What's Behind the Smooth Dividends? Evidence from Structural Estimation

Yufeng Wu*

* University of Illinois at Urbana-Champaign, 470D Wohlers Hall, 1206 South Sixth Street, Champaign, IL, 61820, phone:217-265-0422, e-mail: yufengwu@illinois.edu.

I am deeply indebted to my advisor, Toni Whited, for her continuous support. I thank Heitor Almeida, Guojun Chen, Theodosios Dimopoulos, Ruoyan Huang, Olga Itenberg, Ron Kaniel, Jay Kahn, Boris Nikolov, Robert Parham, Robert Ready, Bill Schwert, Mihail Velikov, Jerry Warner, and all seminar participants at Baruch College, Johns Hopkins University, Indiana University, Washington University in St.Louis, University of South Carolina, University of Illinois, University of Rochester, University of Utah, and University of Washington for helpful discussions and comments. I am very grateful to Dirk Jenter and Charlie Hadlock for sharing with me their data on CEO turnovers.

What's Behind the Smooth Dividends? Evidence from Structural Estimation

Abstract

I study the driving forces behind dividend smoothing by developing a dynamic agency model in which dividends signal the firms' earnings persistence. In equilibrium, managers treat dividends and earnings as informational substitutes, and they smooth dividends relative to earnings to smooth negative news releases and lower their turnover risk. Empirical estimates of the model parameters imply that 36% of observed dividend smoothness among US firms is driven by managers' own career concerns instead of shareholders' preferences. Managers cut investments and adjust external financing policies to accommodate this career concern-based dividend smoothing. These effects destroy firm value by 2.09%.

keywords: Structural Estimation, Dividend Smoothing, Executive Turnover

JEL Classification: G30, G34, G32

1 Introduction

Dividend smoothing is one of the oldest and most puzzling phenomena in corporate finance. On one hand, Miller and Modigliani (1961) show that in a frictionless market, managers cannot add value to a firm by changing the amount or timing of dividend payments. On the other hand, in his seminal work, Lintner (1956) provides survey evidence showing that managers put a high priority on the smoothness of dividends. Linter argues that it is "[a] mix of attitudes and sentiments, pressures and sense of responsibility, standards of fairness and good management performance" that shapes the observed dividend pattern among firms.

A natural question that arises from Linter's argument is what constitutes the "mix of attitudes and sentiment." Is it mostly reflective of the shareholders' preferences? In addition, do managers also have personal interests that induce them to smooth dividends? Although dividend smoothing is widely documented in the literature, little work has been done to disentangle the underlying driving forces behind this phenomenon.

This paper addresses this question by exploring managers' career concerns and dividend smoothing in an information-asymmetric environment. I document that in the data, changes in dividend policy are indeed a strong negative predictor of managerial turnover. Firms that lower their dividends experience on average a-third higher forced executive turnover in the subsequent year. I also build and estimate a dynamic agency model that endogenizes this negative dividend-turnover correlation and show that managers react to it. Having career concerns induces managers to smooth dividends excessively, compared to the level of smoothness that would have been chosen to maximize shareholders' value. This excess dividend smoothing leads to cash hoarding during good times, and it crowds out investment when earnings deteriorate, leading to a 2.09% firm value loss in equilibrium.

The model I consider features an information-asymmetric environment and a team of self-interested managers who face a turnover risk in each period. The managers choose the optimal firm policies to maximize the expected value of their lifetime utility, which is a weighted average of their expected future wage income and the value of their equity stake in the firm. Holding an equity stake aligns managers' incentives with the shareholders', but having career concerns diverges their personal interests. Hence, the managers' optimal choice of firm investment, financing, and payout policies will be different from those that maximize the expected cash flows to the shareholders.

The model yields two channels for dividend smoothing. In the first channel, dividend payments convey information on earnings persistence. Current earnings are deemed to have a higher (lower) persistent component if they are accompanied by dividend increases (cuts). In equilibrium, stock price reacts to this dividend informativeness, and due to the signaling costs, the magnitude of the price reaction to dividend cuts is larger than that to dividend increases. Hence, a stable dividend policy helps to protect the equity value of the firm. This first, signaling channel is frequently mentioned in the dividend smoothing literature, so the inclusion of this mechanism in the model allows me to isolate the career concern-based explanation.

The second channel operates through managers' career concerns, which states that the information conveyed by both earnings and dividends influences decisions on managerial tenure. Therefore, managers treat dividends and earnings as informational substitutes, and this substitutability gives them a separate incentive to smooth dividends. In particular, they will be hesitant to increase dividends as earnings improve, because in such states, they are already far away from the turnover threshold. Thus, further increasing dividends brings them very limited benefit. They will also be extremely reluctant to cut dividends when earnings decline in order to withhold the negative news and keep their turnover risk from increasing. This career concern-based amplification channel is the focus of this paper.

Quantifying the effects of the dividend smoothing channels is difficult in part because firms' turnover decisions and dividend payments are both endogenous. There is no obvious instrumental variable for the managers' career concerns ex-ante at the time when they set the firm policies. In addition, although reduced-form regressions can deliver the directional effects of proxies for career concerns on dividend smoothness, they cannot, by nature, address the extent to which dividend smoothing is accounted for by each potential mechanism.

In this paper, I tackle these empirical challenges by estimating the model via simulated method of moments (SMM) on a set of frequent dividend payers using data for the 1992–2011 period. The estimation results confirm that both types of dividend smoothing are present in the data and have large economic significance. In both the actual and simulated data, the average dismissal rate for top executives increases by roughly one-third following dividend cuts, after controlling for other firm- and executive-level characteristics. Managers choose smoother dividends in order to lower their turnover risk, and this incentive explains approximately 36% of the observed dividend smoothness in the data. Because turnover is only a transfer of wage income from the incumbent to future managers, this type of dividend smoothing is considered excessive from the point of view of the shareholders. Dividends would be markedly more responsive to earnings if set directly by shareholders to maximize firm value.

My estimation provides three further results. First, accounting for the relation between dividend policy and executive turnover is crucial for the model to match moments in the actual data. An alternative model that ignores this relation always fails to generate the small variance of dividends or to reproduce the low responsiveness of dividends to earnings changes. The class of dynamic investment models had been struggling with the dispersions in firm-level payouts. My results show that once I allow the managers' career concerns to directly enter into firms' payout decisions, the fit of the model-generated payouts significantly improves. Second, I perform subsample estimations to revisit the cross-sectional variations in dividend smoothing among firms operating in different information environments. By relying on a structural model, I am able to isolate the correlation between information transparency and dividend smoothness, and establish a negative and robust causal effect. This result is different from evidence found in the literature, where the relation between information and dividend smoothness is at best mixed. I provide explanations on what drives such differences. Lastly, a dynamic agency model provides a natural setting to examine the time series trend in dividend smoothness. I find that over time, managers face increasing turnover risk, which causes a higher degree of dividend smoothing in recent decades.

The observation that firms tend to smooth dividend payments goes back as far as 60 years, with the evidence in Lintner (1956) that managers are primarily concerned with the stability of dividends. They behave as if there were a premium associated with a stable dividend policy. This observation is further confirmed by Fama and Babiak (1968); DeAngelo and DeAngelo (1990), and Brav et al. (2005). In particular, Brav et al. (2005) document that firms may take costly actions to avoid decreasing dividends, such as issuing new equity or even cutting positive net present value projects. This finding contradicts the predictions in a typical Modigliani and Miller world, where dividend changes are among many value-neutral policies that a firm can implement.

Inspired by this observation, subsequent studies use different data and experimental settings to understand why stable dividends are value-enhancing from a firm's perspective. For example, Easterbrook (1984) argues that consistent dividend payments keep a firm in the capital market and motivate efficient public monitoring, which, in turn, raises firm value. In the same spirit, Allen, Bernardo, and Welch (2000) emphasize that ex-post stable high dividends attract more institutional investors, who can process information more efficiently and better discipline the management team. Dewenter and Warther (1998) study the payout policies of *keiretsu* in Japan. Their research also supports the idea that firms smooth dividends to alleviate the cost associated with information asymmetry and to reduce the free cash flow problems. These studies are all based on the implicit assumption that dividends are determined by the shareholders' desire to maximize the stock price. However, in reality, we know that managers, especially top financial executives, exert a significant influence on a firm's payout decisions. They have an extra incentive to smooth dividends if their personal well-being is tied to the firm's dividend stability. Whether this incentive exists and how much effect it has on the smoothness in the data is the main focus of this paper.

Kaplan and Reishus (1990) are among the first to study the implications of dividend policy on managers' wealth. They document that top executives in firms that announce dividend cuts are 50% less likely to be appointed outside directors, and they have a higher probability of losing their outside directorship three years following the dividend cuts. In a more recent study, Parrino, Sias, and Starks (2003) find that firms with dividend cuts or eliminations experience a greater institutional exodus, which reduces the likelihood that a top executive is promoted to CEO internally. These results are all consistent with the idea that dividend instability hurts the top executives' well-being. Given these observations, it is also interesting to examine whether managers react to such incentive by choosing an inter-temporally smoother dividend profile, which is the main research question of this paper.

My paper is most closely related to Lambrecht and Myers (2012, 2014), who are the first to formally model the link between firm earnings and payout using a dynamic agency model. Under certain simplifying assumptions, they derive a closed-form solution for firm-level total payout, which follows Lintner's target adjustment equation. While their paper focuses on the role of cash redistribution as a contracting tool, this paper examines the information content of dividends on firm earnings and managerial turnover. I test the empirical relevance of this career concern-based dividend smoothing channel and present the quantitative as well as the qualitative effects. Mahmudi and Pavlin (2013) also estimate a dynamic model to examine how a firm's payout policy is determined in conjunction with its investment and financing decisions. However, they do not directly tackle the question of why firms smooth payouts, whereas I test and confirm that managerial career concerns are an economically important factor that drives dividend smoothing.

Fudenburg and Tirole (1995) also focus on executives' career concerns. They provide a theoretical framework to assess the influence of career concerns on managers' choice of earnings smoothing. In their model, all reported earnings are paid to the owners period by period as dividends, and hence the smoothness of dividends arises naturally due to earnings smoothing. My paper differs from theirs in two important aspects. First, unlike in their paper where the equilibrium dividend policy is either non-informative or fully revealing, I captures the idea that dividends and earnings signal different aspects of a firm's profitability and hence are informational substitutes. Second, they analyze the smoothing of earnings and dividends against some latent profitability measures instead of focusing on how dividends vary relative to the reported earnings. In the robustness section, I also explore the joint determination of earnings and dividend smoothing. Consistent with Fudenburg and Tirole (1995), I find that managers smooth reported earnings relative to unobservable profit changes. On top of that, they also smooth dividends relative to reported earnings to further delay the information release and mitigate their career concerns. This "two-tier" smoothing behavior generates interesting information dynamics that is absent in the literature.

The remainder of this paper is organized as follows. In Section 2 and 3, I discuss the model and its underlying economic intuitions. In Section 4 , I describe the data. I outline the estimation strategies and report the results in Section 5. Robustness checks are presented in Section 6, and in Section 7, I conclude and indicate some future directions.

2 Model

Models on firm payouts (Allen, Bernardo, and Welch, 2000; Miller and Rock, 1985) are usually based on the implicit assumption that dividend policy reflects shareholders' desire to boost the stock price. However, the recent literature puts increasing emphasis on how managers' self-interest can also shape a firm's financial decisions (Morellec, Nikolov, and Schürhoff, 2012;

Nikolov and Whited, 2014). I follow this literature and build a dynamic model of self-interested managers who set their firms' investment, financing, and payout policies each period to maximize the expected value of their utility. Managers are subject to career concerns, which makes their policies, in general, not the same as the ones that maximize the shareholders' welfare. The model also imbeds an information asymmetry between the inside managers and outside investors. Both the managers and the investors observe the current earnings level, but only the managers know precisely how persistent earnings will be going forward. This information asymmetry gives investors an incentive to extract information out of the announced dividend policy. The dividend informativeness generates endogenous price reactions and managerial turnovers in equilibrium, which is another distinctive feature of the model. The remainder of this section provides more details on the model setup and qualitatively illustrates how managers' utility maximization determines firm-level dividend smoothness. In the following sections, I take the model to data and present the quantitative results.

2.1 The Basic Setup

The backbone of this model is a dynamic investment model with financing frictions. The model is in discrete time and infinite horizon. The timing of events in each period is described in Figure 1.

[Insert Figure 1]

The model focuses on a representative firm that faces decreasing return-to-scale technology and uses capital, K_t, as the only input to generate per period after-tax profit:

$$Y\left(K_t, z_t, s_t\right) = (1-\tau_c) \times e^{z_t} e^{s_t} K_t^{\theta}, \tag{1}$$

in which $\theta < 1$ is the curvature of a firm's production function, and τ_c is the corporate tax rate. z_t represents a shock specific to each firm-management match, which follows an AR(1) process:

$$z_t = \rho_z \times z_{t-1} + \epsilon_{z,t}, \qquad \epsilon_{z,t} \overset{iid}{\sim} N\left(0, \sigma_z^2\right). \tag{2}$$

Note that as in Holmström (1999); Mortensen and Pissarides (1994); Cao and Wang (2013), z_t should be understood as the match quality between the managers and the firm, which is time varying and does not translate directly into some fixed properties of the firm or the person. As Cao and Wang (2013) argue, a high match qualify means that the executive's talent and experience fits well with the size of the firm, its nature of business, strategic direction, and organizational culture in this particular time frame. An manager well-matched with the firm at one point in time may not be well-matched with the same firm at another point in time, due to the change in the above executive or firm-level characteristics.

s_t is a transitory earnings shock, $s_t \overset{iid}{\sim} N\left(0, \sigma_s^2\right)$, which also enters exponentially into the firm's current earnings, but it does not have any implication on future cash flows. At the beginning of each period, the two shocks, z_t and s_t, are realized. Managers observe them separately, and they base the firm's investment, financing, and payout decisions on the realized values. A firm's investment, I_t, is defined as:

$$I_t = K_t - K_{t-1} \times (1 - \delta), \tag{3}$$

where δ is the depreciation rate of physical assets. A firm can either finance investment using its holdings of liquid assets, L_t, or it can go to the capital market and issue new equity. On the other hand, if a firm wants to dispose of idle cash, it can either pay dividends or make repurchases. Dividends, D_t, are subject to a personal tax rate, τ_p, at the time of distribution. Equity issuance and repurchases are associated with a linear-quadratic financing cost. Let $\Lambda(.)$ denote the net cash flow from issuance and repurchases after paying this cost:

$$\Lambda\left(E_t\right) = E_t - \upsilon_1 \times |E_t| - \upsilon_2 \times \tfrac{E_t^2}{K_t}. \tag{4}$$

In equation (4), a positive E_t means that the firm is receiving cash from its investors, while a negative E_t means cash redistributions to the shareholders. $\upsilon_1 \times |E_t|$ and $\upsilon_2 \times \frac{E_t^2}{K_t}$ capture the linear and quadratic components of the financing costs. Empirically, firms pay sizeable fees for investment banking services when they make seasoned equity offerings (Gao and Ritter, 2010) or accelerated share repurchases (Dickinson, Kimmel, and Warfield, 2012). Asquith and Mullins (1986) and Corwin (2003) find that SEOs occur at discounts to market prices and that the discount increases with the size of the equity offering. Related to this idea, a large literature following Vermaelen (1981) documents that firms announce stock repurchases at a premium to the current share price. I summarize the effects of the fees, along with adverse selection in reduced form using equation (4), which implies that the cost of net issuance is monotonically increasing in size and exhibits diseconomies of scale, consistent with the evidence presented in Warusawitharana and Whited (2012).

2.2 Managers' Utility Maximization

In the model, managers are offered compensation contracts consisting of two components: The first component is a fixed wage income per period, contingent on the managers staying with the firm. The second captures the managers' equity stake in the firm. In this paper, I do not discuss the optimality of such a contract. Instead, I take the form of executive compensation in the data and try to infer managers' policy choices based on the structure of their compensation package. Managers in the model are assumed to be risk-neutral. This risk-neutrality assumption captures the idea that the top executives who can influence a firm's policies are usually wealthy individuals, and they have good access to various investment and savings technologies.[1] In each period, managers determine the firm's investment, financing, and payout policies, $\{I_t, E_t, L_t, D_t\}$, to maximize the discounted present value of their utility:

$$U_t = \max_{\{I_t, E_t, L_t, D_t\}} \mathbb{E}\left[\sum_{s\geq t}\left(\prod_{s\geq v\geq t}\beta\left(1-\Phi_v\right)\right)W_t + \kappa_I V_{It} + \kappa_M V_{Mt}\right], \tag{5}$$

[1] The model can also be extended to include risk-averse and habit-persistent agents as in Lambrecht and Myers (2012, 2014). Mathematically, making such an extension is equivalent to introducing a concave transformation on the managers' utility function, which gives them stronger incentives to smooth.

subject to the sources and uses of funds constraint:

$$Y_t + \tau_c \delta K_t + L_{t-1}(1 + r_f - r_f \tau_c) + \Lambda(E_t) \geq I_t + D_t + W_t + L_t. \tag{6}$$

U_t in equation (5) stands for the managers' utility. V_I and V_M represent the intrinsic and the market value of the firm, respectively. The two measures are usually different because investors do not directly observe firm fundamentals. Instead, they form their own forecast of firm fundamentals using available information. Managers care about the intrinsic value of the firm because they have an equity stake in their own firm, the worth of which is tied to the firm's intrinsic value over a long horizon. The market value of the firm is also relevant because managers can inherit the preferences from shareholders who need to trade for liquidity reasons (John and Williams, 1985). κ_I and κ_M captures the weights of firm intrinsic and market value in the managers' utility function. W_t is the managers' wage income, which is modeled as a constant fraction, η, of the firm's steady state asset value[2]. β is the managers' discount rate and Φ_t is a dummy variable indicating forced turnover. Once a manager leaves office, he keeps his equity stake, but forfeits his current, plus the expected value of all future wage income.

2.3 Investors' Information Set and Firm Value

One important friction embedded in the model is the information asymmetry between managers and outside investors. Unlike the managers who directly observe the underlying productivity shocks, the investors only perceive the realized profit, which is jointly determined by the persistent and transitory components. This profit is not a sufficient statistic for predicting the firm's future performance as uncertainty exists regarding the value of each individual shock process. Any additional information that helps to disentangle s_t from z_t improves the shareholders' knowledge of the firm's economic standing and allows them to set more efficient stock prices.

[2] I discuss performance-based wage income in Section 6.7.

In the model, investors are allowed to extract information from all announced firm policies. To make the model solvable and estimable, I focus on the set of time-invariant linear forecasting rules,[3] $\{\gamma_0, \gamma_\pi, \gamma_\Omega, \gamma_\mathcal{F}\}$, based on which the investors predict the value of the persistent profitability component, z_t, as accurately as possible[4]:

$$\hat{z}_t = \gamma_0 + \gamma_\pi \times \pi_t + \gamma_\Omega \times \Omega_t + \gamma_\mathcal{F} \times \mathcal{F}_t \tag{7}$$

$$\{\gamma_0, \gamma_\pi, \gamma_\Omega, \gamma_\mathcal{F}\} = \arg\min \mathbb{E}\,|\hat{z}_t - z_t|,$$

in which $\pi_t = \ln Y_t - \theta \ln K_t - \ln(1-\tau_c)$ is the log of capital- and tax- adjusted firm profit. Ω_t denotes the firm's announced investment, financing, and payout policies. $\mathcal{F}_t$ includes investors' previous forecast, $\hat{z}_{t-1}$, and a noisy signal observed in the current period, $\varphi_t \sim z_t + N(0, \sigma_z^2)$[5]. Given the forecasted profitability processes, the firm's intrinsic and market value, V_I and V_M, can be written recursively as:

$$V_I(K_{t-1}, L_{t-1}, D_{t-1}, z_t, \pi_t) = (1-\tau_p)D_t - E_t - \lambda|\triangle D_t| + \beta\,\mathbb{E}\,V_I(K_t, L_t, D_t, z_{t+1}, \pi_{t+1}) \tag{8}$$

$$V_M(K_{t-1}, L_{t-1}, D_{t-1}, \hat{z}_{t-1}, \pi_t) = (1-\tau_p)D_t - E_t - \lambda|\triangle D_t| + \beta\,\mathbb{E}\,V_M(K_t, L_t, D_t, \hat{z}_t, \pi_{t+1}), \tag{9}$$

where the law of motions for z and $\hat{z}$ are specified in equations (2) and (7). The choice of $\{I_t, E_t, L_t, D_t\}$ that enters into equations (8) and (9) represents the optimal policies that maximize the managers' utility defined in equation (5). This choice is not, in general, the same choice that would be made if the managers were maximizing the expected present value of cash flows to shareholders. For any given state, the firm's equity value is less than it would be in the

[3] This is an ad-hoc assumption imposed to maintain tractability. More rigorously speaking, I should track the investors' entire belief distribution, $P(z)$, as a state variable and specify the conditional distributions of firm policies $P(\Omega|z)$ to allow the investors to update their belief in a Bayesian fashion. This algorithm, however, is computational infeasible given the complexity of the manager's problem.

[4] This forecasting rule is calculated using the Krusell and Smith (1998) algorithm. A detailed discussion on the numerical procedures is in the appendix. In an untabulated robustness check, I introduce quadratic terms and cross-products in the prediction equation. The added terms contribute little to the forecasting accuracy, and they do not alter the quantitative results

[5] In equation (7), I focus on predicting z_t only because s_t is iid across time

absence of the misaligned incentives.

$|\triangle D_t|$ in equation (8) represents the unsigned dividend change from the previous period. A positive λ in equation (8) indicates that dividends are not only associated with a higher tax rate, but also an adjustment cost whenever the prevailing level needs to be altered in future periods. As Brealey, Myers, and Allen (2005) argues in their textbook, setting dividends at \$2 per share is a trivial decision if the last year's dividends were also \$2. However, it can cost substantial managerial time and effort if it entails increasing last year's dividends from \$1.5. Such costly adjustment implies that firms should set dividends to echo the sustainable earnings. No dividend changes should be made if such changes are likely to be reversed in the future. Empirically, Grullon, Michaely, Shlomo, and Thaler (2005) provide direct evidence by documenting that current earnings increases/ declines accompanied by dividend movements in the same direction are less likely to be reversed in the future[6]. The announcement of such dividend movements are also associated with abnormal stock returns after controlling for the changes in other forms of payout (Michaely, Thaler, and Womack, 1995), the firms' investment needs (Ghosh and Woolridge, 1989), and the contemporaneous earnings shocks (Aharony and Swary, 1980). The empirical evidence is consistent with the idea that dividend changes can be used as an effective signal, and the market is actively screening firms based on the information revealed by dividend changes.

2.4 The State-Contingent Turnover Risk

For most dynamic investment models, the managers' turnover rate, $E(\Phi_t)$, is treated as an exogenous parameter and is assumed to be constant across time and states. In this paper, I deviate from this assumption by incorporating state-contingent turnover risk into the model. More specifically, the board pulls the trigger when the intra-period return falls below a certain

[6] Consistent with the prior literature, Grullon et al. (2005) also find that including dividend change does not add to a model's predictive power on future earnings changes. These findings jointly suggest that the informativeness of dividends is more about the persistence of earnings than about the levels.

threshold, consistent with the evidence in Warner, Watts, and Wruck (1988), Weisbach (1988), and Parrino (1997). A low return indicates that the shareholders hold a negative view concerning the managers' fit with the firm. The board of directors is supposed to act in representation of the shareholders. Therefore, they replace the incumbent management team following the price pressure:[7]

$$\Phi_t = \begin{cases} 1, & \dfrac{EV_M(K_t, L_t, D_t, \hat{z}_t, \pi_{t+1})}{V_M(K_{t-1}, L_{t-1}, D_{t-1}, \hat{z}_{t-1}, \pi_t)} \leq \underline{r} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where the numerator denotes the firm's continuation value assuming no turnover takes place, and the denominator is the firm's market price at the beginning of the period. Holding the firm's history constant, equation (7) and (10) jointly determine the factors that affect turnover: The first one is the firm's current profit, which is a function of the realized shocks and is outside the control of the management. The second factor is the firm's announced policy, which could inform the investors of the latent persistence of profitability, and the last factor is a noisy signal observed by the investors, which is orthogonal to their existing information. In Figure 2, I examine how this state-contingent turnover affects managers' decisions.

[Insert Figure 2]

The main intuition that emerges from Figure 2 is that the managers' optimal policy should depend on the convexity of their turnover risk. Let us assume for now that both γ_π and γ_Ω are positive and that the economy has low expected turnover (the economy is in Region III), which implies that the managers of an average firm would face convex turnover risk. This turnover risk profile means that if a firm is hit by a good profitability shock, the expected turnover moves into a even more convex region where its slope with respect to any information release becomes relatively flat. Therefore, the marginal benefit from sending out additional signals via a good

[7] I check how this particular specification fits the data in Section 4 and I discuss alternative ways of modelling turnover in Section 6.

policy change is very limited. On the other hand, when performance deteriorates, the turnover risk increases and it becomes increasingly sensitive to the release of additional negative information, which means if the previous policy change needs to be reversed in the future, the marginal cost could be substantial. Trading off such costs versus benefits implies that the managers have an incentive to set their policies insensitive to earnings.

Now, let us turn to an economy with moderate turnover (the economy is in Region II) where the good and bad information have almost symmetric marginal effects. This symmetry implies that managers are neutral about how strongly earnings and other firm policies should co-move. For any policy change that gets quickly reversed in the future, the net effect on the managers' expected turnover is almost zero. Lastly, I examine a case where the expected turnover rate stays at a high level, and the turnover profile is concave for a representative manager (the economy is in Region I). In this case, the mechanism described previously works in an opposite direction: when earnings improve, managers have a strong incentive to signal with good policy changes. They anticipate large marginal effect from such policy changes which could potentially move them into a "safe" region. At the mean time, they are not afraid of reversing such policies in the future when the standing of the firm worsens as they expect very high turnover risk regardless of their policy choices. Which of these above predictions corresponds to the situation in reality is, of course, an empirical question that I will rely on the data to tell.

If the board decide to retain the management, then the firm directly enters into the next period with its profitability $\{z_t\}$ and $\{s_t\}$ following the law of motion described in equation (2). Otherwise, if the board decide to overturn their management under the criteria described in equation (10), they re-enter the labor market to search for successors at a cost c. The match-specific profitability for any new firm-management pair follows the unconditional distribution: $z_{new} \sim N(-c, \frac{\sigma_z^2}{1-\rho_z^2})$, which implies having executive turnover, on one hand, allows the firm to eliminate unproductive matches. On the other hand, it also disrupts the firm's normal operations and entails an opportunity cost, c.[8] For each pair of $\{c; K_t, L_t\}$, the board of directors has a unique choice of $\underline{r}$, which maximizes the firm's market value.[9]

3 Equilibrium Characterization

In this section, I discuss the solution to the baseline model described in Section 2. The model can be condensed into a two-sided decision-making problem. An equilibrium is characterized by the following two incentive compatibility conditions: First, given the shareholders' forecasting decision, $\{\gamma_0, \gamma_\pi, \gamma_\Omega, \gamma_\mathcal{F}\}$, managers set firm investment, financing, and payout policies in each period to maximize their expected utility. The second condition states that knowing the managers' decision-making process, $\{I_t, E_t, L_t, D_t\}$, investors choose the optimal forecasting rule in order to make the best possible predictions of the underlying profitability processes. When both conditions are satisfied, no party has an incentive to deviate from their current strategies. Hence, an equilibrium is achieved.

Using the equilibrium model solution, I can examine the consequences of dividend changes on the managers' personal well-being. The results illustrate why it is utility-enhancing for the managers to choose a smooth dividend path. Table 1 summarizes the parameters and lists the values used for this exercise.

[Insert Table 1]

To see how managers are hurt by the announcement of dividend reductions, I simulate 100,000 hypothetical firms. I sort out the firms that cut dividends in year 1 and track their economic conditions for the subsequent ten periods. I also create a matched sample by choosing a set of non-dividend-cutting firms who have on average the same year-1 reported earnings as

[8] This section deals with a case where the board and outside shareholders share the same information about firm future profitability, and they perceive the same cost for firing a top executive. In Section 6, I extend the model by 1) adding an information asymmetry between the board and the outsiders, and 2) allowing the board of directors to bear a personal cost from executive turnover, which can be understood as the disutility of losing a golf partner in Taylor (2010).

[9] This result can be easily derived following the strict monotonicity of a firm's market value in $\hat{z}$.

in the dividend-cutting sample and use them as controls.

Figure 3 shows that dividend decreases are typically associated with well below average contemporary earnings. When firms reduce dividends, it does not directly signal future decreases in earnings. Instead, their actions imply that large earnings shocks have already been realized and will have persistent effects on the firms' future performance. Consequently, such firms experience slower productivity reversals, and it usually takes longer for their earnings to converge back to the steady state. These outcomes are consistent with the empirical evidence in Grullon et al. (2005).

[Insert Figure 3]

Moreover, Figure 3 also suggests that dividend changes influence the managers' turnover risk. Firms that decrease dividends experience on average one-third higher rates of forced turnover, compared with firms in the matching sample who report similar earnings but manage to increase or maintain their dividend levels. I repeat the above exercise for the sample of dividend-increasing firms and find similar qualitative effects. However, quantitatively, the magnitudes of the effects are much weaker. This is because the parameters in Table 1 predict an average turnover rate of approximately 2.6% per year (consistent with the actual data), which corresponds to the low and strongly convex turnover profile described by Region III in Figure 2. This convexity implies that the benefit from raising dividends when earnings increases is very limited, while announcing a dividend cut accompanied by earnings deteriorations may have a larger negative impact on the expected turnover.

In anticipation of these effects, managers will be reluctant to raise dividends when earnings improve as well as to cut dividends when earnings decline, leading to a low responsiveness of dividends to earnings. Figure 4 illustrates this result graphically, in which I consider two model specifications: the baseline model and an alternative case where I assume that the shareholders extract information from the dividends, but the managers ignore this effect on their tenure.

Instead, they believe that their turnover probability is constant across time and states. This assumption brings me back to the first-best solution, in which the managers behave as if their personal interests are fully aligned with the best interests of the shareholders when setting the optimal policies.

[Insert Figure 4]

Figure 4 shows that when the managers anticipate the state-contingent turnover risk, they become more conservative in setting the rate of payments, which lowers the level of dividends. At the same time, dividends are markedly less responsive to earnings changes, reflected by a flatter slope of the dividend policy. This difference in slopes captures the amount of dividend smoothing that stems entirely from managers' career concerns, which is not desirable from the shareholders' point of view.

Anticipating a state-contingent turnover risk also influences the managers' other policy choices. More specifically, they will hoard cash instead of paying out dividends in cash-rich states, which is costly for the firm because interest is taxed. They will also avoid cutting dividends in an attempt to withhold bad information from investors in a low cash-flow state, which may require issuing new equity or cutting investments. Figure 4 shows that managers choose lower and stickier dividends, issue slightly more equity, and cut investment. These differences make the equilibrium firm value lower than in the alternative case where firm policies do not reflect executives' career concerns.

4 Data

This section offers a brief discussion of the data sets used to quantify the model. The data come from four sources: firm fundamentals come from Compustat; executive compensation data are from ExecuComp; dividend announcement dates and returns are from the CRSP daily file; and the top executive turnovers are from a hand-collected dataset based on *Businessweek*, *Equilar*, and *The Wall Street Journal*.

4.1 Sample Construction

To construct the sample, I start with all non-financial and non-utility firms in the merged CRSP and Compustat database from 1992 to 2011. Analyzing firms' dividend smoothing behavior requires that they provide sufficient dividend payment records, so that a reliable measure on the smoothness can be calculated. Therefore, I restrict the sample to the set of frequent dividend-paying firms following three steps. In the first step, I remove all observations before a firm announces its first dividend and after it makes its last dividend payment. In the second step, I divide the sample into 11 overlapping 10-year sub-periods. For any given 10-year subsample, I only retain the firms that have made at least six positive dividend payments. Any firm with consecutive zero payments is dropped. Those firms are likely to differ systematically from the firms with consistent positive dividends and, therefore, I do not consider zero payments as a special form of smoothed dividends. In the third step, I compare the annualized split-adjusted dividend per share from Compustat and CRSP. I drop those observations where the reported data from the two sources are significantly different (the difference being larger than 10%).

Next, I calculate the magnitude of the dividend price effect by focusing on the observations where the changes in split-adjusted dividend per share are larger than 10%. I obtain the dates for the dividend change announcements from the CRSP daily event file. I check whether these firms make any earnings disclosure in a 10-day window prior to the dividend announcements and exclude the observations where the two types of events overlap. I calculate the three-day cumulative abnormal return (CAR) around the dividend changes and use it to quantify the price effect. The CAR is slightly below 1% and insignificant for dividend increases, and it averages -3% for dividend cuts. In terms of magnitude, these results lies in proximity to earlier studies (Aharony and Swary, 1980; Nissim and Ziv, 2001).

4.2 Executive Turnover

ExecuComp tracks top executives' compensation starting from 1992 onwards, where I retrieve data on the five highest paid managers' total annual compensation, their percentage of stock holdings, and the percentage of non-vested versus vested stock options. However, ExecuComp is not a good source for the turnover data for two reasons: 1) It does not always report the date when an executive leaves office and 2) the reason for departure indicated by the dataset is often vague and inaccurate. To overcome these issues, I hand-collect data from *Businessweek*, *Equilar*, and *The Wall Street Journal*. I define the year of "turnover" as the one in which the firm announces the departure of a top executive. Following Warner, Watts, and Wruck (1988); Parrino, Sias, and Starks (2003); Jenter and Kanaan (2015), I classify a turnover as "forced" if a manager leaves a firm and does not find another executive position within the next year, or if a manager is reported to have retired before the age of 60. I also do a Google search to supplement the data. If any reliable source points out that the turnover is performance-based, then I interpret it as "forced." If, on the other hand, the turnover is due to health issues, I classify it as "voluntary". I discuss the limitations of this algorithm in Section 6. The final sample consists of 10,827 distinct firm-executive pairs and 11,626 firm-year observations from 1992 to 2011. The summary statistics are reported in Table 2.

[Insert Table 2]

4.3 Dividend Smoothness

Following Leary and Michaely (2011), I measure a firm's dividend smoothness using the speed of adjustment (SOA), which equals the estimated coefficient β in the following regression:

$$D_{i,t} - D_{i,t-1} = \alpha + \beta \times (TPR_i \times E_{i,t} - D_{i,t-1}) + \epsilon_{i,t}, \tag{11}$$

in which $D_{i,t}$ and $E_{i,t}$ refer to firm i's dividend and earnings per share, respectively, at time t after adjusting for stock splits. TPR_i is the firm's target dividend payout ratio, which is defined as the median dividend-to-earnings ratio for firm i over the 10-year window. In a hypothetical case where a firm always lets its dividends fluctuate proportionately with earnings, β will have an estimated value of 1; if, on the other extreme, a firm keeps its dividend per share constant regardless of its earnings changes, then β will take the value of 0. In reality, a firm's dividend adjustment usually lies in between these two extremes, with a lower SOA implying that the dividends are smoother and less responsive to earnings changes. In Figure 5, I plots the time-series changes for dividend smoothing over the past 25 years.

[Insert Figure 5]

Figure 5 shows that the SOA of dividends is consistently around 0.2, indicating that shocks to firms' earnings do not translate into proportional changes in dividends. Over time, the SOA of dividends has decreased slightly,[10] while the level of dividend per share has increased by roughly 50%. This evidence suggests that the frequent dividend payers have been increasing their distributions over time, and they distribute in an increasingly "smoother" fashion.

5 Results

In this section , I take the model to the data and present the quantitative results. Based on the results, I perform two counterfactual exercises to quantify the amount of career concern-based dividend smoothing and explore its relation with the information opaqueness and corporate governance.

5.1 Identification

I estimate most parameters using simulated method of moments (SMM), the objective of which is to pick the set of parameters that make the simulated data track the actual data as closely as possible. For the rest of the parameters, I calculate their values separately outside of the

[10] This finding has also been documented by Skinner (2008) and Leary and Michaely (2011).

model. For example, I set the risk-free rate, r_f, equal to the average real three-month Treasury bill rate. I set the dividend income tax, τ_p, equal to the average tax disadvantage of personal income relative to capital gains, which is approximately 15% over my sample period. I set the corporate tax rate, τ_c, to 35%. In addition, I calibrate κ_I to 1.3% using the sum of executives' stock-based compensations, and I set the per-period wage, η, to 0.10% of the steady-state firm value. These parameter choices match the average levels of executive compensation reported in ExecuComp. I estimate the value of the dividend adjustment cost parameter λ by equating the endogenous model-predicted dividend announcement return with the five-day CAR surrounding dividend cuts in the actual data. With $\lambda = 1.37\%$, the model generates -3% and 0.41% abnormal returns around dividend cuts and increases, respectively, consistent with what is found in the actual data.[11] Note that I do not estimate λ together with the other 9 parameters (discussed below) via SMM because the identification for λ comes from a five-day event window, while the identifications for the other parameters come from data at annual frequencies, which makes it difficult to compute their relative weights in an optimal weighting matrix.

I estimate the remaining 9 parameters $\{\rho_z, \sigma_z, \sigma_s, \theta, \nu_1, \nu_2, \delta, \kappa_M, c\}$[12] within the model by matching 17 moments. The success of this strategy depends critically on choosing the moments that are sensitive to variations of underlying structural parameters. On the other hand, I avoid "cherry-picking" by focusing on the moments that reflect important characteristics of the data.

The first two moments correspond to the two coefficients, $\{\beta_k, \beta_y\}$, in the following regression:

$$\ln(Y_{i,t}) = \beta_Y \times \ln(Y_{i,t-1}) + \beta_K \times \ln(K_{i,t}) - \beta_Y \times \beta_K \times \ln(K_{i,t-1}) + \epsilon_{i,t}, \quad (12)$$

[11] I only target the five-day CAR surrounding dividend cuts because the price reaction at dividend increases is statistically insignificant. If I introduce asymmetric adjustment costs, the CAR at dividend cuts and increases could both be matched precisely. Since the price reaction to dividend increases is relatively small and statistically insignificant in the data, this alternative estimation strategy yields qualitatively very similar results.

[12] ρ_z and σ_z are the persistence and standard deviation of a firm's match-specific shock, respectively; σ_s is the standard deviation of the transitory shock; θ is the curvature of a firm's production function; δ is the depreciation rate of physical capital; $\{\nu_1, \nu_2\}$ represents the linear-quadratic cost for net equity issuance; κ_M measures to what extent managers care about the firm's market value; and c captures the opportunity cost for executive turnover.

where $Y_{i,t}$ is a firm's operating income and $K_{i,t}$ denotes the stock of physical capital. As argued by Cooper and Haltiwanger (2006), equation (12) can be derived as an auxiliary equation from the firm's production function equation (1) and the profitability shock processes.[13] Hence, these moments are sensitive to the underlying parameter changes, and they map monotonically into the parameters of interest. When estimating equation (12), I focus on the first-order difference to deal with firm-fixed effects. I use twice-lagged profit, as well as lagged and twice-lagged capital stock as instruments. I impose a complete set of year dummies to absorb the time series heterogeneity in the data.

The next four moments are the standard deviations and AR(1) coefficients of a firm's investment and operating income. These four moments help to identify the dispersions of the shock processes. Keeping all else constant, increases in both σ_s and σ_z will increase the variance of the investment and operating income while only σ_s has a dampening effect on the estimated AR(1) coefficients. The third set of moments includes the mean of investment, which is used to determine the depreciation rate, δ, as well as the mean and variance of net equity issuance, which are used to pin down the fixed and quadratic equity issuance costs $\{\nu_1, \nu_2\}$. I then add the frequency of turnover and the correlation between turnover and earnings to help identify the opportunity cost of firing, c. Lastly, I include the mean and standard deviation of the market-to-book ratio, the mean and standard deviation of the dividend payments, and the SOA of dividends. These are the "catch-all" moments in the model. However, the market-to-book related moments are most sensitive to variations in κ_M, which measures to what extent managers care about the market value of the firm.

5.2 Main Results

Panel A of Table 3 presents the moment conditions. The results show that the model provides a good overall fit to the data. Only two simulated moments, the standard deviation of operating

[13] A step-by-step derivation is provided in the Appendix.

income and the mean market-to-book ratio, are statistically different from the corresponding actual moments at the 10% level. Those two differences, although statistically significant, are not large in terms of the economic magnitude. An over-identification test fails to reject the model with a P-value of 0.195.

[Insert Table 3]

Panel B of Table 3 reports the structural parameter estimates. On the real side, a firm's production function has substantial curvature and the productivity shock, $\{z_t\}$, has a moderate degree of persistence. On the financial side, the costs for net issuance average 8%, which is roughly the sum of gross spread and percentage discount in seasoned equity offerings (Gao and Ritter, 2010).

The opportunity cost from turnover, c, is positive and significant, suggesting that turnover disrupts a firm's operations and induces it to produce at below-capacity levels for subsequent periods. Having a significant turnover cost implies that the firm will keep its incumbent managers for most of the times. Forced turnover will be triggered only infrequently when there is substantial bad information conveyed by earnings and dividends.

I run the following logit regression on both the actual and simulated turnover data to examine these predictions:

$$Prob(\Phi_t = 1) = F(\beta_0 + \beta_Y \times Y_t + \beta_D \times \lambda \triangle D_t + \beta_X \times X_t + \epsilon_t), \tag{13}$$

where F is the cumulative distribution function (CDF) of logistic distribution, Y_t is a measure for firm profitability, and $\triangle D_t$ is the change in dividend payments from the previous period. When the regression is run on the actual data, X_t includes the commonly-used performance and governance control variables in the executive turnover literature. When the regression is run on the simulated data, X_t consists of other announced firm policies.

Table 4 suggests that the model consistently predicts a negative earnings-turnover correlation. After controlling for earnings, dividend changes still have a strong predictive power for managerial turnover. For firms whose current earnings are around the median, reducing dividend per share by 25% will increase the expected rate of turnover by 10.4%, while for firms whose earnings are around the lower 10th percentile, the increase in expected turnover will be one third. These results are economically large, and they remain quantitatively very similar even after I add additional controls or include higher degree polynomials of earnings. The estimated regression coefficients on the simulated data also closely track their counterparts on the actual data. This finding, on one hand, serves as an external validation of the model: Although I do not directly include these coefficients in my moment matching process, the predictions arising from the model naturally parallels the marginal effects of earnings and dividends on turnover in reality. One the other hand, the results can also be used to illuminate why we observe more frequent turnovers among dividend-cutting firms. This result is not mechanical due to mismeasurement in earnings. Instead, it is because dividends reveal important information on how likely the current earnings will be repeated in the future. Therefore, the board members should base their executive turnover decisions on the dividend informativeness.

[Insert Table 4]

5.3 Counterfactuals

This section contains two counterfactual exercises. In the first case, I re-estimate an alternative model specification by turning off the effect of dividends on turnovers. Running a horse race between the baseline model and this nested alternative model highlights the importance of incorporating career concern-based dividend smoothing in order to explain the observed data patterns. In the second case, I re-simulate data under different degrees of dividend price effects and managerial career concerns to quantify the sensitivity of firm payout policy to such informational and agency frictions.

In Table 5, I report the estimation results for a model in which the executive turnover rate equals 2.63% at all times and in all states. The results show that the overall model fit becomes significantly worse under this constant turnover specification. In particular, the model is not able to match the autocorrelation of operating income, investment, or the mean of market-to-book ratio. The model also fails to predict the level, and the low speed of adjustment for dividends. This is because setting turnover risk to a constant brings me back to a standard dynamic investment model with differential discount rates by the managers and shareholders, but with no agency career concerns. In this case, the price effect of dividends alone is not strong enough to induce sufficient smoothing. Note that the cost of executive turnover, c, also becomes smaller in magnitude and is statistically insignificant. This is because the identification of this cost parameter mainly comes from the frequency of executive turnovers and the correlation between executive turnover and firm performance. Once the turnover rate is fixed exogenously, no other moments can effectively pinpoint the value of c. The J-test result shows that the simulated moments under this alternative specification are significantly different from those on the actual data, and the model is rejected at lower than the 1% level.

[Insert Table 5]

An alternative way for the model to match the observed degree of dividend smoothing without relying on managerial career concerns is to put higher penalties on dividend cuts. In this way, decreasing dividends will further depress the stock price and negatively affect the returns to shareholders who need to trade their shares today for liquidity reasons. Having anticipated this effect, shareholders will have a stronger preference to smooth dividends ex ante, while managers will set payout policies to echo such preferences. To examine how this alternative approach works, I first reset the cost parameter λ. Larger λ means a higher real cost associated with dividend changes, and it endogenously generates a stronger signaling effect. As a result, an increase in λ leads to both larger CARs around dividend announcements and inter-temporally, smoother dividends. I relax the constraint that the model matches the observed dividend announcement effect in the data and let λ vary freely to generate the desired degree of smoothness. Second, I adjust the weight matrix in my estimation. Instead of using the optimal weighting matrix [14], I alter it by increasing the weights corresponding to the dividend level and smoothness measures so that these moments are more precisely matched in the outcome.

The results in Table 6 confirm that this approach leads to closer matches on the dividend-related moments. The results are not surprising because I give such moments high weights. Note that with the added degree of freedom, the overall fit of the model does not improve significantly. This is because I am twisting the optimal weighting matrix by giving the less precisely estimated moments higher weights. As a result, some other moments receive relatively lighter weights in the estimation process and end up being further away from the targets. However, those moments tend to have smaller standard deviations, which means the difference for the actual versus simulated data, scaled by the standard deviations, could be large.

[Insert Table 6]

Panel B in Table 6 shows that in order for the model to match the smoothness of dividends, the estimated equity issuance cost needs to be as large as 18%. In addition, the negative return to dividend cuts will have to be more than -6%. These values are different from the baseline estimates because in the original setting, a turnover risk induces the managers to view dividends and earnings as informational substitutes. Therefore, managers become most unwilling to increase dividends precisely when earnings turn good, due to the fact that the marginal benefit from additional good information releases is the smallest in such states. Managers also become extremely reluctant to cut dividends when earnings deteriorate in order not to reveal the persistent poor economic prospects of the firm. This substitutability naturally maps into a lower responsiveness of dividends to earnings. This effect, however, can not be duplicated by simply pushing up the cost of dividend adjustments. In the counterfactual case, I need to impose a

[14] The details on constructing the optimal weighting matrix are included in the appendix.

huge cost on external financing in the model and more than double the dividend announcement return in order to generate enough penalty for volatile dividends. The magnitudes of these frictions suggest that these model estimates are not reasonable.

To quantify to what extent managers' career concerns determine the firm-level dividend smoothness, I re-simulate data under several scenarios: 25%, 50%, and 100% decreases in $\frac{W_t}{(\kappa_I+\kappa_M)}$, which captures the relative weight of the fixed wage versus shareholders' welfare in the managers' utility function, and 25%, 50%, and 100% decreases in γ_Ω, which controls, in equilibrium, how informative dividends are and to what extent they influence the market prices of shares.

[Insert Table 7]

Table 7 shows that the smoothness of dividends (measured by 1-SOA) decreases gradually when either $\frac{W_t}{(\kappa_I+\kappa_M)}$ or γ_Ω decreases. The SOA of dividends hits 0.48 when the career concern-related channel is completely shut off, and it further rises to 0.97 when the stock price effect is also eliminated. The speed of adjustment being close to unity indicates that shocks to a firm's cash flows are almost proportionately reflected by the firm's dividend policy. Comparing rows (1), (4), and (7) in Panel A, I come to the conclusion that 63.7% of the observed dividend smoothness is related to the shareholders' signaling incentives, while the remaining 36.3% is driven by managers' career concerns. Though earnings and dividends should co-move positively due to the sources and uses of funds constraint, they act as substitutes in information production. This substitutability leads the managers to choose a lower responsiveness of dividends.

Moreover, since turnover is only a transfer of wage income from the incumbent to future managers, this type of dividend smoothing is not desirable from the shareholders' point of view. Panel B confirms this idea by showing that having career concern-based dividend smoothing reduces the equilibrium market-to-book ratio by 2.09%. This value reduction comes from two sources. First, managers choose on average less efficient investment and financing policies to maintain dividend stability. For example, they hoard cash instead of raising dividends in cash-rich states, which leads to higher tax payments on interests. They also avoid dividend cuts in low cash-flow states, which incurs the cost of issuing new equity or cutting investments. Second, dividend smoothing allows managers to withhold private information, which implies that the realized turnover decisions will come from a more restricted information set, and they are less efficient compared with the case where dividends are fully revealing.

The counterfactual results in Table 7 allows me to answer what is behind the smooth dividends—it is a combination of 60% of shareholders' preference and 40% of managers' self-interests. The first 60% of dividend smoothing benefit the shareholders by making their share values stable over time. However, having extra smoothness beyond this level starts to destroy shareholders' welfare by restricting information and distorting firm policies. To my knowledge, this is the first attempt to disentangle how different stakeholders' incentives impact dividend smoothing and discuss separately their welfare implications.

5.4 Subsample Estimations

In this section, I confront my model with the cross-sectional and time-series dispersions of dividend smoothing. Although dividend smoothing is prevalent among the sample of frequent payers, there is a wide heterogeneity in terms of the extent to which firms smooth dividends. Such heterogeneity provides a natural setting to test my model and examine whether it can be used to generate predictions consistent with the data.

I first split the sample using two-digit SIC codes and re-estimate the models based on the 17 industries with over 300 observations. In Panel A of Figure 6, I report the simulated versus actual industry-level dividend payments under the baseline model. In Panel B, I report the industry estimation results for an alternative model assuming constant turnover risk. Figure 6 suggests that, on the one hand, the baseline model slightly undershoots the average rate of dividends, but overall, both models do a reasonable job in matching the cross-industry dispersion in dividend levels. On the other hand, the constant turnover model systematically overshoots the variance of dividends. It predicts more volatile dividends than the actual data suggests. The constant turnover risk model performs even worse when it comes to the SOA of dividends. In particular, it more than doubles the responsiveness of dividends to earnings, and it is not able to preserve the rank ordering of dividend smoothness across industries. These results sharply contrast the performance of the baseline model, highlighting the importance to account for managers' career concerns not only to help explain the average low SOA of dividends, but also to predict the wide cross-sectional dispersion of dividend smoothness among subsamples of firms.

[Insert Figure 6]

To further explore the cross-sectional heterogeneity of dividend smoothing, I sort firms based on their executives' reputation and compensation structure. An executives' reputation is measured using the firm's average earnings decile within the industry from year -2 to the year when the executive first joins the firm, and his compensation structure is measured by the ratio of total cash- versus stock-based compensations awarded to him during his stay with the firm. I re-estimate the model and redo the calculations in Table 7 using parameter estimates from each subsample. The (untabulated) results suggest that the career concern-based dividend smoothing is 16% lower among firms run by more reputable executives. This is because the good reputation puts these executives in better positions to "absorb" the negative news released by earnings and dividends instead of smoothing it. The amount of career concern-based smoothing is also 11% lower when top executives receive a higher fraction of stock-based compensation. High stock-based compensation induces the executives to behave similarly to the equity holders. Hence, they will choose a lower degree of career concern-based dividend smoothing, which is value-destroying from the shareholders' perspective.

In Table 8, I present the subsample results by splitting firms based on time. The "Early" subsample contains all firm-year observations prior to 2002. The "Late" subsample contains all observations from 2003 onwards. The sample is split in this way so there is roughly the same number of years on each side of the cutoff. More importantly, there is a major tax reform in 2003 that changes the relative tax disadvantage of individual income to capital gains. Breaking the sample at 2003 ensures that I can parameterize the effects of the tax change explicitly, so that my estimation is not confounded by the structural break in tax code.

[Insert Table 8]

Columns 1 and 4 in Table 8 present summaries for the two subsamples of firms based on the actual data. The results suggest that over time, top executives face higher turnover rate, and they choose slightly lower and smoother dividends. However, firms' investment opportunities and external financing frictions are also likely to have changed over time; hence it is difficult to ascertain whether the two trends on dividends and turnovers bear any causal relation. To address this concern, I re-estimate the baseline model on the two subsamples and calculate the magnitude of the career concern-related dividend smoothing channel based on the estimated model parameters. First, comparing the actual versus simulated data in Table 8 suggests that the model fits the data well for both subsamples, especially for the moments related to firms' payout policy. Without this result, it would be hard to claim that the model could illuminate what drives the time series changes in dividend smoothing. Secondly, the fit of the model is slight better for the Late subsample, when the managers have a greater incentive to smooth dividends based on their personal interests. Lastly, I find that the career concern-based dividend smoothing is 14% higher in the Late subsample. This result establishes a causal link between the time series changes in turnover and dividend smoothness. It suggests that, over time, managers face more severe career concerns, and they attempt to mitigate this effect by smoothing dividends even further.

5.5 Information Environment and Dividend Smoothing

Theories on dividend smoothing are typically built on the idea that managers rely on smooth dividends as a signal to reveal private information to the market participants. However, the empirical literature has found at best mixed evidence for the effect of information environment on dividend smoothness.[15] In this section, I revisit this relation by relying on the estimates from a structural model. The purpose of this exercise is three fold. First, I present evidence on how the dividend smoothness varies across subsamples with different information opaqueness measures. Second, I disentangle the driving forces behind such cross-sectional variations in dividend smoothness. Last and most importantly, I provide explanations on what could potentially confound the relation between information and dividend smoothing and reconcile my findings with those in the prior literature. I consider five information opaqueness measures: the analyst forecast dispersion, firms' marginal q, the asset tangibility, the percentage of shares held by institutional investors, and the abnormal accruals calculated under the modified Jones model.

[Insert Table 9]

In Panel A of Table 9, I focus on the SOA of dividends observed in the data. For firms with larger abnormal accruals and more dispersed analyst forecasts, the SOA of dividends is almost monotonically decreasing, and the correlation between the information opaqueness measures and the dividend smoothness is approximately -0.7, suggesting that firms operating in more opaque environments have greater incentives to smooth. On the other hand, if I measure information asymmetry by firms' marginal q or asset tangibility, the SOA of dividends exhibits a U-shaped pattern. Most surprisingly, if I look at the percentage of institutional investors measure, the results imply that firms smooth dividends to a larger extent when they have more institutional holdings, and are hence more transparent. This result seems to have disqualified the information story, according to which, firms with more asymmetric information should have

[15] see Dewenter and Warther (1998), Leary and Michaely (2011), and Michaely and Roberts (2012) for a comprehensive discussion.

stronger signaling incentives and hence smooth dividends more.

Consistent with the literature, I document mixed evidence on the relation between the informational environment and dividend smoothing. Next, I rely on my model estimates within each subsample to investigate what drives such patterns. Recall that in the model, information environment has a direct effect on dividend smoothing: a more opaque environment makes it harder for the investors to gauge a firm's persistent earnings, and such difficulty gives rise to a stronger incentive for managers to signal via dividends. However, when we put this idea to test, one challenge is to find a good proxy that isolates the effect of information asymmetry. Unfortunately, the commonly used information measures, such as marginal q and institutional shareholders, are all correlated with other firm characteristics, such as the strength of corporate governance. This imperfect proxy problem creates a parallel governance effect: those firms with stronger governance usually impose "stricter" turnover rules, which induce managers to smooth dividends even more based on their own career concerns. This parallel governance channel therefore indicates that firms with more transparent information (and stronger governance) should have less responsive dividends. In reality, how dividend smoothing varies across subsamples depends on the relative effects of the two counteracting channels, as well as how strongly the specific information proxy is correlated with these two channels.

In Panel B of Table 9, I isolate the percentage of dividend smoothing driven by managers' career concerns following the algorithm in Table 7. As I move from Quintile 5 (most opaque firms) to Quintile 1 (most transparent firms), the career concern-based dividend smoothing increases by an average of 7%. The evidence is consistent with the idea that the information measures I consider are correlated with the strength of corporate governance. When firms become better governed, the risk of being terminated becomes a more important concern for the top managers, and hence, managers will increase dividend smoothing based on this concern. The magnitude of this increase is largest when I focus on institutional shareholders or abnormal accruals, which suggests that these two measures are most correlated with the other driving force behind the smooth dividends.

Lastly, I examine what would happen if these information measures are designed to capture solely the degree of information asymmetry. To this end, I calculate the model-predicted SOA assuming that managers are free from career concerns. In the results presented in Panel C of Table 9, any cross-sectional dispersion in SOA is driven entirely by the differences in the information environments. Panel C shows that there is a steady decline in the responsiveness of dividends among firms with higher information symmetry. Comparing the results in Panel A and C highlights that using endogenous proxies could dampen the relation between information and dividend smoothing. Once the confounding effects are eliminated, I uncover a strong negative link between information transparency and dividend smoothness for four out of the five information opaqueness measures.

6 Robustness

In the baseline model discussed above, the key friction that I consider is the inseparability between a firm's dividend payout and the information revelation, combined with managerial career concerns. I impose a list of simplifying assumptions with respect to how a firm's operational and turnover decisions are formed to make the model tractable and estimable. In this section, I relax the assumptions one by one to ensure that the quantitative effects based on the model are not sensitive to any of these simplifications.

[Insert Table 10]

6.1 Lumpy Investments

In Section 2, I present a simple dynamic investment model with no capital adjustment costs and perfectly reversible investments. These features do not generate the lumpiness of investments that is observed in the data. To resolve this discrepancy, I add capital adjustment costs,

$A_t(I_t, K_{t-1})$, to the model:

$$A_t(I_t, K_{t-1}) = \alpha_0 \times K_{t-1} + \alpha_1 \times \left(\frac{I_t}{K_{t-1}}\right)^2 \times K_{t-1} - \alpha_2 \times |I_t| \times \mathbf{1}\{I_t < 0\}, \tag{14}$$

which captures the fixed and convex costs of investments, as well as the asymmetric prices in buying and selling capital. In particular, because the cost $\alpha_0 \times K_{t-1}$ is independent of the amount of investment, it induces optimal lumpy behavior, where the firm is inactive for long spells before investing a large amount. I resolve the model, using the baseline parameter estimates reported in Table 3 and the adjustment cost parameter values estimated by Cooper and Haltiwanger (2006). This model with adjustment costs generates investment spikes. Managers' personal interests account for 39.84% of dividend smoothing even in the presence of lumpy investments.

6.2 Stochastic Equity Issuance Cost

In this section, I extend the model in Section 6.1 by incorporating a stochastic equity issuance cost. Following Eisfeldt and Muir (2012), I model a firm's ex-post equity issuance cost by $\xi_t \times \nu_1 \times E_t$, where $\{\xi_t\}$ is stochastic and follows an AR(1) process with a mean of 1, autocorrelation of ρ_ξ, and an unconditional variance of $\frac{\sigma_\xi^2}{1-\rho_\xi^2}$.[16] On one hand, having stochastic issuance cost allows a firm to "time the market" by floating more equity when it is relatively cheap to do so and staying away from the external capital market when the perceived costs increase. Thus, a firm's cash flows will be more volatile across time, making it harder for managers to smooth dividends. On the other hand, having stochastic issuance cost also introduces another source of uncertainty—increasing the value of learning and hence the benefit to signal via dividends. My results suggest that these joint effects lead to a 4% increase in the estimated career concern-based dividend smoothing.

[16] I resolve the model, using the stochastic issuance cost parameters in Nikolov and Whited (2014).

6.3 Linear-Quadratic Cost for Repurchases

In this section, I separate a firm's cost for issuing new equity versus the cost of conducting share repurchases. I use a simple linear function $\nu_1 \times E_t \times \mathbf{1}\{E_t > 1\}$ to capture a firm's equity issuance cost. I include a fixed, as well as linear and convex costs for repurchasing shares:

$$F(E_t) = (f_0 + f_1 \times |E_t| + f_2 \times \tfrac{E_t^2}{K_{t-1}}) \times \mathbf{1}\{E_t < 1\}. \tag{15}$$

Empirically, dividends and repurchases are close substitutes. I keep the form of the repurchase costs in equation (15) parsimonious enough to ensure that the model matches the time series variation of a firm's repurchase activities while predicting stable dividends. Hence, the model does not generate the smoothness of dividends mechanically by shifting volatility to the other form of cash distribution. Compared with the baseline estimation presented in Table 3, the model described above calls for on average lower, but more convex cost for conducting share repurchases. The results imply that firms will repurchase more frequently, but the scale of the program will be smaller in cash-rich states. In a model with linear-quadratic repurchase cost, managers still have strong incentives to smooth dividends. I found that their personal incentives contribute to 37% of the observed dividend smoothness in the data.

6.4 Earnings Management

In this section, I examine a case where managers smooth earnings by managing accounting accruals. Earnings management offers the managers an extra degree of freedom to "window dress" performance and smooth the release of negative information. I model the aggregate earnings management as a separate choice variable $\{M_t\}$. The amount of earnings management performed in the current period is denoted by $M_t - M_{t-1}$, which is additive to the reported earnings but has no real effect on the firm's sources and uses of funds constraint. Earnings management is associated with a quadratic cost $q \times \left(\frac{M_t}{K_{t-1}}\right)^2 \times K_{t-1}$. This cost function captures the idea that larger cumulative earnings management is more likely to be detected, and hence is perceived more costly by the firm. I calibrate the earnings management cost parameter, q, separately by matching the average unsigned abnormal accruals calculated under the modified Jone's model (Dechow, Sloan, and Sweeney, 1995). Earnings management generates two opposing effects on the model-predicted dividend smoothing. First, it implies that a firm's earnings are more likely to reflect managerial manipulation instead of underlying economic conditions, and are thus noisier. Anticipating this effect, the shareholders will put a greater weight on the firm's dividend policy (which is more verifiable) when they forecast productivity. This effect induces managers to smooth dividends to a greater extent, holding all else equal. Second, earnings management offers a separate channel for managers to withhold negative news, which implies that the managers become less reliant on dividend smoothing, and hence they will smooth less. The results in Table 10 suggest that overall, the second channel dominates. Mangers' personal interests have a slightly weaker effect on dividend smoothing once earnings management is incorporated into the model, but the effect is still quantitatively large and highly significant (27.34%). The result also confirms that managers follow a "two-tier" smoothing decision: They first smooth reported earnings relative to the true realized earnings, and then, they further smooth dividends relative to the reported earnings. Managers operate on both margins to enhance their control of the information release and achieve higher utility.

6.5 Managers' Outside Opportunities

In the baseline model, I do not account for managers' outside opportunity; I implicitly assume that upon a forced turnover, the managers' wage rate drops to zero and stays at this level thereafter. This assumption captures the idea that there is large rent income loss upon executives' departures, but quantitatively, it introduces an upward bias in their personal cost estimate. Eckbo, Thorburn, and Wang (2012) study executive turnover decisions for firms under financial distress. They find that 37% of the CEOs who are forced to leave office are able to find new jobs within three years. Conditional on finding full-time executive positions, the median annualized income change for these CEOs is -47%. To check whether similar effects exist in my sample, I randomly pick 100 top executives who experience forced turnover and track their career path for the subsequent five years. Consistent with Eckbo, Thorburn, and Wang (2012), I only identify 24 cases where a top executive moves on to another full-time position in a different company. The average income loss among these 24 executives is 33%. As a robustness check, I use the expected executives' subsequent wage decreases as a measure for their rent income, which captures how much more they are able to extract above and beyond their second best options. I re-write the executives' utility as the sum of their rent income stream, plus the value of their equity stake, and I re-estimate the model. All parameter values remain quantitatively similar.

6.6 Executives' Stock Options

In the baseline model, I combine the managers' holdings of stocks and stock options into their equity stake in the firm. In reality, one may worry about this grouping because stock options typically do not come with dividend rights. This feature of compensation contracts imposes a personal cost of dividend distributions on the managers and makes them favor stock repurchases over dividends (Hall and Liebman, 2000). To alleviate this concern, I separate the executives' stock holdings, vested and non-vested stock options (averaging at 0.76%, 0.22%, and 0.15% of a firm's total shares outstanding, respectively), and re-estimate the model. The estimation yields parameter values roughly in line with the baseline case, except that the costs for repurchase have increased. These higher costs offset the managers' reluctance to pay dividends and make the model-predicted payout levels track the actual data. Despite the changes in parameter estimates, the magnitude of career concern-based dividend smoothing remains significant and quantitatively similar.

6.7 Executives' Risk Aversion

In the baseline model, I assume risk neutrality for both the shareholders and the managers. This assumption is built on the idea that the top executives have sufficiently large outside wealth and can achieve good income smoothing without relying on the firm. Now, I relax this assumption by considering risk-averse managers with habit formation. Managers are assumed to have exponential utility, $u_t = 1 - \frac{1}{\iota}e^{-\iota(W_t - h \times W_{t-1})}$. The risk-aversion parameter is $\iota = 2$, while the habit persistence parameter is $h = 0.74$. I further assume that managers do not have access to personal savings or investment opportunities, so that their per period consumption equals their wage income. Unlike in the baseline case where managers get fixed paychecks, I now model the managers' salaries, $\eta_z \times \hat{z}_t$, as fully performance-dependent, and proportional to the shareholders' perceived match-specific profitability, $\hat{z}_t$. I estimate the value of η_z to match the correlation between executive total salary and firm return on assets. Consistent with Lambrecht and Myers (2012, 2014), risk-averse, habit-persistent managers tend to smooth dividends more aggressively in order to smooth their marginal utility. The managers' utility maximization incentive alone accounts for roughly 50% of the observed dividend smoothing in the data. This result, combined with the one presented for the baseline model, defines the range of excessive dividend smoothing in cases where the managers can smooth their income stream to some degree, but not perfectly.

6.8 Board's Entrenchment

In reality, executive turnover decisions are determined by the level of entrenchment established between the board members and top executives (Taylor, 2010), as well as by the degree of stock price pressure (Warner, Watts, and Wruck, 1988) or relative firm performance (Parrino, 1997). To examine this effect, I follow Taylor (2010) and assume that the board faces a personal turnover cost, c_p, which is constant across time and independent of the firm's economic conditions. I further assume that the entrenched board has the same information set as the executives and can observe the firm's underlying productivity processes. The board will initiate a turnover if the criterion described in equation (10) is satisfied, or if the net benefit from turnover conditional on their information set exceeds c_p. I re-estimate the model with an entrenched board, and the results suggest that the board faces substantial personal cost from firing an executive. Thus, they tend to under-utilize their information and delay turnover decisions. On the other hand, having an entrenched board does not change the model predictions quantitatively. Managers still have strong incentives to smooth the information released by dividends based on their career concerns, and such incentives explain 34.56% of dividend smoothness in the data.

6.9 Asymmetric Learning Speed

Lastly, I consider a case where the investors extract a different amount of information from dividend increases versus cuts. This assumption implies that the forecasting coefficient, γ_Ω in equation (7), can vary based on the direction of the announced dividend change. Estimating this alternative specification suggests that shareholders tend to learn faster when the dividend announcements convey negative information, which gives the managers stronger incentives to stay away from dividend cuts. This result is consistent with the survey evidence documented in Brav et al. (2005). Table 10 suggests that a model with asymmetric learning speed generates parameter estimates very similar to the baseline case, and it only predicts a slightly higher degree of career concern-based dividend smoothing (34.76%). I also estimate models where the learning is based on having dividend increases or declines instead of on the magnitude of the changes. The results (untabulated) are also quantitatively similar.

7 Conclusion

This paper provides a framework to study the interaction of dividend smoothing, firm value, and managers' well-being. I build and estimate a dynamic agency model in which managers distribute dividends not only to signal the persistence of firm earnings, but also to influence their own turnover risk. This setup departs from most of the dividend-signaling literature by explicitly modeling the difference in the agent and principle's signaling incentives. It is also the first in the literature to disentangle how different stakeholders' incentives impact dividend smoothing and discuss the individual welfare implication.

Estimating the model yields three major findings. First, a model that embeds the managers' career concern-based dividend smoothing fits the data much better than a model ignoring this channel. This model can be used to understand the variations of dividend smoothing across a wide spectrum of industries as well as across time. Second, I parameterize the model according to the estimations obtained from the actual data and use it to disentangle the underlying forces behind the observed smooth dividends. The results suggest that the smoothness is driven 60% by the shareholders' preference to stabilize price and 40% by the managers' attempt to ease their career concerns.

Last and most importantly, I analyze the welfare implication of this career concern-based dividend smoothing and find that it distorts firm policy and lowers the equilibrium firm value by 2.09%. This result, however, is not suggesting that this type of career concern-based dividend smoothing represents an inefficiency ex-ante on the firm side because we do not have a benchmark to evaluate what is the minimal contracting cost for this agency friction. Finding an answer to this question entails incorporating the current model into a dynamic contracting framework, which could be an interesting topic for future research.

References

Aharony, J. and Swary, I. Quarterly dividend and earnings announcements and stockholders' return: An empirical analysis. *The Journal of Finance*, 35:1–11, 1980.

Allen, F., Bernardo, A. E., and Welch, I. A theory of dividends based on tax clienteles. *The Journal of Finance*, 55(6):2499–2536, 2000.

Asquith, P. and Mullins, D. W. Equity issues and offering dilution. *Journal of Financial Economics*, 15(1):61–89, 1986.

Brav, A., Graham, J. R., Harvey, C. R., and Michaely, R. Payout policy in the 21st century. *Journal of Financial Economics*, 77:483–527, 2005.

Brealey, R. A., Myers, S. C., and Allen, F. Principles of corporate finance, mcgraw-hill, new york. 2005.

Cao, M. and Wang, R. Optimal ceo compensation with search: theory and empirical evidence. *The Journal of Finance*, 68(5):2001–2058, 2013.

Cooper, R. and Haltiwanger, J. On the nature of capital adjustment costs. *Review of Economics Studies*, 73:611–633, 2006.

Corwin, S. A. The determinants of underpricing for seasoned equity offers. *Journal of Finance*, pages 2249–2279, 2003.

DeAngelo, H. and DeAngelo, L. Dividend policy and financial distress: An empirical investigation of troubled NYSE firms. *The Journal of Finance*, 45:1415–1431, 1990.

Dechow, P. M., Sloan, R. G., and Sweeney, A. P. Detecting earnings management. *The Accounting Review*, pages 193–225, 1995.

Dewenter, K. L. and Warther, V. A. Dividends, asymmetric information, and agency conflicts: evidence from a comparison of the dividend policies of Japanese and U.S. firms? *The Journal of Finance*, 53:879–904, 1998.

Dickinson, V., Kimmel, P., and Warfield, T. The accounting and market consequences of accelerated share repurchases. *Review of Accounting Studies*, 17(1):41–71, 2012.

Easterbrook, F. H. Two agency-cost explanations of dividends. *American Economic Review*, 74:650–659, 1984.

Eckbo, B. E., Thorburn, K. S., and Wang, W. How costly is corporate bankruptcy for top executives? *Working paper, Dartmouth College*, 2012.

Eisfeldt, A. and Muir, T. Aggregate issuance and savings waves. *Working paper, UCLA*, 2012.

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

Fama, E. F. and Babiak, H. Dividend policy: An empirical analysis. *Journal of the American Statistical Association*, 63:1132–1161, 1968.

Fudenburg, D. and Tirole, J. A theory of income and dividend smoothing based on incumbency rents. *Journal of Political Economy*, 103:75–93, 1995.

Gao, X. and Ritter, J. R. The marketing of seasoned equity offerings. *Journal of Financial Economics*, 97(1):33–52, 2010.

Ghosh, C. and Woolridge, J. R. Stock-market reaction to growth-induced dividend cuts: Are investors myopic? *Managerial and Decision Economics*, 10:25–35, 1989.

Grullon, G., Michaely, R., Shlomo, B., and Thaler, R. Dividend changes do not signal changes in future profitability. *Journal of Business*, 78(5):1956–1982, 2005.

Hall, B. J. and Liebman, J. B. The taxation of executive compensation. In *Tax Policy and the Economy, Volume 14*, pages 1–44. MIT Press, 2000.

Holmström, B. Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182, 1999.

Jenter, D. and Kanaan, F. Ceo turnover and relative performance evaluation. *The Journal of Finance*, 2015.

John, K. and Williams, J. Dividends, dilution, and taxes: A signalling equilibrium. *Journal of finance*, pages 1053–1070, 1985.

Kaplan, S. N. and Reishus, D. Outside directorships and corporate performance. *Journal of Financial Economics*, 27(2):389–410, 1990.

Krusell, P. and Smith, A. A., Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, 1998.

Lambrecht, B. M. and Myers, S. C. A Lintner Model of payout and managerial rents. *The Journal of Finance*, 67:1761–1810, 2012.

Lambrecht, B. M. and Myers, S. C. The dynamics of investment, payout and debt. *Working paper, MIT Sloan*, 2014.

Leary, M. T. and Michaely, R. Determinants of dividend smoothing: Empirical evidence. *Review of Financial Studies*, 24:3197–3249, 2011.

Lintner, J. Distribution of incomes of corporations among dividends, retained earnings and taxes. *American Economic Review*, 46:93–117, 1956.

Mahmudi, H. and Pavlin, M. Corporate payout policy, cash savings, and the cost of consistency: Evidence from a structural estimation. *Financial Management*, 42(4):843–874, 2013.

Michaelides, A. and Ng, S. Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators. *Journal of Econometrics*, 96(2): 231–266, 2000.

Michaely, R. and Roberts, M. R. Dividend smoothing, agency costs, and information asymmetry: Lesson from the dividend policies of private firms. *Cornell University Working Paper*, 2012.

Michaely, R., Thaler, R. H., and Womack, K. L. Price reactions to dividend initiations and omissions: Overreaction or drift? *The Journal of Finance*, 50:573–608, 1995.

Miller, M. H. and Modigliani, F. Dividend policy, growth, and the valuation of shares. *the Journal of Business*, 34(4):411–433, 1961.

Miller, M. H. and Rock, K. Dividend policy under asymmetric information. *The Journal of Finance*, 40(4):1031–1051, 1985.

Morellec, E., Nikolov, B., and Schürhoff, N. Corporate governance and capital structure dynamics. *Journal of Finance*, 67(3):803–848, 2012.

Mortensen, D. T. and Pissarides, C. A. Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415, 1994.

Nikolov, B. and Whited, T. M. Agency conflicts and cash: Estimates from a dynamic model. *The Journal of Finance*, 2014.

Nissim, D. and Ziv, A. Dividend changes and future profitability. *The Journal of Finance*, 56 (6):2111–2133, 2001.

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

Parrino, R., Sias, R. W., and Starks, L. T. Voting with their feet: institutional ownership changes around forced CEO turnover. *Journal of Financial Economics*, 68(1):3–46, 2003.

Skinner, D. J. The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics*, 87:582–609, 2008.

Tauchen, G. and Hussey, R. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica: Journal of the Econometric Society*, pages 371–396, 1991.

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

Vermaelen, T. Common stock repurchases and market signalling: An empirical study. *Journal of financial economics*, 9(2):139–183, 1981.

Warner, J. B., Watts, R. L., and Wruck, K. H. Stock prices and top management changes. *Journal of financial Economics*, 20:461–492, 1988.

Warusawitharana, M. and Whited, T. Equity market misvaluation, financing, and investment. *Working Paper, University of Rochester*, 2012.

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

Appendix

A.1 Numerical Solution

In this appendix, I describe how the equilibrium allocation is solved. A graphical illustration of the solution algorithm is presented in Figure A.1

[Insert Figure A.1]

The algorithm is similar to that described in Krusell and Smith (1998). As a preliminary step, I descritize the five state variables $\{z, s, K, L, D\}$. The net asset value lies between 0 and $\bar{K}$, where $\bar{K}$ is the maximal capital that a firm will hold in the first-best case. A firm's dividends lie between 0 and $(0.1 \times \bar{K})$. The shocks to a firm's productivity are transformed into discrete states using the quadrature method described in Tauchen and Hussey (1991). After defining the grids, I solve the equilibrium via the following steps: (1) I guess the shareholders' optimal forecasting rule, $\Gamma = \{\gamma_0, \gamma_\pi, \gamma_\Omega, \gamma_\mathcal{F}\}$. Taking this rule as given, I solve the managers' optimal policy, $\{I', E', L', D'\}$. In equations (5) and (8), managers' utility and the firm's value are interdependent, so they have to be determined simultaneously. (2) To achieve this, I first set the firm's value function to the first-best case:

$$V_{FB}(K, L, D, z, \pi) = \max_{\{I',E',L',D'\}} (1 - \tau_p)D' - E' + \beta \times \mathbb{E}\, V'_{FB}(K', L', D', z', \pi'), \tag{A.1}$$

subject to the sources and uses of funds constraint:

$$Y - I' + \tau_c \delta K' + L \times [1 + r_f(1 - \tau_c)] - L' + \Lambda(E') - D' - W \geq 0. \tag{A.2}$$

Because the model does not include any adjustment cost, I can collapse a firm's holdings of liquid assets and physical capital into a single state variable, A:

$$A = L \times [1 + r_f \times (1 - \tau_c)] + K \times (1 - \delta), \tag{A.3}$$

and I refer to A as the firm's net worth. I can rewrite equations (A.1) and (A.2) as functions of this new state variable:

$$V_{FB}(A, D, z, \pi) = \max_{\{K',A',E',D'\}} (1 - \tau_p)D' - E' + \beta \times \mathbb{E}\, V'_{FB}(A', D', z', \pi'), \tag{A.4}$$

$$s.t. \quad Y + A + (\delta\tau_c - \tfrac{r_f \times (1-\tau_c)}{1+r_f \times (1-\tau_c)})K' + \delta\tau_c K' + \Lambda(E') - D' - W - \tfrac{1}{1+r_f \times (1-\tau_c)} A' \geq 0. \tag{A.5}$$

Conditional on $V_M = V_I = V_{FB}$, I solve the managers' value function, $U(.)$, using value function iteration:

$$U(A, D, z, \hat{z}, \pi) = \max_{\{K', A', E', D'\}} \mathbb{E}[W + \beta(1-\Phi)U(A', D', z', \hat{z}', s') \\ + \kappa_I \triangle V_I(A, D, z, \pi) + \kappa_M \triangle V_M(A, D, \hat{z}, \pi)], \tag{A.6}$$

in which Φ represents the executive turnover decision:

$$\Phi = \begin{cases} 1, & \dfrac{EV_M(K', L', D', \hat{z}', \pi')}{V_M(K, L, D, \hat{z}, \pi)} \leq \underline{r} \\ 0, & \text{otherwise} \end{cases} \tag{A.7}$$

If the firm decides to dismiss its incumbent executives, it takes a random draw from the initial unconditional distribution: $z_{new} \sim N(-c, \frac{\sigma_z^2}{1-\rho_z^2})$ and resets: $z = z_{new}$. Solving equation (A.6) subject to the constraint equation (A.5) yields the managers' optimal decision, $\{K', A', E', D'\}^*$. Based on this optimal decision, I can update the firm's value function:

$$V_I(A, D, z, \pi) = (1-\tau_p)D' - E' - \lambda|\triangle D'| + \beta\, \mathbb{E}\, V_I'(A', D', z', \pi'). \tag{A.8}$$

$$V_M(A, D, \hat{z}, \pi) = (1-\tau_p)D' - E' - \lambda|\triangle D'| + \beta\, \mathbb{E}\, V_M'(A', D', \hat{z}', \pi'). \tag{A.9}$$

I iterate until the value functions $\{U, V_M, V_I\}$ converge. (3) I then generate a panel of firms according to the optimal policy, $\{K', A', E', D'\}^*$ and calculate what forecasting decision best describes the simulated data:

$$\hat{z}' = \gamma_0 + \gamma_\pi \times \pi' + \gamma_\Omega \times \Omega + \gamma_{\mathcal{F}} \times \mathcal{F}. \quad \{\gamma_0, \gamma_\pi, \gamma_\Omega, \gamma_{\mathcal{F}}\} = \arg\min \mathbb{E}\, |\hat{z}' - z|. \tag{A.10}$$

Obtaining the optimal forecasting rule is essentially finding the least absolute deviation estimates in the following regression:

$$z = \gamma_0 + \gamma_\pi \times \pi' + \gamma_\Omega \times \Omega + \gamma_{\mathcal{F}} \times \mathcal{F} + \epsilon. \tag{A.11}$$

Estimating the regression also provides a measure for the goodness-of-fit. (4) I stop if the estimates converge to the initial guess and the process yields a reasonable goodness-of-fit. If the estimates converge, but the goodness-of-fit is poor, I add in additional determinants and try different functional forms of equation (A.11).

A.2 Estimation Procedure

In this section, I briefly outline the estimation procedure. Let $x_{i,t}$ represent the real data vector and let $y_{i,t,s}(\beta)$ represent the simulated data, where $i = (1, 2, 3...n)$ denotes the number of firms, $t = (1, 2, 3...T)$ indicates the number of time periods, and $s = (1, 2, 3...S)$ represents the number of simulated data sets. I explicitly write $y_{i,t,s}$ as a function of β to emphasize the dependence of simulated data on the deep structural parameters. Michaelides and Ng (2000) find that good finite-sample performance of a simulation estimator requires a simulated sample that is approximately ten times as large as the actual data sample. I set $S = 10$ following their suggestion.

Equation (12) contains the first two moments in the estimation process. As Cooper and Haltiwanger (2006) argue, this equation can be derived as an auxiliary equation from the firm's production function and the profitability shock processes. The detailed procedures are illustrated as follows: First, I take the log of a firm's production function,

$$\ln Y_t = \ln z_t + \ln s_t + \theta \times \ln K_t. \tag{A.12}$$

I substitute z_t and z_{t-1} with $(\rho \times z_{t-1} + \epsilon_{z,t})$ and $(\ln Y_{t-1} - \theta \ln K_{t-1} - \ln s_{t-1})$, respectively and rewrite equation (A.12) as:

$$\ln Y_t = \rho \times (\ln Y_{t-1} - \theta \ln K_{t-1} - \ln s_{t-1} + \epsilon_{z,t}) + \ln s_t + \theta \times \ln K_t. \tag{A.13}$$

I rearrange terms in equation (A.13) to get:

$$\ln Y_t = \rho \times \ln Y_{t-1} + \theta \times \ln K_t - \rho \times \theta \times \ln K_{t-1} + (\ln s_t - \rho \times \ln s_{t-1} + \epsilon_{z,t}). \tag{A.14}$$

equation (A.14) is in the same format as equation (12), which can be directly estimated on the simulated data. For the actual data, I remove the firm-fixed effects by focusing on the first difference, and I include a complete set of year dummies to absorb the aggregate shocks. Following the suggestion in Cooper and Haltiwanger (2006), I estimate Equation (A.14) as a nonlinear GMM system, and I use lagged and twice-lagged capital, as well as twice-lagged profit as instruments.

I add another set of 15 moments to pin down the 9 underlying parameters: $\{\rho_z, \sigma_z, \sigma_s, \theta, \nu_1, \nu_2, \delta, \kappa_M, c\}$. The choice of moments and the corresponding parameter estimates are reported in Table 3. The model is estimated using simulated method of moments (SMM). SMM chooses the parameter values to minimize the distance between simulated moments and the corresponding actual moments. Let $m(x_{i,t})$ and $y_{i,t,s}$ denote the moments calculated based on the real and simulated data, respectively. I can write the sample moment condition as:

$$g(x_{i,t}, \beta) = \frac{1}{nT} \sum_{i=1,n} \sum_{t=1,T} \left[m(x_{i,t}) - \frac{1}{S} \sum_{s=1,S} m(y_{i,t,s}(\beta)) \right]. \tag{A.15}$$

The simulated method of moment estimator $\hat{\beta}$ is then obtained by solving:

$$\hat{\beta} = \arg\min_{\beta} g(x_{i,t}, \beta)' \hat{W} g(x_{i,t}, \beta), \tag{A.16}$$

in which $\hat{W}$ is a positive definite matrix that converges in probability to a deterministic positive definite matrix W. I use the inverse of the sample covariance matrix of the moments to construct $\hat{W}$. The calculation follows the influence-function approach described in Erickson and Whited (2002). The simulated method of moment estimator is asymptotically normally distributed:

$$\hat{\beta} - \beta \to N\left(0, avar(\hat{\beta})\right). \tag{A.17}$$

Let Ω denote the variance-covariance matrix of the moment conditions, then the asymptotic variance $avar(\hat{\beta})$ can be expressed as:

$$avar(\hat{\beta}) = \left(1 + \frac{1}{S}\right) \times \left[\frac{\partial g}{\partial \beta} W \frac{\partial g'}{\partial \beta}\right]^{-1} \left[\frac{\partial g}{\partial \beta} W \Omega W \frac{\partial g'}{\partial \beta}\right] \left[\frac{\partial g}{\partial \beta} W \frac{\partial g'}{\partial \beta}\right]^{-1}. \tag{A.18}$$

Figure 1: Timeline

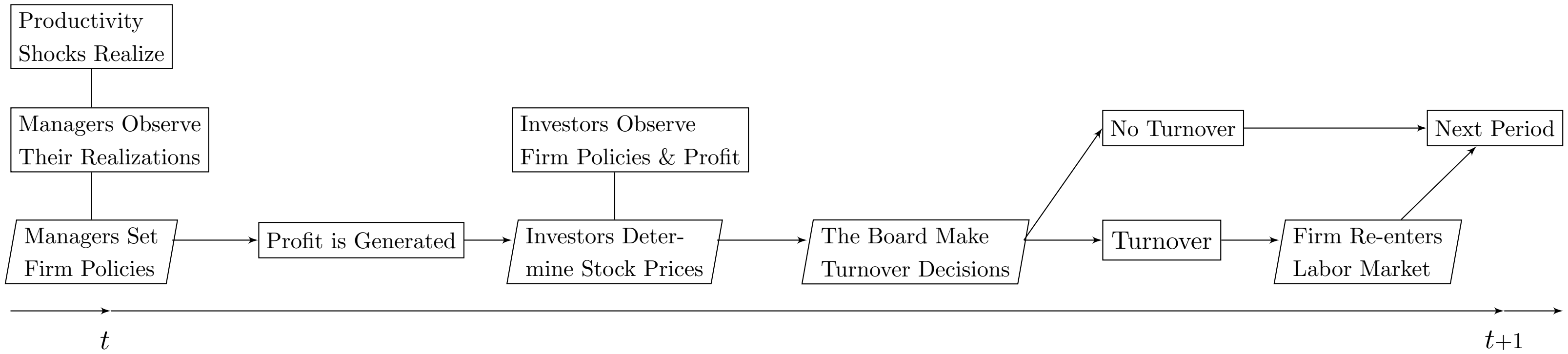

Figure 1 summarizes the timeline of the model. At the beginning of each period, the productivity shocks are realized. Managers observe their realizations and base the firm's investment, financing, and payout decisions on this information. The investors do not directly see the underlying profitability processes. Instead, they extract information from the realized profit and the reported firm policies, and they determine the firm's market price. The board of directors are in charge of the firm's executive turnover decisions.

Figure 2: Endogenous Turnover Rate for Managers

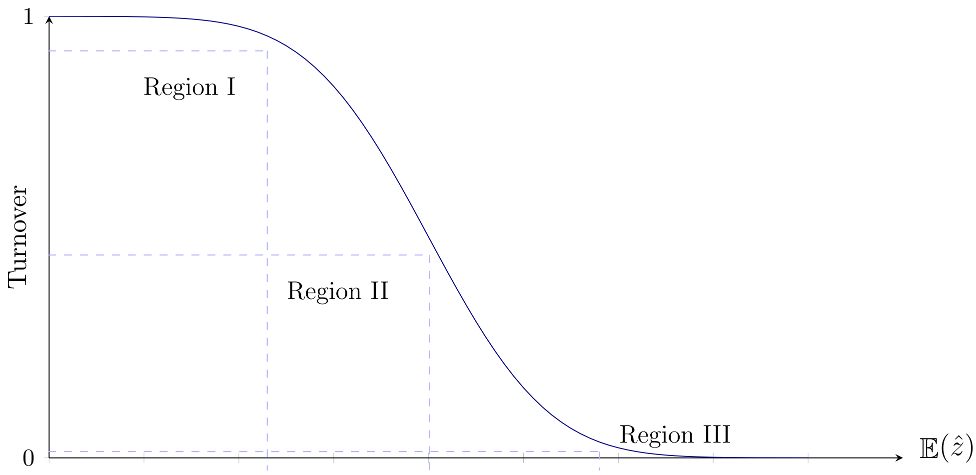

Figure 2 depicts the expected turnover rate for the managers, holding constant a firm's history. Region I, Region II, and Region III correspond to economies with low, moderate, and high average turnovers. The y-axis contains the expected turnover rate, which lies between 0 and 1. The x-axis contains the investors' assessment of the match quality between the firm and executives. Notice that $\mathbb{E}(\hat{z})$ instead of $\hat{z}$ is used as the measurement because $\hat{z}$ contains a signal unobservable to the managers at the time when they choose the optimal policies.

Figure 3: The Information Content of Dividends

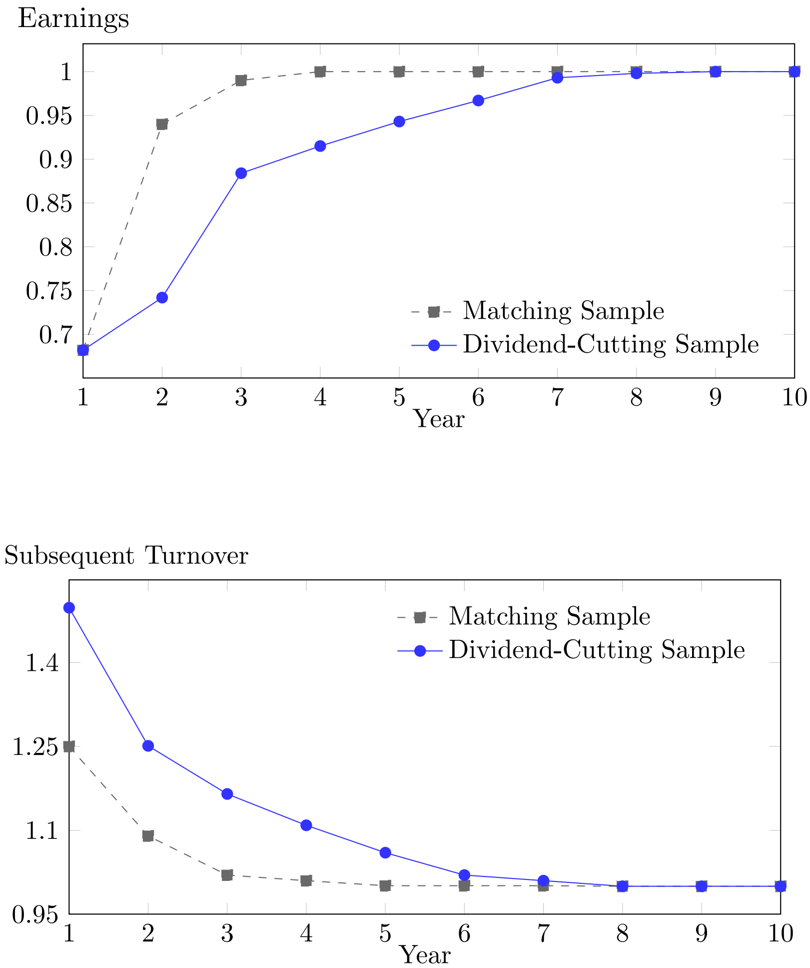

Figure 3 reports the model predicted firm profit and executive turnovers following dividend cuts. The dividend-cutting sample consists of simulated firms who have cut dividends in year one; the matching sample is constructed from simulated firms that have the same year 1 reported earnings as the dividend-cutting sample but have managed to maintain or increase dividends. The x-axis corresponds to the number of years after dividend cuts. On the y-axis, firms' profit and forced executive turnover rates are expressed as fractions of their corresponding steady state levels.

Figure 4: Managers' Career Concerns and Firm Policies

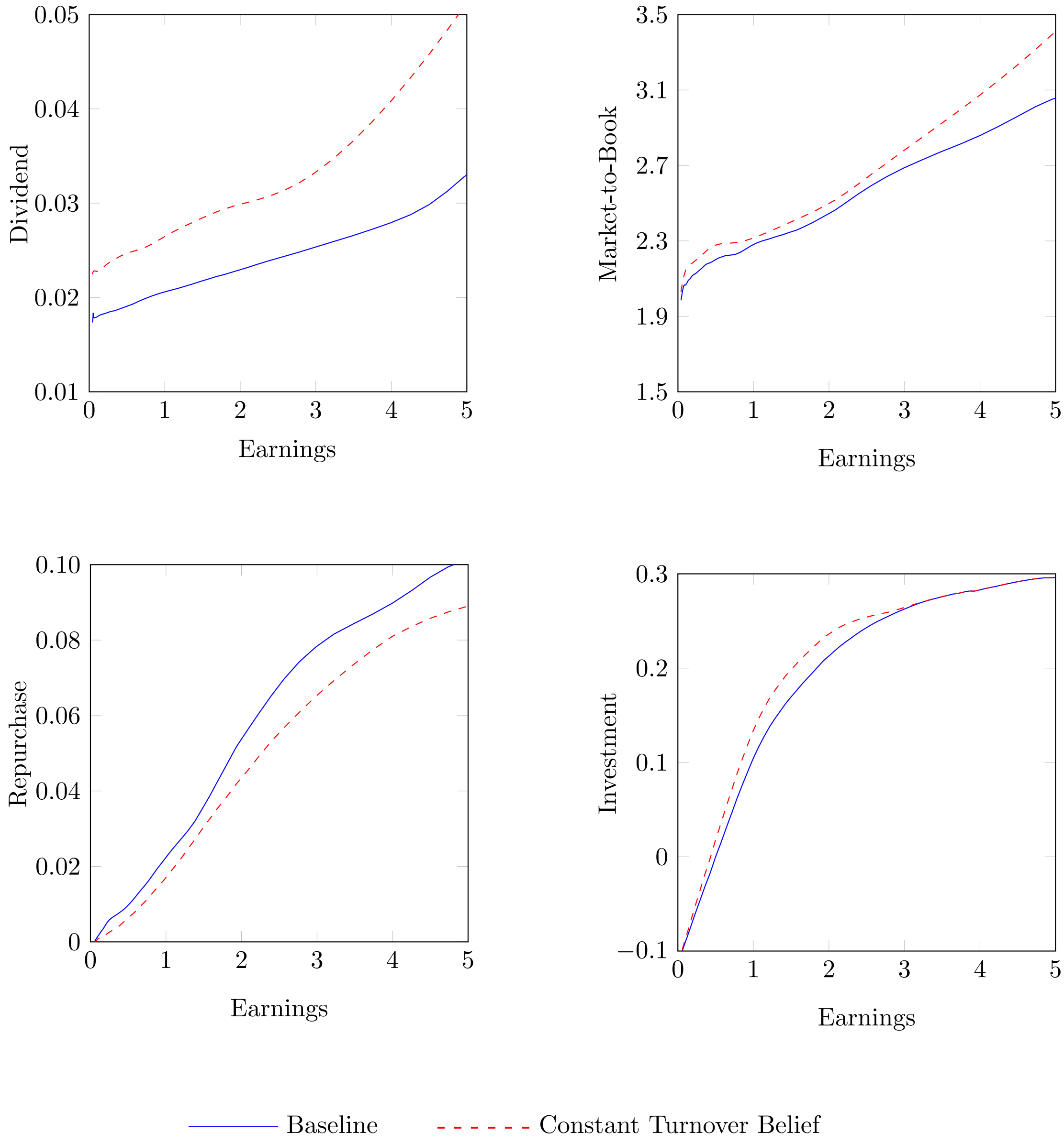

Figure 4 presents the firm policy functions and market-to-book ratio under two alternative model specifications. The solid line corresponds to the baseline model described in Section 2.3; the dashed line is a case where the managers have constant turnover belief. On the x-axis, earnings are normalized by the steady-state level under the constant turnover model. On the y-axis, dividends, repurchases, and equity issuance are scaled by the firm's total assets.

Figure 5: Dividend Smoothing over Time

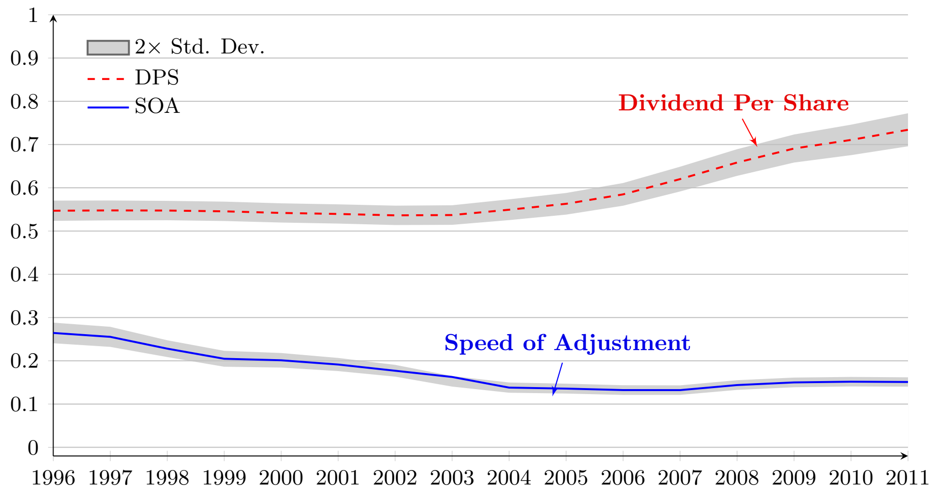

Figure 5 shows the time trend for dividend per share and dividend smoothness among non-financial and non-utility firms. Calculations are based on rolling 10-year windows from 1987 to 2011. x-axis corresponds to the last year in each 10-year subample. Dividend smoothness is measured by the speed of adjustment (SOA), which equals the estimated coefficient β in the following regression: $D_{i,t} - D_{i,t-1} = \alpha + \beta \times (TPR_i \times E_{i,t} - D_{i,t-1}) + \epsilon_{i,t}$, where $D_{i,t}$ and $E_{i,t}$ are the dividend and earnings per share, respectively. TPR_i is a firm's target dividend payout ratio over the surrounding 10-year period.

Figure 6: Industry Estimation

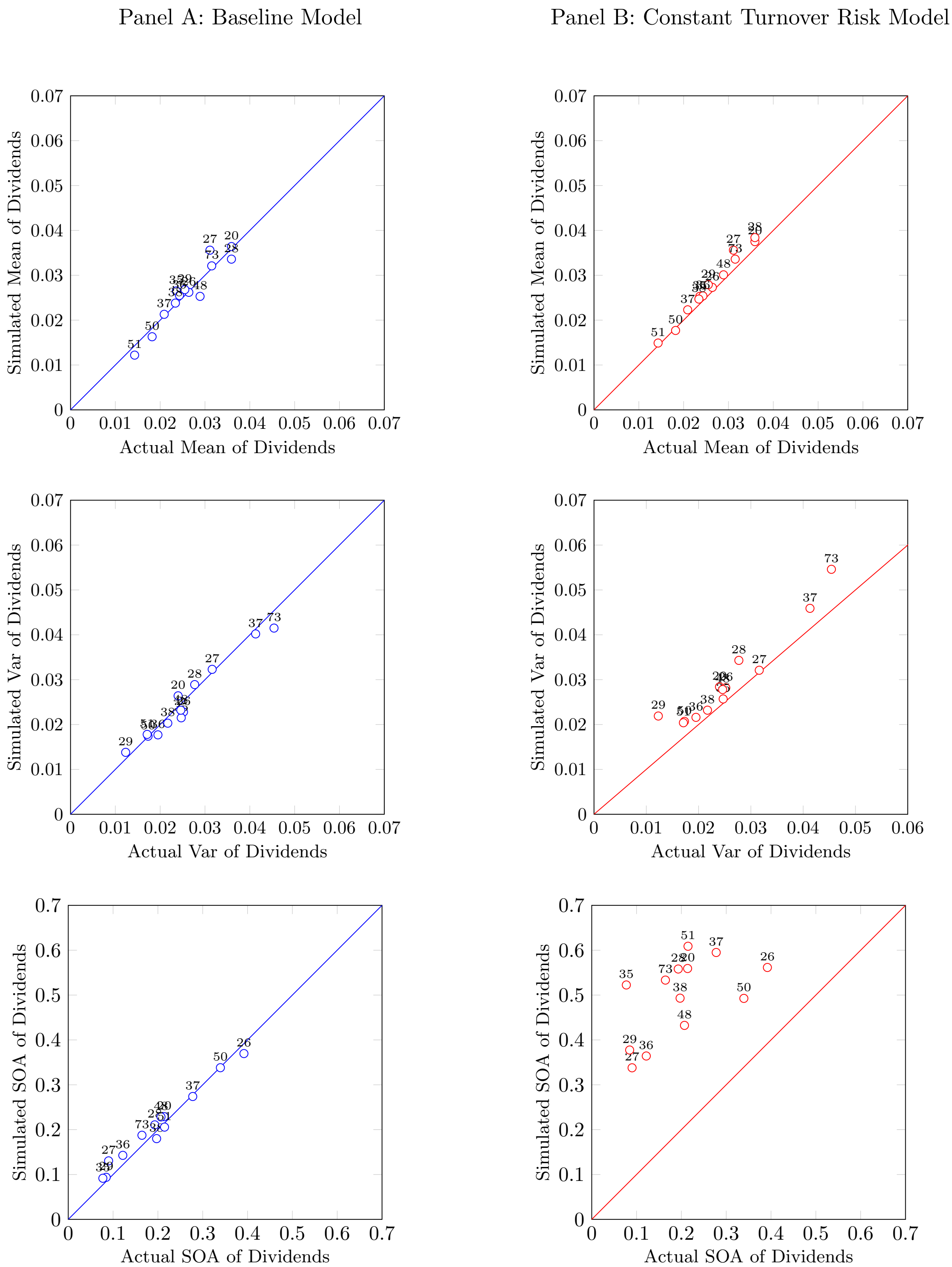

Figure 6 plots the mean, variance, and speed of adjustment for dividends in the actual versus simulated data. The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file over the period 1992–2011, and it is split into industry subsamples based on two-digit SIC codes. The model is re-estimated on the 17 industries with more than 300 firm-year observations. Panel A reports the moments under the baseline model; Panel B shows the results under a restricted model that assumes constant executive turnover rates across times and states. Dividends are deflated by total book assets in both the simulated and actual data.

Figure A.1: Numerical Strategy

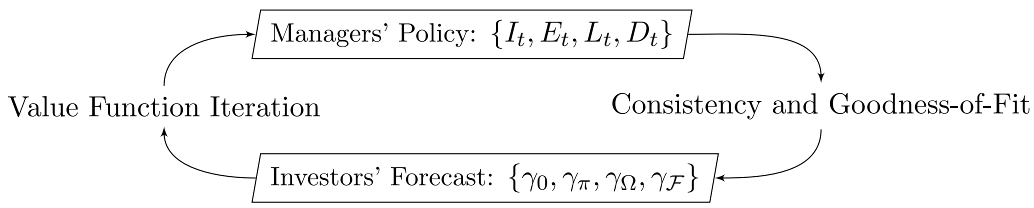

Figure A.1 illustrates how the model described in Section 2 is solved numerically. The process starts by guessing a forecasting rule for the productivity processes and assigning it to the investors. Taking this rule as given, managers set the firm policies to maximize their utility. In anticipation of the managers' decision-making process, the investors choose the optimal forecasting rule so that they make the best possible predictions of the underlying productivity processes. An equilibrium is achieved when both the managers' policies and the investors' forecasting decisions converge simultaneously, which guarantees that no party has any incentive to deviate from their current strategies.

Table 1: Variable Description

Table 1 presents the variable definitions and the parameter values used to calculate the baseline solution.

Panel A: Calculated outside of the model		
r_f	One-year risk-free interest rate	2%
τ_c	Corporate income tax rate	35%
τ_p	Personal tax rate on dividends	15%
η	Executive total compensation (% steady state asset)	0.0978
κ_I	Executives' stock and option holdings (% shares outstanding)	1.1268
λ	Cost for dividend adjustment	1.37%
Panel B: Solved in equilibrium allocation		
$\{K_t\}$	Stock of physical capital	/
$\{L_t\}$	Holdings of liquid assets	/
$\{E_t\}$	Net equity issuance	/
$\{D_t\}$	Regular dividend distributions	/
$\{\hat{z}_t\}$	Investor' forecasted match-specific (persistent) profitability	/
$\{\Phi_t\}$	Dummy variable indicating forced executive turnover	/
Panel C: Estimated within the model		
ρ_z	Autocorrelation of the persistent profitability	0.7021
σ_z	Standard deviation of innovations to the persistent profitability	0.2926
σ_s	Standard deviation of the transitory shock	0.0783
θ	Curvature of a firm's production function	0.5524
δ	Capital depreciation rate	0.0983
ν_1	Linear external financing cost	0.0663
ν_2	Quadratic external financing cost	0.7104
κ_M	Weight of firm market value in executives' utility function	0.6223
c	Cost for executive turnover	0.4607
Panel D: Shock Processes		
$\{\ln(z_t)\}$	Match-specific (persistent) profitability shock	$\sim \rho_z \times \ln(z_{t-1}) + N(0, \sigma_z^2)$
$\{\ln(s_t)\}$	Pure transitory shock to profit	$\sim N(0, \sigma_s^2)$
$\{\varphi_t\}$	Additional signal observed by shareholders	$\sim N(\ln(z_t), \frac{\sigma_z^2}{1-\rho_z^2})$

Table 2: Summary Statistics

The sample is constructed from the 2013 Compustat and CRSP file for the period 1992–2011. Non-financial and non-utility firms who declare at least six positive annual dividends in the surrounding 10-year window are included. This sample is merged with executive compensation data from the ExecuComp, institutional ownership data from Thomson Financial, and a hand-collected dataset on top executive turnovers. This process yields 10,827 distinct firm-executive pairs and 11,626 firm-year observations.

	Mean	Std. Dev.	25%	50%	75%
Firm Investment and Financial Characteristics					
Size	8.0269	1.6111	6.8491	7.9384	9.1646
Operating Income	0.1630	0.0790	0.1091	0.1490	0.2067
Leverage	0.2542	0.1427	0.1510	0.2597	0.3546
Market-to-Book	2.7443	2.4076	1.3271	2.0057	3.1630
Cash	0.0707	0.0911	0.0120	0.0339	0.0906
Equity Issuance	0.0112	0.0186	0.0005	0.0041	0.0164
Investment	0.0985	0.0684	0.0591	0.0796	0.1256
Dividend	0.0240	0.0179	0.0114	0.0198	0.0311
Repurchase	0.0297	0.0331	0.0000	0.0042	0.0319
Dividend to Earnings Ratio	0.1542	0.0984	0.0813	0.1399	0.2094
Dividend per Share	0.6581	0.5930	0.2141	0.4797	0.9133
Earnings per Share	4.5869	3.8815	2.0428	3.5942	5.8738
Stock Return	0.1226	0.4556	-0.1040	0.0820	0.2775
Asset Tangibility	0.4627	0.3530	0.2220	0.4328	0.6367
Institutional Holdings	0.4287	0.2420	0.2581	0.4611	0.6403
Analyst Forecast Dispersion	0.0067	0.0161	0.0049	0.0063	0.0085
Abnormal Accrual	0.0641	0.0768	0.0252	0.0526	0.0992
Managerial Characteristics					
Salary	0.0251	0.0317	0.0050	0.0133	0.0318
Bonus	0.0129	0.0211	0.0010	0.0053	0.0150
Total Compensation	0.0735	0.0981	0.0173	0.0399	0.0864
% Stock Holdings	0.6041	1.5196	0.0318	0.0909	0.3124
% Vested Options	0.1826	0.2169	0.0364	0.1067	0.2467
% Unvested Options	0.1327	0.1682	0.0224	0.0762	0.1708
Tenure	8.9538	10.8480	3	9	12

Table 3: Simulated Moments Estimation: Full Sample

The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file for the 1992–2011 sample period. The estimation is based on the baseline model described in Section 2, and estimation is done with simulated method of moments (SMM). Panel A reports the simulated versus the actual moments, along with the t-statistics for the pairwise differences. The J-statistic tests the over-identification constraint for the moment conditions. Panel B reports the parameter estimates with clustered standard errors in parentheses. ρ_z, σ_z, and σ_s govern the persistence and standard deviation of the firms' shock processes. θ is the curvature of the firms' production function. δ is the rate of depreciation. $\{\nu_1, \nu_2\}$ represents the linear-quadratic costs for net equity issuance. κ_M measures to what extent managers care about firms' market price, and c captures the opportunity cost for executive turnovers.

Panel A: Moments

	Actual Moments	Simulated Moments	T-test for Difference
Coefficient γ_y in equation (12)	0.6514	0.7002	-0.5918
Coefficient γ_k in equation (12)	0.5404	0.5811	-0.6023
Mean of Operating Income	0.1630	0.1727	-1.0336
Std of Operating Income	0.0738	0.0874	-2.0786
AR(1) Coefficient of Operating Income	0.6368	0.6631	-0.4074
Mean of Investment	0.0985	0.1059	-1.1205
Std of Investment	0.0645	0.0794	-1.6566
AR(1) Coefficient of Investment	0.6149	0.5608	1.5046
Mean of Net Equity Issuance	-0.0185	-0.0210	0.5027
Std of Net Equity Issuance	0.0171	0.0167	0.0436
Mean of Market-too-Book	2.7443	2.4120	1.8407
Std of Market-to-Book	2.2892	2.1883	0.6977
Mean of Dividend	0.0240	0.0253	-1.2137
Std of Dividend	0.0102	0.0119	-0.1499
SOA of Dividend	0.2040	0.2014	0.0564
Mean Turnover	0.0263	0.0259	0.2488
Corr btw Return and Turnover	-0.0674	-0.0652	-0.1165
		J-statistics : 11.97	P-val : 0.194

Panel B: Parameter Estimates

ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c
0.6977	0.3188	0.8362	0.5903	0.9914	0.6079	0.6466	0.5562	0.4282
(0.0941)	(0.0095)	(0.0556)	(0.1291)	(0.0104)	(0.2057)	(0.1014)	(0.1309)	(0.0198)

Table 4: Predicting Forced Turnover

The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file for the 1992–2011 period. The sample contains 10,827 distinct firm-executive pairs and 47,563 executive-year observations. Table 4 reports the coefficient estimates from equation (13). The outcome variable is an indicator of forced executive turnover. $Profit_{-i}$ measures the firm's ROA in the previous ith year; *% Insider* is the firm-level aggregate insider holdings; *% Executive* is the executive-level holdings of the firm's own stocks. *Net Repurchase* represent the firm's stock repurchases minutes new equity issuance. Both *Investment* and *Net Repurchase* are scaled by the firm's total book assets. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The significance levels in columns (1)-(4) are calculated based on standard errors clustered by industry and year. The significance levels in columns (5) and (6) are calculated using bootstrapped standard errors. The economic significance of the predictors are reported in brackets, which measures the probability effect on the outcome if the predictor increases from the lower to the upper 10th percentile of its distribution.

	RealData				Simulated Data	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-3.947***	-4.428***	-4.289***	-4.428***	-3.701***	-3.581***
Profit	-3.540***	-3.502***	-3.556***	-3.502***	-2.275***	-3.391***
	[-0.609]	[-0.602]	[-0.612]	[-0.602]	[-0.467]	[-0.581]
△ DPS	-0.543**	-0.650***	-0.631***	-0.667***	-0.925***	-0.466**
	[-0.423]	[-0.514]	[-0.498]	[-0.529]	[-0.741]	[-0.373]
Investment						-0.169*
						[-0.058]
Net Repurchase						-0.851***
						[-0.071]
Size		-0.017	0.004	0.005		
		[-0.134]	[0.028]	[0.042]		
$Profit_{-1}$		0.304	0.397	0.743*		
		[0.047]	[0.061]	[0.113]		
$Profit_{-2}$		-1.095***	-0.829**	0.261		
		[-0.175]	[-0.132]	[0.040]		
Tenure			-0.032***	-0.033***		
			[-0.552]	[-0.562]		
% Institution			0.758***	0.842***		
			[0.529]	[0.581]		
% Insider			0.246**	0.174*		
			[0.094]	[0.067]		
% Executive			-0.078*	-0.077*		
			[-0.682]	[-0.667]		
SIC/Year FE	N	N	N	Y	N	N

Table 5: Simulated Moments Estimation: No Agency Career Concerns

The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file for the 1992–2011 period. The estimation is based on an alternative model where executives are assumed to face constant turnover rates across all times and states, and the estimation is done with simulated method of moments (SMM). Panel A reports the simulated versus the actual moments, along with the t-statistics for the pairwise differences. The J-statistic tests the over-identification constraint for the moment conditions (excluding the last two moments). Panel B reports the parameter estimates with clustered standard errors in parentheses. ρ_z, σ_z, and σ_s govern the persistence and standard deviation of the firms' shock processes. θ is the curvature of the firms' production function. δ is the rate of depreciation. $\{\nu_1, \nu_2\}$ represents the linear-quadratic costs for net equity issuance. κ_M measures to what extent managers care about firms' market price, and c captures the opportunity cost for executive turnovers.

multicolumn5lPanel A: Moments

	Actual Moments	Simulated Moments	t-test for Difference
Coefficient γ_y in equation (12)	0.6514	0.6480	0.0383
Coefficient γ_k in equation (12)	0.5404	0.5212	0.2908
Mean of Operating Income	0.1630	0.1731	-1.3400
Std of Operating Income	0.0738	0.0692	0.8403
AR(1) Coefficient of Operating Income	0.6368	0.5124	1.8638
Mean of Investment	0.0985	0.0988	-0.0469
Std of Investment	0.0645	0.0749	-1.5759
AR(1) Coefficient of Investment	0.6149	0.6838	-1.9153
Mean of Net Equity Issuance	-0.0185	-0.0147	-0.6970
Std of Net Equity Issuance	0.0171	0.0090	0.9174
Mean of Market-too-Book	2.7443	2.2654	2.6231
Std of Market-to-Book	2.2892	2.4259	-1.0061
Mean of Dividend	0.0240	0.0287	-3.2027
Std of Dividend	0.0102	0.0148	-0.4996
SOA of Dividend	0.2040	0.3283	2.2973
Mean Turnover	0.0263	0.0263	/
Corr btw Return and Turnover	-0.0674	/	/
		J-statistics : 49.87	P-value : <0.01

Panel B: Parameter Estimates

ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c
0.6685	0.3070	0.0859	0.5342	0.0941	0.0681	0.6334	0.5231	0.1389
(0.0462)	(0.0291)	(0.0121)	(0.1291)	(0.0134)	(0.0374)	(0.2843)	(0.1078)	(0.0923)

Table 6: Simulated Moments Estimation with Alternative Weight Matrix

The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file for the 1992–2011 period. This table considers an alternative model where executives are assumed to face constant turnover rates across all times and states. The estimation is done with simulated method of moments (SMM), and it puts high weights on moments related to dividend levels and smoothness. Panel A reports the simulated versus the actual moments, along with the t-statistics for the differences. The J-statistic tests the over-identification constraint for the moment conditions (excluding the last two moments). Panel B reports the parameter estimates with clustered standard errors in parentheses. ρ_z, σ_z, and σ_s govern the persistence and standard deviation of the firms' shock process. θ is the curvature of the firms' production function. δ is the rate of depreciation. $\{\nu_1, \nu_2\}$ represents the linear-quadratic costs for net equity issuance. κ_M measures to what extent managers care about firms' market price, and c captures the opportunity cost for executive turnovers.

Panel A: Moments

	Actual Moments	Simulated Moments	t-test for Difference
Coefficient γ_y in equation (12)	0.6514	0.6716	0.9269
Coefficient γ_k in equation (12)	0.5404	0.5632	-0.3666
Mean of Operating Income	0.1630	0.1434	2.6093
Std of Operating Income	0.0738	0.0650	1.6970
AR(1) Coefficient of Operating Income	0.6368	0.5958	1.7120
Mean of Investment	0.0985	0.0979	0.0503
Std of Investment	0.0645	0.0722	-1.1291
AR(1) Coefficient of Investment	0.6149	0.6916	-2.2319
Mean of Net Equity Issuance	-0.0185	-0.0083	-4.2116
Std of Net Equity Issuance	0.0171	0.0045	1.7570
Mean of Market-too-Book	2.7443	2.3305	2.4253
Std of Market-to-Book	2.2892	2.4805	-1.3022
Mean of Dividend	0.0240	0.0242	-0.1568
Std of Dividend	0.0102	0.0109	-0.0701
SOA of Dividend	0.2040	0.2040	-0.0001
Mean Turnover	0.0263	0.0263	/
Corr btw Return and Turnover	-0.0674	/	/
		J-statistics : 54.80	P-value : <0.01

Panel B: Parameter Estimates

ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c
0.6741	0.2529	0.0837	0.5701	0.0917	0.1241	0.8562	0.4702	0.2239
(0.0396)	(0.0102)	(0.0228)	(0.0661)	(0.0183)	(0.0397)	(0.4087)	(0.1621)	(0.0853)

Table 7: Counterfactuals

Table 7 reports the model-predicted equilibrium firm value and dividend smoothness under different parameterizations: 25%, 50%, and 100% decreases in $\frac{W}{\kappa_I+\kappa_M}$, which captures the conflict of interest between shareholders and manager, and 25%, 50%, and 100% decreases in γ_Ω (while holding $\frac{W}{\kappa_I+\kappa_M}$ at 0), which controls, in equilibrium, how informative dividends are and to what extent they influence the market prices of shares.

Panel A: Degree of Dividend Smoothing			
	SOA	Total Diff (%)	Incremental Diff (%)
(1) Baseline	0.2047	/	/
(2) 25% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	0.2436	5.08%	5.08%
(3) 50% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	0.4376	30.41%	25.32%
(4) 100% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	0.4828	36.30%	5.90%
(5) 25% decrease in Γ_Ω	0.8307	65.73%	29.44%
(6) 50% decrease in Γ_Ω	0.8992	90.65%	24.91%
(7) 100% decrease in Γ_Ω	0.9708	100.00%	9.37%

Panel B: Equilibrium Firm Value			
	Firm Value	Total Diff (%)	Incremental Diff(%)
(1) Baseline	2.4120	/	/
(2) 25% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	2.4250	0.54%	0.54%
(3) 50% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	2.4428	1.28%	0.73%
(4) 100% decrease in $\frac{W_t}{(\kappa_I+\kappa_M)}$	2.4625	2.09%	0.81%
(5) 25% decrease in Γ_Ω	2.5130	3.36%	1.27%
(6) 50% decrease in Γ_Ω	2.5506	5.75%	2.39%
(7) 100% decrease in Γ_Ω	2.5896	7.37%	1.62%

Table 8: Subsample Estimation: Early versus Late Sample

Table 8 presents the estimation results obtained by splitting the sample based on time. The sample consists of non-financial and non-utility firms in the 2013 Compustat and CRSP file for the 1992–2011 period. The "Early" subsample includes all firm-year observations in or before 2002; the "Late" subsample consists of all firm-year observations from 2003 onwards. The estimation is done with simulated method of moments (SMM), which chooses structural parameters by matching the moments from a simulated panel of firms to the corresponding moments from the actual data. The J-statistic tests the over-identification constraint for the moment conditions within each subsample. Based on parameter estimates obtained from each subsample, the table reports the percentage of dividend smoothness driven by managers' career concerns following the algorithm in Table 7.

	Actual	Simulated	T-stat	Actual	Simulated	T-stat
Coefficient γ_y in Equation (12)	0.6909	0.6310	0.6624	0.5531	0.5797	-0.2905
Coefficient γ_k in Equation (12)	0.4027	0.4842	-1.0089	0.5578	0.5128	0.4873
Mean of Operating Income	0.1685	0.1603	1.0216	0.1445	0.1578	-1.8291
Std of Operating Income	0.0502	0.0692	-2.6359	0.0815	0.0827	-0.1527
AR(1) Coefficient of Operating Income	0.7462	0.6233	1.5921	0.6940	0.6492	0.6269
Mean of Investment	0.0869	0.0786	1.1874	0.1306	0.1339	-0.3826
Std of Investment	0.0591	0.0513	0.8500	0.0381	0.0484	-1.3906
AR(1) Coefficient of Investment	0.6524	0.5715	2.2283	0.4720	0.4782	-0.1479
Mean of Net Equity Issuance	-0.0176	-0.0157	-0.3211	-0.0212	-0.0209	-0.0619
Std of Net Equity Issuance	0.0175	0.0160	0.1688	0.0153	0.0162	-0.0619
Mean of Market-too-Book	2.7641	2.7614	0.0110	2.7230	2.7975	-0.3526
Std of Market-to-Book	2.4931	2.1405	2.2707	2.2998	2.3202	-0.1203
Mean of Dividend	0.0241	0.0255	-0.3058	0.0229	0.0247	-1.2456
Std of Dividend	0.0177	0.0159	0.1362	0.0184	0.0156	0.1722
SOA of Dividend	0.2163	0.2205	-0.0675	0.1664	0.1491	0.2395
Mean Turnover	0.0237	0.0221	1.0313	0.0283	0.0329	-1.9871
Corr btw Return and Turnover	-0.0869	-0.0887	0.0674	-0.0470	-0.0557	0.2964
J-statistics			25.53			12.01
% Career Concern-Based Dividend Smoothing			23.02%			37.98%

Table 9: Information Environment and Dividend Smoothing

In Table 9, I split firms into subsamples based on commonly-used information asymmetry measures. All variables in the table are self-explanatory, except for Abnormal Accruals, which is measured using the residuals from the modified Jones model, and Marginal q, which is constructed using the Erickson-Whited high-order moment estimator. Q1 corresponds to firms with the best information transparency, while Q5 includes firms with most opaque information. Panel A reports the actual Speed of Adjustment (SOA) of dividends; Based on the parameter estimates within each subsample, Panel B reports what fraction of the dividend smoothness is driven by managers' career concerns; Panel C predicts what the SOA of dividends would have been if managers' career concerns are eliminated. Corr(Info,·) is the correlation coefficient between the information measures and the indicated variable of interest in each panel.

Panel A: Speed of Adjustment (SOA) in the Data

Split By	Transparent Quintile 1	Q2	Q3	Q4	Opaque Quintile 5	Corr(Info,SOA)
Analyst Forecast Dispersion	0.2404	0.2003	0.2380	0.2214	0.1782	-0.7756
Marginal q	0.2394	0.2485	0.2764	0.2081	0.2357	-0.2731
Tangibility	0.2454	0.1781	0.1555	0.1750	0.2291	-0.1046
Institutional Holdings	0.1123	0.1888	0.1451	0.1618	0.2634	0.7411
Abnormal Accruals	0.2628	0.2129	0.2207	0.1657	0.2092	-0.6245

Panel B: Percentage of Career Concern-based Dividend Smoothing (%)

Split By	Transparent Quintile 1	Q2	Q3	Q4	Opaque Quintile 5	Corr(Info,%)
Analyst Forecast Dispersion	33.66 %	35.14 %	23.41 %	29.57 %	27.62 %	-0.4048
Marginal q	30.94 %	36.37 %	27.73 %	28.23 %	29.72 %	-0.4830
Tangibility	34.68 %	40.75 %	31.97 %	27.70 %	38.54 %	-0.1379
Institutional Holdings	48.18 %	35.69 %	39.18 %	21.69 %	27.81 %	-0.8564
Abnormal Accruals	30.94 %	33.72 %	27.73 %	28.23 %	22.72 %	-0.8463

Panel C: The Direct Information Effect on the Speed of Adjustment ($\widehat{\text{SOA}}$)

Split By	Transparent Quintile 1	Q2	Q3	Q4	Opaque Quintile 5	Corr(Info,$\widehat{\text{SOA}}$)
Analyst Forecast Dispersion	0.4984	0.4847	0.4223	0.4521	0.4084	-0.7614
Marginal q	0.4761	0.5267	0.4793	0.4376	0.4710	-0.5939
Tangibility	0.5082	0.5222	0.4270	0.4064	0.4854	-0.2977
Institutional Holdings	0.5469	0.4786	0.4864	0.3467	0.4713	-0.6484
Abnormal Accruals	0.4926	0.4788	0.4386	0.4028	0.4172	-0.8658

Table 10: Robustness

Table 10 reports the parameter estimates and the magnitude of the career concern-based dividend smoothing under alternative model specifications. The clustered standard errors for the estimated coefficients are reported in parentheses. In the first two robustness checks, I re-calculate the equilibrium allocation, using the parameter estimates from the baseline case and from the literature; In the remaining seven tests, I re-estimate the model parameters based on each alternative specification. α_0, α_1, & α_2 in (1) represent the fixed and convex capital adjustment costs, and the asymmetry in buying and selling prices. σ_ξ and ρ_ξ in (2) measure the persistence and standard deviation of the stochastic equity issuance cost. $\{f_0, f_1, f_2\}$ in (3) captures the fixed, linear, and convex costs for conducting repurchases. q in (8) captures the cost for earnings management. η_z in (7) captures the sensitivity of executive compensation to the perceived match quality, $\hat{z}_t$. c_p in (8) stands for the board's personal turnover cost.

												Career Concern-Based Dividend Smoothing
1. Lumpy investment												
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c	a_0	$a_1 \times 10^2$	a_2	
0.702	0.293	0.078	0.552	0.098	0.066	0.710	0.622	0.461	0.043	0.400	0.967	39.84%
2. Stochastic issuance cost												
ρ_z	σ_z	σ_s	θ	δ	ν_1	κ_M	c	σ_ξ	$\rho_\xi \times 10^3$			
0.702	0.293	0.078	0.552	0.098	0.066	0.622	0.461	0.400	0.967			42.89%
3. Linear-quadratic cost for repurchases												
ρ_z	σ_z	σ_s	θ	δ	ν_1	f_0	f_1	f_2	κ_M	c		
0.647	0.216	0.049	0.558	0.106	0.072	0.014	0.077	0.849	0.536	0.479		37.43%
(0.174)	(0.051)	(0.010)	(0.187)	(0.027)	(0.015)	(0.005)	(0.109)	(0.046)	(0.127)	(0.021)		(6.12%)
4. Earning management												
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c	q			
0.688	0.297	0.073	0.570	0.104	0.084	0.010	0.826	0.487	1.970			27.34%
(0.145)	(0.010)	(0.008)	(0.153)	(0.053)	(0.013)	(0.074)	(0.145)	(0.036)	(0.024)			(2.75%)
5. Executives' outside opportunities												
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c				
0.735	0.287	0.058	0.524	0.099	0.045	0.734	0.576	0.423				30.56%
(0.034)	(0.061)	(0.039)	(0.090)	(0.064)	(0.010)	(0.070)	(0.075)	(0.085)				(8.57%)

6. Executives' stock options										
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c		
0.683	0.333	0.068	0.579	0.141	0.077	0.757	0.587	0.431		34.12%
(0.050)	(0.083)	(0.026)	(0.122)	(0.043)	(0.018)	(0.144)	(0.086)	(0.031)		(4.10%)
7. Executives' risk aversion										
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c	η_z	
0.695	0.300	0.060	0.641	0.141	0.073	0.030	0.388	0.234	0.082	53.56%
(0.045)	(0.066)	(0.007)	(0.085)	(0.038)	(0.008)	(0.061)	(0.104)	(0.087)	(0.007)	(4.58%)
8. Board's entrenchment										
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c	c_p	
0.741	0.269	0.096	0.534	0.055	0.088	0.708	0.667	0.413	4.683	34.56%
(0.060)	(0.094)	(0.012)	(0.130)	(0.002)	(0.008)	(0.050)	(0.041)	(0.064)	(0.581)	(8.37%)
9.Asymmetric learning speed										
ρ_z	σ_z	σ_s	θ	δ	ν_1	ν_2	κ_M	c		
0.730	0.299	0.114	0.533	0.100	0.052	0.716	0.623	0.447		34.76%
(0.040)	(0.077)	(0.009)	(0.061)	(0.019)	(0.045)	(0.213)	(0.191)	(0.069)		(5.34%)