_\varepsilon = 0.025$, $\beta = 0.96$, $\beta_c = 0.85$, $\lambda = 0.35$, $\kappa = 0.025$, $\underline{w} = 0.0$, $\alpha_0 = 4$, $\alpha_1 = -0.4$, $T = 50$.

the moral hazard problem becomes less harsh (reduction in $\lambda$). However, both changes would lead to different reactions in the incentive compensation policies (left panels). Higher replacement costs needs of a more intense use of incentive compensation to offset the lower frequency of dismissals. In contrast, less harsh incentive problems need of less intense use of incentive compensation packages.

This set of target moments related to CEO departures provides identification on the replacement cost parameter $\kappa$ and the exogenous departure parameters $\alpha_0$ and $\alpha_1$. $\alpha_0$ and $\alpha_1$ are separately identified by the difference between the hazard rates at tenure-year 1 and the rest of hazards at tenure-years $t > 1$.

I achieve the separate identification of the replacement cost parameter $\kappa$ from the exogenous departure coefficients $\alpha_0$ and $\alpha_1$ as follows. Consider a change in $\alpha_0$ or $\alpha_1$ that increase the frequency of CEO departures in the same manner as a reduction of $\kappa$. These changes are going to have a differential impact on incentives. A reduction in $\kappa$ makes it cheaper for firms to replace CEOs after poor performance. Hence, investors can provide a higher share of incentives through termination threats. This response of investors reduces average wages and the response of wages to performance, i.e., $E(w_{it})$, and $\delta_1$ become smaller.

In contrast, the change in $\alpha_0$ or $\alpha_1$ that increases the exogenous departure frequencies makes deferred compensation $u$ less effective in providing incentives, i.e., the CEO is more likely to leave in any period. Therefore, the CEO must receive higher flow payments and the performance sensitivity of flow compensation must increase to guarantee incentive-compatibility. This effect is the opposite of a reduction of $\kappa$.

Besides, $\alpha_0$ or $\alpha_1$ are separately identified from $\lambda$ since $\alpha_0$ generates parallel shifts in the hazard function, while $\alpha_1$ changes the slope of the departure hazard function. In contrast, changes in $\lambda$ affect the slope at tenure-years one and two and generates parallel shifts in later tenure-years. Specifically, an increase in $\lambda$ increases the threshold dismissal level of deferred compensation, so at early tenure-years, the CEO is more likely to depart. In later tenure-years the CEO is more likely to fall in the dismissal region after a few periods of bad performance.

I depict this behavior in Figure D.2 where I compare the behavior of the CEO departure hazard function for different values of $\lambda$, $\kappa$, $\alpha_0$, and $\alpha_1$. The figure shows how changes in $\lambda$ and $\kappa$ change the slope of the function at early years of CEO tenure, while changing the level of the function at later years. In contrast, changes in $\alpha_0$ affect the level of the function at all tenure years, while changes in $\alpha_1$ also change its slope at all tenure years.

The parameters $\lambda$, moral hazard parameter, and $\underline{w}$, fixed salary, can be identified with the flow compensation moments. The parameter $\lambda$ has a similar impact than $\underline{w}$ on the average flow compensation. Investors must offset a more severe moral hazard problem must through higher levels of flow and deferred compensation. However, $\lambda$ and $\underline{w}$ have a different impact on the variance of flow compensation, since investors tackle the moral hazard problem by increasing the dispersion of current flow payments. Conversely, an increase of $\underline{w}$ generates a reduction in the variance of flow compensation.

An increase in $\beta_c$ has the same directional effect on the size and the variance of flow compensation than a reduction of $\lambda$. However, notice that an increase in $\beta_c$ reduces the size and dispersion of deferred compensation as an increase of $\lambda$. Although $u$ is unobserved, I can use as a proxy the tenure wage slope $\delta_2$. The intuition works as follows. The more impatient the CEO becomes relatively to investors, the earlier the CEO must receive flow compensation. Since I compute the equation for CEOs beyond tenure-year 3, this would imply a reduction in $\delta_2$. On the contrary, when $\lambda$ increases $\delta_2$ should increase since long-tenured CEOs accumulate greater deferred compensation and thus receive higher flow compensation levels.

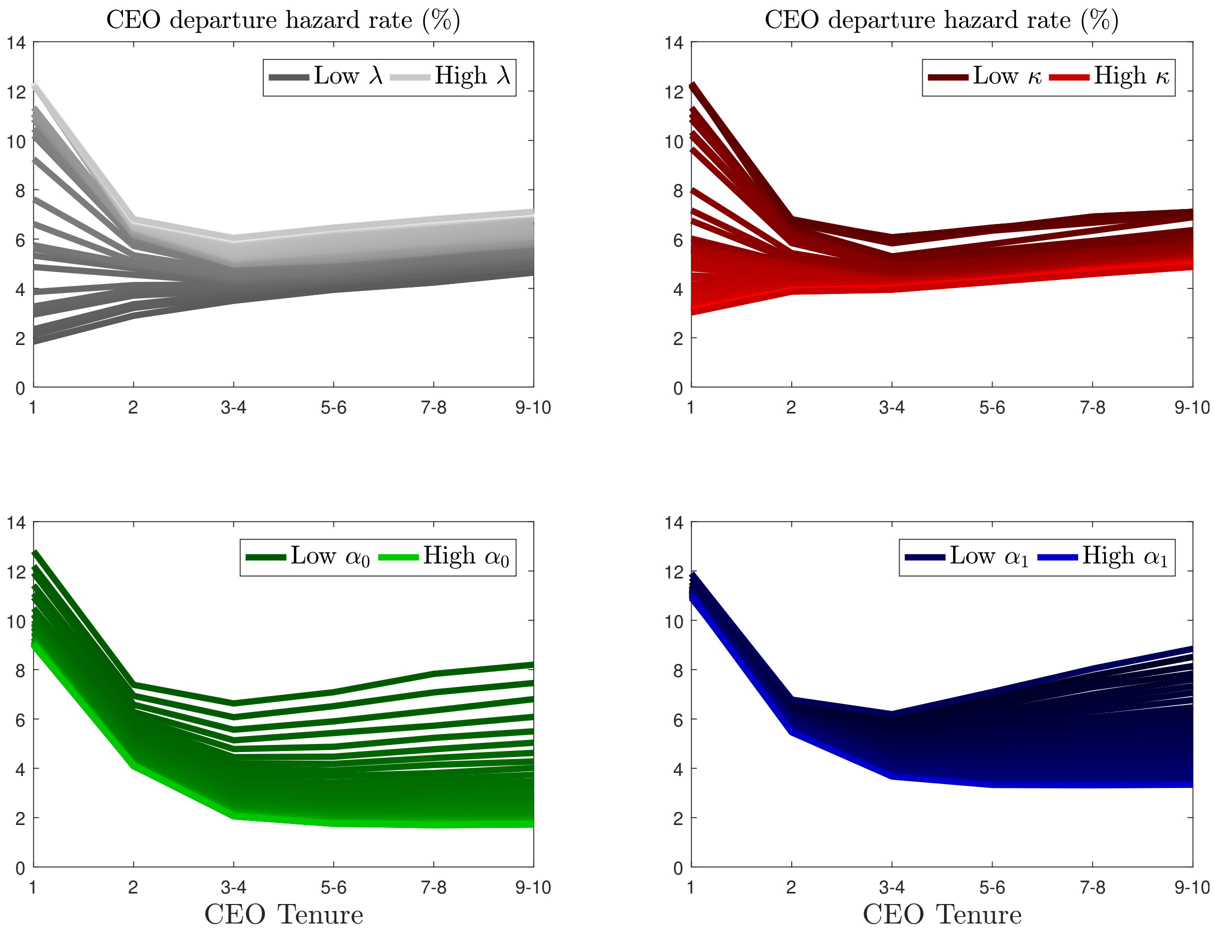


**Figure D.2:** Comparative statics. Identification of $\lambda$, $\kappa$, $\alpha_0$, and $\alpha_1$. Model parameters in this exercise are $\mu = 0.0$, $\rho = 0.0$, $\sigma_\eta = 0.05$, $\sigma_\varepsilon = 0.025$, $\beta = 0.96$, $\beta_c = 0.85$, $\lambda = 0.35$, $\kappa = 0.025$, $\underline{w} = 0.0$, $\alpha_0 = 4$, $\alpha_1 = -0.4$, $T = 50$.

Lastly, $\beta_c$ has a similar impact than $\alpha_0$ and $\alpha_1$ on compensation through the relative impatience of the CEO. However, $\alpha_0$ and $\alpha_1$ have an impact on CEO departure frequencies that is absent when $\beta_c$ changes. A reduction in $\beta_c$ means that deferral decisions must be more sensitive to firm performance and the distribution of $u$ becomes more disperse. The increased dispersion of $u$ implies that bad performers are more likely to be dismissed, but good performers move away from the termination region. Hence, $\beta_c$ can be separately identified.

# E Details on estimation and standard errors

I estimate the model parameters $\theta$ solving problem (6) with an optimization algorithm that combines the simplex method and the simulated annealing procedure (see Goffe et al., 1994 and Press et al., 1996, Chapter 10). This stochastic optimization algorithm avoids the optimization process to get stuck at local minima. I follow a penalty function approach to avoid economically irrelevant evaluations of the parameters.

The construction of the weighting matrix $W$ in problem (6) follows an influence function approach as in Erickson and Whited (2002) and Li et al. (2016). After computing the $J = 15$ target moments I compute the influence function $\psi_{itj}$ for each CEO spell $i$, tenure-year $t$ and moment $j$. Let $\psi_{ij} = (d_{ij1}\psi_{ij1}, ..., d_{ijT_i}\psi_{ijT_i})$ denote the column vector of all the influence functions for moment $j$ and CEO spell $i$. $d_{ijt} \in \{0, 1\}$ is an indicator variable denoting that observation $t$ of spell $i$ enters in the computation of moment $j$. Let $N$ denote the total number of CEO spells in the sample. Let $N_j = \sum_{i=1}^{N} \sum_{t=1}^{T_i} d_{ijt}$ denote the number of observations involving the computation of each moment.

The influence functions can be computed as follows. If moment $j$ is the sample mean of variable $y$, denoted by $\widehat{E}(y)$, the influence function for CEO spell $i$ at tenure-year $t$ is given by

$$\psi_{itj} = y_{it} - \widehat{E}(y)$$

For the variance moments $\widehat{Var}(y)$, the influence function is given by

$$\psi_{itj} = \left(y_{it} - \widehat{E}(y)\right)^2 - \widehat{Var}(y)$$

For the covariance moments $\widehat{Cov}(y, z)$, the influence function is given by

$$\psi_{itj} = (y_{it} - \widehat{E}(y))(z_{it} - \widehat{E}(z)) - \widehat{Cov}(y, z)$$

For the OLS coefficients $\delta$ from regressing $k$ variables $x = (x_1, ..., x_k)$ on variable $y$ their influence functions are given by the $k$-dimensional vector

$$\psi_{itj,j+k} = \left(\frac{1}{N_j} X'X\right)^{-1} x'_{it}(y_{it} - x_{it}\widehat{\delta})$$

where

$$X = \begin{pmatrix} x_{111} & \cdots & x_{11k} \\ \vdots & \ddots & \vdots \\ x_{1T_1 1} & \cdots & x_{1T_1 k} \\ \vdots & \ddots & \vdots \\ x_{CT_c 1} & \cdots & x_{CT_c k} \end{pmatrix}$$

and $N_j$ is the total number of observations entering the computation of $\widehat{\delta}$, and $C$ denotes the number of CEO spells. Besides, CEO departure hazard rates are computed also as OLS coefficients from the regression of CEO tenure dummies on CEO turnover indicators.

The estimation of the model is meant to capture and represent variation at the individual CEO spell level. To remove common year components and cross-sectional heterogeneity I compute the influence functions for means and variances after demeaning each variable first at the year level and then— after adding the full sample means— at the CEO spell level. I do not demean the remaining moments since their computation already takes care of any potential cross-sectional heterogeneity.

The weighting matrix $W$ represents the inverse of the $J \times J$ variance-covariance matrix of the target moments

$$W = \begin{pmatrix} \frac{1}{N_1}\sum_i \psi'_{i1}\psi_{i1} & \cdots & \frac{1}{\sqrt{N_1}\sqrt{N_J}}\sum_i \psi'_{i1}\psi_{iJ} \\ \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{N_J}\sqrt{N_1}}\sum_i \psi'_{iJ}\psi_{i1} & \cdots & \frac{1}{N_J}\sum_i \psi'_{iJ}\psi_{iJ} \end{pmatrix}^{-1}$$

This matrix takes into consideration potential heteroskedasticity and contemporaneous cross-moment correlations within CEO spells. The standard errors of the parameters follow from Pakes and Pollard (1989):

$$\sqrt{N}(\widehat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}\left(0, \left(1 + \frac{1}{S}\right)(GWG')^{-1}\right)$$

Where $G$ is the Jacobian matrix of the moment conditions with size equal to # parameters$\times$#

moments. I estimate $\widehat{G}$ by numerical differentiation at the estimated values of the parameters. Define

$$b_N(\widehat{\theta}) = \hat{m} - \frac{1}{S}\sum_{s=1}^{S} m_s(\widehat{\theta})$$

We have that

$$\sqrt{N} b(\theta_0) \xrightarrow{d} \mathcal{N}\left(0, \left(1 + \frac{1}{S}\right)\Lambda\right)$$

where the asymptotic standard errors computed as

$$\frac{\widehat{\Lambda}}{N} = \begin{pmatrix} \frac{1}{N_1^2}\sum_i \psi'_{i1}\psi_{i1} & \cdots & \frac{1}{N_1 N_J}\sum_i \psi'_{i1}\psi_{iJ} \\ \vdots & \ddots & \cdots \\ \frac{1}{N_J N_1}\sum_i \psi'_{iJ}\psi_{i1} & \cdots & \frac{1}{N_J^2}\sum_i \psi'_{iJ}\psi_{iJ} \end{pmatrix}$$

Therefore, the test for overidentifying restrictions satisfies

$$\frac{S}{1+S} b_N(\theta_0)' \left(\frac{\widehat{\Lambda}}{N}\right)^{-1} b_N(\theta_0) \xrightarrow{d} \chi^2_{\#moments - \#parameters} .$$

# F Full set of target moments and estimation results

This Appendix presents the set of target moments of all samples and the full set of parameter estimations analyzed in the main text. Table F.1 reports the target moments used in the estimations, altogether with the number of firms and CEO spells that appear in each subsample. Table F.2 reports the estimation results for all subsamples, including the estimated CEO dismissal rate and the test for overidentifying restrictions.

**Table F.1:** Full set of target moments

| Parameter | Full sample | Bottom size quartile | | | Top size quartile | | |
|---|---|---|---|---|---|---|---|
| | | Full sample | Board independence | | Full sample | Board independence | |
| | | | ≤ median | > median | | ≤ median | > median |
| Mean of cash flows | -0.640 | -0.912 | -0.463 | -0.514 | -0.672 | -0.202 | -0.004 |
| AR(1) coefficient of cash flows | 0.488 | 0.576 | 0.420 | 0.498 | 0.381 | 0.458 | 0.417 |
| Variance of cash flows | 43.515 | 64.650 | 72.979 | 48.213 | 27.003 | 17.882 | 17.609 |
| Covariance of cash flows | -4.309 | -6.034 | -5.201 | 0.256 | -2.348 | -2.080 | 0.167 |
| Mean CEO flow pay | 0.246 | 0.482 | 0.552 | 0.466 | 0.064 | 0.062 | 0.045 |
| Variance of first diff. CEO flow pay | 0.096 | 0.215 | 0.264 | 0.189 | 0.014 | 0.006 | 0.004 |
| PPS of CEO flow pay | 0.009 | 0.014 | 0.019 | 0.015 | 0.003 | 0.003 | 0.001 |
| Tenure slope of CEO flow pay | 0.051 | 0.084 | -0.095 | 0.183 | 0.018 | 0.022 | 0.014 |
| CEO turnover rate | 11.409 | 12.723 | 13.051 | 11.447 | 10.726 | 9.979 | 9.577 |
| Hazard rate tenure 1 | 9.956 | 13.449 | 16.306 | 13.842 | 8.024 | 6.698 | 8.328 |
| Hazard rate tenure 2 | 9.668 | 13.050 | 12.542 | 9.172 | 7.620 | 7.454 | 5.084 |
| Hazard rate tenure 3-4 | 11.576 | 11.972 | 11.999 | 11.194 | 10.974 | 11.606 | 8.190 |
| Hazard rate tenure 5-6 | 13.810 | 11.925 | 10.655 | 16.220 | 17.556 | 12.936 | 19.137 |
| Hazard rate tenure 7-8 | 13.765 | 11.958 | 9.622 | 14.178 | 15.095 | 14.668 | 15.818 |
| Hazard rate tenure 9-10 | 13.913 | 11.659 | 14.164 | 3.977 | 16.985 | 19.732 | 20.366 |
| Number of firms | 2,235 | 684 | 195 | 111 | 606 | 304 | 160 |
| Number of CEOs | 4,098 | 1,025 | 239 | 137 | 1,024 | 407 | 217 |

The Table shows the target moments for the full sample an all subsamples. The table reports the number of firms and CEO spells in each subsample.

**Table F.2:** Parameters and estimation results

| Parameter | Full sample | Bottom size quartile | | | Top size quartile | | |
|---|---|---|---|---|---|---|---|
| | | Full sample | Board independence | | Full sample | Board independence | |
| | | | ≤ median | > median | | ≤ median | > median |
| $\beta_c$ | 0.800 | 0.776 | 0.728 | 0.946 | 0.827 | 0.881 | 0.802 |
| | (0.009) | (0.022) | (0.027) | (0.030) | (0.039) | (0.036) | (0.031) |
| $\mu$ | -0.635 | -0.910 | -0.439 | -0.457 | -0.727 | -0.234 | 0.210 |
| | (0.042) | (0.376) | (0.315) | (0.673) | (0.076) | (0.090) | (0.199) |
| $\rho$ | 0.534 | 0.600 | 0.310 | 0.833 | 0.448 | 0.493 | 0.752 |
| | (0.035) | (0.063) | (0.122) | (0.233) | (0.097) | (0.148) | (0.067) |
| $\sigma_\eta$ | 5.562 | 6.930 | 6.863 | 5.958 | 4.322 | 3.616 | 3.152 |
| | (0.081) | (0.173) | (0.462) | (0.388) | (0.164) | (0.223) | (0.247) |
| $\sigma_\varepsilon$ | 1.067 | 1.656 | 2.407 | 2.269 | 0.827 | 0.631 | 0.516 |
| | (0.047) | (0.164) | (1.009) | (0.416) | (0.160) | (0.085) | (0.113) |
| $\lambda$ | 24.036 | 27.357 | 19.804 | 30.998 | 7.470 | 10.716 | 9.746 |
| | (1.191) | (1.598) | (8.266) | (7.760) | (2.322) | (2.409) | (2.846) |
| $\underline{w}$ | 0.000 | 0.170 | 0.165 | 0.295 | 0.001 | 0.001 | 0.003 |
| | (0.009) | (0.022) | (0.034) | (0.052) | (0.015) | (0.010) | (0.008) |
| $\kappa$ | 3.417 | 1.686 | 0.491 | 3.575 | 2.240 | 4.624 | 0.567 |
| | (0.959) | (0.668) | (0.351) | (1.106) | (0.469) | (2.613) | (0.130) |
| $\alpha_0$ | 2.543 | 2.778 | 1.814 | 2.277 | 2.780 | 2.815 | 3.545 |
| | (0.089) | (0.253) | (0.193) | (0.199) | (0.099) | (0.154) | (0.165) |
| $\alpha_1$ | -0.244 | -0.277 | 0.572 | 0.006 | -0.406 | -0.410 | -0.552 |
| | (0.039) | (0.079) | (0.085) | (0.019) | (0.059) | (0.087) | (0.087) |
| Simulated firms | 2235 | 684 | 195 | 111 | 606 | 304 | 160 |
| Estimated CEO turnover and CEO dismissal rates | | | | | | | |
| CEO turnover rate (%) | 11.132 | 12.328 | 12.819 | 11.299 | 10.372 | 9.991 | 8.828 |
| | (0.027) | (0.105) | (0.657) | (0.736) | (0.078) | (0.148) | (0.243) |
| CEO dismissal rate (%) | 1.185 | 3.964 | 5.009 | 2.047 | 0.367 | 0.185 | 2.495 |
| | (0.004) | (0.044) | (0.209) | (0.145) | (0.004) | (0.004) | (0.102) |
| Test of overidentifying restrictions | | | | | | | |
| $\chi^2$ | 24.982 | 7.635 | 7.560 | 7.594 | 42.655 | 8.671 | 15.999 |
| p-value | 0.000 | 0.178 | 0.182 | 0.180 | 0.000 | 0.123 | 0.007 |

The Table shows the parameter estimates for the full set of estimations. Robust standard errors appear in parentheses. The bottom part of the table shows the $\chi^2$ statistic and corresponding p-value for the test of overidentifying restrictions, which tests the null hypothesis that the empirical and simulated moments are jointly equal. Moreover, the table reports the estimated CEO dismissal rate with the corresponding Monte Carlo standard errors in each sample.

# G Estimation results in the remaining subsamples

This Appendix presents the set of target moments and parameter estimations for the remaining subsamples of firm size. Table G.1 reports the target moments used in the estimations, altogether with the number of firms and CEO spells that appear in each subsample. Table G.2 reports the estimation results for all, including the estimated CEO dismissal rate and the test for overidentifying restrictions.

**Table G.1:** Full set of target moments

| Parameter | 2nd size quartile | | | 3rd size quartile | | |
|---|---|---|---|---|---|---|
| | Full sample | Board independence | | Full sample | Board independence | |
| | | ≤ median | > median | | ≤ median | > median |
| Mean of cash flows | -0.509 | 0.025 | -0.483 | -0.483 | 0.108 | 0.139 |
| AR(1) coefficient of cash flows | 0.483 | 0.538 | 0.557 | 0.431 | 0.414 | 0.420 |
| Variance of cash flows | 48.025 | 50.467 | 47.044 | 34.718 | 37.760 | 24.829 |
| Covariance of cash flows | -5.412 | -5.573 | -7.486 | -3.368 | -4.854 | -0.932 |
| Mean CEO flow pay | 0.284 | 0.326 | 0.279 | 0.162 | 0.189 | 0.154 |
| Variance of first diff. CEO flow pay | 0.101 | 0.130 | 0.074 | 0.052 | 0.069 | 0.036 |
| PPS of CEO flow pay | 0.009 | 0.009 | 0.009 | 0.006 | 0.008 | 0.008 |
| Tenure slope of CEO flow pay | 0.057 | 0.025 | 0.042 | 0.081 | 0.102 | 0.013 |
| CEO turnover rate | 11.497 | 12.427 | 8.925 | 10.726 | 10.276 | 9.650 |
| Hazard rate tenure 1 | 10.878 | 11.285 | 8.861 | 7.693 | 7.125 | 6.434 |
| Hazard rate tenure 2 | 9.622 | 10.858 | 6.045 | 8.766 | 9.193 | 7.423 |
| Hazard rate tenure 3-4 | 11.648 | 14.165 | 9.011 | 11.878 | 11.307 | 8.199 |
| Hazard rate tenure 5-6 | 12.038 | 12.877 | 11.435 | 13.453 | 12.688 | 17.369 |
| Hazard rate tenure 7-8 | 15.405 | 13.867 | 16.246 | 12.670 | 8.314 | 14.762 |
| Hazard rate tenure 9-10 | 14.240 | 11.377 | 8.613 | 13.219 | 11.869 | 7.773 |
| Number of firms | 753 | 298 | 175 | 749 | 281 | 205 |
| Number of CEOs | 1.024 | 368 | 210 | 1.025 | 325 | 253 |

The Table shows the target moments for the remaining set of subsamples. The table reports the number of firms and CEO spells in each subsample.

**Table G.2:** Parameters and estimation results

| Parameter | 2nd size quartile | | | 3rd size quartile | | |
|---|---|---|---|---|---|---|
| | Full sample | Board independence | | Full sample | Board independence | |
| | | $\leq$ median | $>$ median | | $\leq$ median | $>$ median |
| $\beta_c$ | 0.754 | 0.774 | 0.743 | 0.898 | 0.899 | 0.784 |
| | (0.032) | (0.022) | (0.024) | (0.015) | (0.019) | (0.039) |
| $\mu$ | -0.540 | -0.002 | -0.641 | -0.491 | 0.086 | 0.124 |
| | (0.101) | (0.361) | (0.399) | (0.103) | (0.188) | (0.208) |
| $\rho$ | 0.496 | 0.584 | 0.588 | 0.469 | 0.434 | 0.456 |
| | (0.064) | (0.136) | (0.099) | (0.066) | (0.102) | (0.126) |
| $\sigma_\varepsilon$ | 0.496 | 0.584 | 0.588 | 0.469 | 0.434 | 0.456 |
| | (0.064) | (0.136) | (0.099) | (0.066) | (0.102) | (0.126) |
| | 5.855 | 6.234 | 5.820 | 4.962 | 5.086 | 3.914 |
| | (0.149) | (0.386) | (0.299) | (0.158) | (0.252) | (0.270) |
| $\sigma_\varepsilon$ | 1.095 | 0.997 | 1.401 | 0.786 | 0.944 | 0.780 |
| | (0.090) | (0.099) | (0.124) | (0.140) | (0.091) | (0.122) |
| $\lambda$ | 24.405 | 30.242 | 17.667 | 27.103 | 27.984 | 16.848 |
| | (2.592) | (3.275) | (1.506) | (5.211) | (3.778) | (3.705) |
| $\underline{w}$ | 0.001 | 0.000 | 0.068 | 0.000 | 0.015 | 0.001 |
| | (0.014) | (0.021) | (0.023) | (0.021) | (0.014) | (0.020) |
| $\kappa$ | 3.381 | 4.671 | 1.789 | 5.581 | 4.545 | 6.692 |
| | (0.504) | (0.821) | (0.437) | (2.268) | (2.410) | (4.263) |
| $\alpha_0$ | 2.638 | 2.259 | 3.288 | 2.673 | 2.770 | 2.767 |
| | (0.142) | (0.113) | (0.321) | (0.148) | (0.134) | (0.167) |
| $\alpha_1$ | -0.322 | -0.147 | -0.395 | -0.344 | -0.333 | -0.257 |
| | (0.070) | (0.043) | (0.125) | (0.098) | (0.088) | (0.073) |
| Simulated firms | 753 | 298 | 175 | 749 | 281 | 205 |
| Estimated CEO turnover and CEO dismissal rates | | | | | | |
| CEO turnover rate (%) | 11.413 | 12.169 | 8.638 | 10.607 | 10.035 | 8.700 |
| | (0.075) | (0.220) | (0.278) | (0.068) | (0.186) | (0.243) |
| CEO dismissal rate (%) | 1.330 | 0.905 | 2.247 | 0.489 | 0.828 | 0.272 |
| | (0.012) | (0.021) | (0.097) | (0.004) | (0.021) | (0.010) |
| Test of overidentifying restrictions | | | | | | |
| $\chi^2$ | 2.692 | 1.976 | 5.732 | 4.761 | 2.749 | 12.845 |
| p-value | 0.747 | 0.852 | 0.333 | 0.446 | 0.739 | 0.025 |

The Table shows the parameter estimates for the remaining set of subsamples. Robust standard errors appear in parentheses. The bottom part of the table shows the $\chi^2$ statistic and corresponding p-value for the test of overidentifying restrictions, which tests the null hypothesis that the empirical and simulated moments are jointly equal. Moreover, the table reports the estimated CEO dismissal rate with the corresponding Monte Carlo standard errors in each sample.