A Theory of Dividend Smoothing

Ilan Guttman  
Graduate School of Business  
Stanford University  
iguttman@stanford.edu

Ohad Kadan  
John M. Olin School of Business  
Washington University in St. Louis  
kadan@wustl.edu

Eugene Kandel  
School of Business Administration  
and Department of Economics  
Hebrew University  
mskandel@mscc.huji.ac.il

This version: September 2008

1We would like to thank three anonymous referees, Effi Benmelech, Sugato Bhattacharyya, Martin Cripps, Peter DeMarzo, Zsuzsanna Fluck, Paolo Fulghieri (the Editor), Armando Gomes, Milt Harris, Kose John, Avner Kalay, Praveen Kumar, Jeremy Stein, Jaime Zender, Jeff Zwiebel; seminar participants at Hebrew University, University of Houston, Rutgers University, Stanford University, Tel Aviv University, and Washington University in St. Louis; and participants of the Utah Winter Finance 2007, Gerzensee 2007, FIRS 2008, and WFA 2008 conferences for helpful comments and suggestions. Mike Borns and Janice Fisher provided excellent editorial assistance. Kandel expresses gratitude for financial assistance from the Kruger Center for Finance at Hebrew University.
Abstract

Dividend smoothing remains a puzzle for financial economists. We present a model in which smoothing of dividends arises as an equilibrium outcome. A manager who cares about the intrinsic value of the firm as well as its current stock price has to decide how to allocate earnings between investments and dividends. Since the stock price is determined by uninformed investors, the manager has an incentive to inflate dividends and lower investment relative to the first best. We show that there is a continuum of equilibria in which the dividend is constant for a range of realized earnings. Compared to the standard separating equilibrium, this partial pooling induces higher firm value, lower average dividends, and lower deviation from the first-best investment. We argue that the previous year’s dividend can serve as a “focal point,” allowing managers and investors to coordinate on just one in a continuum of equilibria. We conclude that dividend smoothing provides a partial remedy to underinvestment resulting from information asymmetries. We also offer several new testable predictions relating dividend smoothing to investors’ mix, managerial incentives, and investment.
1 Introduction

Dividends have long puzzled financial economists (see Allen and Michaely (2003) for an extensive survey). In this paper we focus on one aspect of this puzzle: dividend smoothing. Lintner (1956) interviewed managers from 28 companies and found that rather than setting dividends each year independently based on that year’s earnings, they first decide whether to change dividends from the previous year’s level. Managers claimed to reduce dividends only when they had no other choice, and increase dividends only if they were confident that future cash flows could sustain the new dividend level. Two beliefs were expressed strongly: that investors put a premium on companies with stable dividends, and that markets penalize firms that cut dividends. Furthermore, Lintner found that managers were setting the dividend policy first, while adjusting other cash-related decisions to the chosen dividend level. Almost fifty years later, in a survey of 384 financial executives, Brav, Graham, Harvey, and Michaely (2005) found that similar considerations still play a dominant role in determining dividends in publicly traded firms. By contrast, Michaely and Roberts (2007) found that dividend smoothing is significantly less likely in private firms.

For the purposes of this paper, we define dividend-smoothing as keeping the dividends per-share constant over two or more consecutive years. This definition is stronger than the one implied by Lintner’s paper, which requires only that the variation in dividends is lower than the variation in earnings. Figure 1 illustrates the dividend-smoothing phenomenon based on our definition. The figure shows the distribution of the proportions of annual dividend changes over the 40-year period 1966-2005 for all Compustat firms that paid dividends over two consecutive years.\(^1\) The figure shows an over 25% mass-point at zero change.\(^2\)

While dividend smoothing is an empirical regularity, it is quite perplexing theoretically. Why would well-diversified investors value a smoothed stream of dividends? After all, any idiosyncratic risk of dividend changes can be diversified away. Why do investors view dividend cuts as bad news in excess of what they have learned from the reported cash flows and earnings? Why don’t managers adjust dividends frequently to better reflect the level of current earnings? Why are changes in cash flows and

\(^1\)Note that over this time period the proportion of dividend-paying firms out of the total number of listed firms has significantly declined. See Fama and French (2001).

\(^2\)In the figure, dividend per share (adjusted for splits and other changes in the number of shares) is calculated as the ratio between Compustat items 26 and 27.
earnings reflected in investments and capital structure, but only weakly in dividend changes? Finally, why is dividend smoothing prevalent in public firms, and not in private firms?

In this paper we attempt to address these questions. We present a theoretical model showing that dividend smoothing can evolve endogenously as an equilibrium outcome. The key is that keeping dividends constant from one year to another while earnings have changed implies that firms follow a “partially pooling” dividend policy: the same dividend is announced for a range of different earnings realizations. Figure 1 provides good intuition for this pooling behavior. The only way in which a mass-point at zero dividend change can occur is if different realizations of earnings result in the same dividend being paid; that is, there is a pooling dividend that equals last year’s dividend. We show that this pooling behavior alleviates underinvestment problems that arise due to information asymmetries between managers and investors.

Our setup is similar to Miller and Rock (1985). An informed manager chooses the allocation of the available earnings between investments and dividend payout.\textsuperscript{3}

\textsuperscript{3}Similar to Miller and Rock (1985) and Kumar (1988), the term “dividends” in the model refers to any type of payout net of financing. We discuss this issue further below.
The total earnings and the investment are private information and cannot be credibly conveyed to investors; thus the observed dividend payment serves as a signal. The manager cares about the short-term stock price in addition to its long-term (intrinsic) value. Such objective functions appear to be well documented in reality. While we assume this objective function as given exogenously, we also propose several scenarios under which it can evolve in an endogenous manner. Linking the manager’s compensation to short-term stock price induces her to raise the dividends in order to signal higher earnings, resulting in underinvestment relative to the first-best level.

In the equilibria studied in Miller and Rock (1985), the dividend reveals all the private information of the manager to investors. These perfectly separating equilibria are inefficient: the manager overpays dividends and underinvests, and yet receives no informational rents. In our setting, while a unique separating equilibrium exists, it is just one of a multitude of equilibria. Our focus is on more efficient equilibria that are consistent with dividend smoothing over time.

The equilibria we consider are partially pooling: they have full revelation of earnings for very low and very high outcomes, but for all “intermediate outcomes” the same dividend is announced. Thus, for a wide range of earnings realizations the manager chooses exactly the same dividend. We show the existence of a continuum of such partially pooling equilibria. All these equilibria Pareto-dominate the fully separating equilibrium. Both the manager and the investors benefit from dividend pooling. The manager prefers any of the partially pooling equilibria to the standard fully revealing equilibrium, since the investment in the former is closer to first-best. Investors are assumed to be fully diversified, and thus effectively risk-neutral. As a result, the stock price reflects, on average, the increased firm value resulting from the more efficient investment level in the partially pooling equilibrium.

Next we discuss how the multiplicity of partially pooling equilibria in a static model is translated into dividend smoothing over time. The challenge for the manager and the investors is to coordinate on just one equilibrium out of a continuum. We first show that our static results can be generalized to a dynamic setting with myopic managers and investors. We then argue that the previous year’s dividend is a

---

4 These equilibria are somewhat similar to those studied in Harrington (1987) in the context of limit pricing; in Bernheim (1994) in the context of conformity, customs, and fads; and in particular in Guttman, Kadan and Kandel (2006) in the context of accounting earnings management.

5 If investors are risk averse then the Pareto-dominance result becomes more subtle. In that case, partial revelation of information exposes investors to more risk, which has to be weighed against the lower underinvestment.
natural candidate for a “focal point” that enables investors and managers to coordinate on just one equilibrium. All players coordinate on the partially pooling dividend strategy that yields a pooling dividend equal to the dividend paid in the previous year. Deviations from the last-year dividend are observed only if the earnings are either too low or too high, so that the last-year dividend can no longer be supported. This is consistent with Lintner’s findings. Thus a partially pooling dividend policy may yield smoothing of dividends over time. We use a detailed example to illustrate this process.

To summarize: our argument has three steps. First, we note that keeping dividends constant when earnings have changed implies some pooling behavior. We thus show that in a Miller-Rock-type model there exists a continuum of partially pooling equilibria. Second, we show that these equilibria Pareto-dominate the separating one, and therefore are more likely to prevail. Finally, we argue that managers and investors choose the last-year dividend to coordinate on one of these equilibria. This combination predicts that dividends, once announced, persist over time, until the earnings change to the extent that they no longer support the smoothed dividend. Then the dividend is cut or increased, and the process starts over again.

Our model offers several testable implications. First, we show that adverse selection and stock-based compensation are important determinants of dividend smoothing. Hence, dividend smoothing is more likely in public firms, as shown by Michaely and Roberts (2007). Furthermore, our model shows that dividend smoothing is associated with measures of managerial myopia such as short-term incentives. The model also suggests that better investment opportunities result in more dividend smoothing. Moreover, periods of smoothing are associated with higher investment and are followed by periods of higher profitability. We also study an extension that admits retained earnings and offers insights on the correlations between dividends, investments, and retained earnings in the presence of dividend smoothing.

A shortcoming of the Miller-Rock setup is that it does not enable us to formally distinguish between different types of payout. Our intention, however, is to focus on dividends and not consider stock repurchases, which are an alternative form of payout (e.g., Grullon and Michaely (2002)). The empirical literature does distinguish between the two. For example, Jagannathan, Stephens, and Weisbach (2000) show that stock repurchases are primarily used to pay out transitory, non-operating cash flows (although this pattern may be changing), whereas dividends are paid out of
operating cash flows, which are the focus of the model. Additionally, dividends are typically announced and paid on a regular basis, and a dividend announcement represents a binding commitment. In contrast, stock repurchase programs are announced occasionally, and offer just an option to repurchase: they do not commit the firm. Thus, while repurchases and dividends are often viewed as close substitutes, for the purposes of our study they are not, as they are associated with different kinds of cash flows, and their information content is different.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 develops the basic setup and presents the separating equilibrium in which no smoothing occurs. In Section 4 we develop and discuss our newly suggested partially pooling equilibria. Section 5 discusses dividend smoothing over time, and how it evolves from the partially pooling equilibria. Section 6 presents extensions to the model. The empirical predictions are discussed in Section 7. Section 8 concludes. Proofs are in the appendices.

2 Related Literature

Our model adds to the theoretical literature on costly dividend signaling. Important early contributions to this literature are Bhattacharya (1979), John and Williams (1985), Miller and Rock (1985), Ofer and Thakor (1987), Ambarish, John, and Williams (1987), Bernheim (1991), and Hausch and Seward (1993). These papers use perfectly separating equilibria, and do not consider dividend smoothing.

John and Nachman (1987) and Kumar and Spatt (1987) present dynamic models with smoothing of dividends. John and Nachman study an inter-temporal signaling model in which dividends have an adverse tax treatment. In their model, optimal financing and dividend strategies are determined by an endogenous structure of signaling costs, which is related to a permanent component of earnings. They show that many aspects of dividend smoothing emerge endogenously. Kumar and Spatt

---


7A well-known criticism of the dividend signaling literature is Dybvig and Zender (1991), who argue that the inefficiencies associated with asymmetric information can be avoided by optimal contracting. See Persons (1994) and Baranchuk, Dybvig, and Yang (2007) for the on-going discussion of this criticism.
(1987) analyze a dynamic model with risk-averse investors and firms whose investments are subject to serially correlated shocks. Firms in the model have private information on their prospects and idiosyncratic risk. The authors show that under certain conditions, firms will have incentives to smooth dividends and create a reputation for having low systematic risk. Our paper offers a different rationale for dividend smoothing, which relies on the efficiency of partially pooling equilibria.

Kumar (1988) is the first to point out the connection between dividend smoothing and partition pooling. He presents a signaling model in which managers and investors differ in their level of risk aversion. In his model, dividends serve as a coordination device between managers and investors, as in the “cheap talk” literature originated by Crawford and Sobel (1982). Therefore, in his model there is no separating equilibrium; the only equilibria are either completely pooling (“babbling”) or “step-function” equilibria. Our approach is different: we adhere to the costly signaling approach, starting from a standard model as in Miller and Rock (1985). We show that the separating equilibria, the focus of prior studies, form just a subset of all possible equilibria. Moreover, separating equilibria turn out to be inefficient compared to equilibria with pooling. This insight enables us to provide a new rationale for dividend smoothing.

Finally, Allen, Bernardo, and Welch (2000) suggest an explanation for dividend smoothing relying on the differences in taxation between individuals and institutions. In their model, taxable dividends attract informed institutions. Dividend reduction indicates a desire to reduce institutional ownership. Firms that benefit from institutional ownership avoid cutting their dividends.

3 Model

3.1 Basic Setup

We develop a two-period model in the spirit of Miller and Rock (1985). The earnings of a firm in period \( t = 1, 2 \), which are denoted by \( x_t \), are determined by the previous period investment and a random shock:

\[
x_t = F(I_{t-1}) + \varepsilon_t,
\]

\( \varepsilon_t \) are independent and identically distributed random variables.

Kumar and Lee (2001) offer a dynamic generalization of Kumar’s (1988) model.
where $I_{t-1}$ is the investment in period $t-1$, $F(\cdot)$ is a production function, and $\epsilon_t$ is a random shock. The initial investment $I_0$ is given exogenously, as is the dividend from the previous period $D_0$.

For simplicity, investments are assumed to be non-negative in each period ($I_t \geq 0$). The production function is assumed to be twice continuously differentiable, increasing, and concave: $F' > 0$ and $F'' < 0$. We further assume $\lim_{I_t \to 0} F'(I_t) = \infty$ and $\lim_{I_t \to \infty} F'(I_t) = 0$. The random shocks $\epsilon_t$ are drawn from normal distributions with means 0 and variances $\sigma^2_t$. Hence $\tilde{\epsilon}_t$ is normal with mean $F(I_{t-1})$ and variance $\sigma^2_t$ ($t = 1, 2$). We denote the densities of $\tilde{\epsilon}_t$ by $g_t$ and the cumulative distributions by $G_t$.

For simplicity, the random shocks are assumed to be uncorrelated: $\text{Cov}(\epsilon_1, \epsilon_2) = 0$.

At the beginning of period 1, the manager privately observes the realized earnings of the firm, $x_1$. At times, we refer to $x_1$ as the manager’s “type.” Given his type, the manager decides how to allocate $x_1$ between dividend payments ($D_1$) and investment ($I_1$). Since most of the “action” in the model is in period 1, we typically suppress the subscripts and use $I$ and $D$ instead of $I_1$ and $D_1$. Thus,

$$x_1 = I + D.$$  \hfill (1)

Investors are risk-neutral. They observe neither $x_1$ nor $\epsilon_1$, and use the dividend as a signal. All the parameters of the model are common knowledge. The firm is liquidated in period 2, and shareholders receive a liquidating dividend equal to $x_2$.

Given a period 1 dividend payment of $D$ and investment of $I$, the present value of the expected cash flows to the investors, which we refer to as the intrinsic firm value, is

$$V^I(x_1, D) = D + \frac{1}{1+i}[F(I) + E(\tilde{\epsilon}_2|x_1)] = D + \frac{1}{1+i}F(x_1 - D),$$  \hfill (2)

where $i \geq 0$ is an appropriate risk-adjusted discount rate.

The first-best dividend/investment decision is obtained if the manager seeks to maximize the intrinsic value. We denote the first-best investment level by $I^{FB}$. It equates the marginal return on investments with the inter-temporal opportunity cost:

$$F'(I^{FB}) = 1 + i \quad \text{or} \quad I^{FB} = F^{t-1}(1 + i).$$  \hfill (3)

---

9 It is easy to incorporate negative investment (asset sales) into the model.

10 Introducing positive correlation or even slightly negative correlation between the random shocks does not change the results.
Investors use the information contained in the dividend to price the stock. The market value of the firm at the end of period 1 is

\[ V^M(D) = D + \frac{1}{1+i}E(F(x_1 - D)|D). \] (4)

We assume that the manager is compensated based on both the stock price in period 1 and the liquidating value in period 2. The fact that some of this compensation vests in period 1 gives rise to a certain degree of managerial myopia (as in Stein (1989)). Thus, instead of maximizing the intrinsic value, the manager chooses dividend/investment to maximize a weighted average of the intrinsic value and the short-term market value:

\[ U(x_1, D) \equiv \alpha V^M(D) + \beta V^I(x_1, D), \] (5)

where \( \alpha, \beta > 0 \), and \( V^M \) and \( V^I \) are given by (4) and (2) respectively. The assumption that \( \alpha > 0 \) requires some justification, as it clearly diverts the manager away from the first-best investment level. For the ease of exposition we defer this discussion to Section 6, where we discuss ways to motivate and endogenize this objective function.

The manager has two conflicting interests: on the one hand he would like to boost the stock price by announcing a high dividend, resulting in lower than the first-best investment. On the other hand, he does not want to underinvest too much, because the marginal cost of underinvestment is increasing due to the concavity of \( F(\cdot) \).

To gain better intuition we will rewrite the manager’s objective function in a way that emphasizes the trade-offs he faces. First, note that (5) can be written as follows:

\[ U = \alpha \left( D + \frac{1}{1+i}E(F(I)|D) \right) + \beta \left( D + \frac{F(I)}{1+i} \right). \] (6)

In what follows it is useful to define a separate variable that captures the extent of underinvestment \( \Delta \equiv I^{FB} - I \). It is also useful to define a function \( h : \mathbb{R} \to \mathbb{R} \) capturing the real cost of underinvestment measured as the difference between the net present value under the first-best investment and the NPV under the actual investment,

\[ h(\Delta) \equiv \left( \frac{F(I^{FB})}{1+i} - I^{FB} \right) - \left( \frac{F(I)}{1+i} - I \right). \] (7)

The properties of \( F(\cdot) \) imply that \( h(\Delta) \) satisfies standard properties of loss functions: \( h(0) = 0, h'(0) = 0; h'(\Delta) > (\prec)0 \) iff \( \Delta > (\prec)0; h''(\Delta) > 0; \lim_{\Delta \to I^{FB}} h'(\Delta) = \infty. \)
That is, the first-best investment implies zero loss, while any deviation from first-best implies increasing marginal losses.

We can now rewrite the manager’s payoff function (6) as

\[ U = \alpha D - \beta h(\Delta) + B(D) + C(x_1), \]

where

\[ B(D) \equiv \frac{\alpha}{1+i} E(F(I|D)) \text{ and } C(x_1) \equiv \beta \left[ x_1 - I^{FB} + \frac{1}{1+i} F(I^{FB}) \right]. \]

The first component, \( \alpha D \), captures the manager’s direct benefit from the dividend, ignoring any informational effects. The second component, \(-\beta h(\Delta)\), is the real cost to the manager resulting from suboptimal investment, again ignoring informational effects. The third component, \( B(D) \), depends on investors’ beliefs about the investment given the dividend. The last component, \( C(x_1) \), depends neither on investors’ beliefs nor on the manager’s investment decision, but only on the realized earnings \( x_1 \).

A useful feature of the manager’s objective function is the Milgrom and Shannon (1995) single-crossing property.

**Definition 1** A utility function \( U : \mathbb{R}^2 \rightarrow \mathbb{R} \) satisfies the Milgrom–Shannon Single-Crossing Property (SCP) in \((x, D)\) if, for all \( x^H > x^L \) and \( D^H > D^L \), if \( U(x^L, D^H) \geq U(x^L, D^L) \), then \( U(x^H, D^H) > U(x^H, D^L) \).

That is, SCP is satisfied if, whenever a low type manager weakly prefers a high dividend to a low dividend, then a high type manager strictly prefers a high dividend to a low dividend. The following lemma is an immediate consequence of the convexity of \( h(\cdot) \).

**Lemma 1** For any given investors’ beliefs, the manager’s utility \( U(x_1, D) \) satisfies the Milgrom–Shannon SCP in \((x_1, D)\).

**Equilibrium Definition.** A dividend policy is a mapping \( \Lambda : \mathbb{R} \rightarrow \mathbb{R} \) assigning a dividend \( D = \Lambda(x_1) \) to any realization of period 1 earnings. Given any dividend \( D \), investors’ beliefs are a probability distribution over \( \tilde{x}_1 \).

A Perfect Bayesian Equilibrium is a combination of a dividend policy and investors’ beliefs such that:
1. For all \( x_1 \), \( \Lambda(x_1) \in \arg\max_D U(x_1, D) \), where the expectations conditional on dividend \( D \) are calculated using the investors’ beliefs.

2. Investors’ beliefs are consistent with \( \Lambda(\cdot) \) using Bayes rule, whenever applicable.

The following is an immediate implication of the single-crossing property.

**Lemma 2** In any equilibrium, the dividend policy is non-decreasing in manager’s type.

### 3.2 Separating Equilibrium

Observe first that a dividend policy that results in first-best investment is not an equilibrium.\(^{11}\) In a similar model, Miller and Rock (1985) show the existence of a multitude of separating equilibria. An equilibrium is separating if the manager’s type is always revealed given the dividend payment. Our setup differs from Miller and Rock’s only in that the normality of \( \tilde{\xi}_t \) implies an unbounded support of types. This assumption results in a unique, linear, and tractable separating equilibrium. The separating equilibria studied by Miller and Rock are non-linear and could not be given in a closed form. Tractability is important, as it allows us to derive analytically the main results regarding the partially pooling equilibria.\(^{12}\)

To derive the linear separating equilibrium we first conjecture its existence, and then verify that our conjecture is correct. As the first-best investment is independent of the realized earnings, we conjecture the existence of a linear equilibrium in which investment is fixed as well. Formally, we conjecture the existence of an equilibrium dividend policy of the form

\[
\Lambda(x_1) = x_1 - I^*,
\]

where \( I^* \) is a constant level of investment. If investment is fixed, then investors’

\(^{11}\)If \( \Lambda(x_1) = x_1 - I^PB \) for all \( x_1 \in \mathbb{R} \), then \( B(D) \) in (8) becomes a constant. Thus, both \( B(D) \) and \( C(x_1) \) are independent of the dividend/investment decision. It follows that the manager’s choice of dividend is governed by the trade-off between the first two terms in (8). The marginal benefit from underinvestment is \( \alpha > 0 \), while the marginal cost is \( \beta h'(\Delta) \). Since \( h'(0) = 0 \), the marginal benefit of underinvestment outweighs the marginal cost for all sufficiently small underinvestment levels.

\(^{12}\)The unbounded support is crucial because there is no “lowest” type. Such a type does not have an incentive to underinvest, and so the linearity of equilibrium in settings with a bounded support of types is precluded. We do not see why partially pooling equilibria similar to those derived in this paper would not exist also in environments with a bounded support of types. However, we are not able to derive such equilibria analytically in those settings.
beliefs about investment must reflect this, and \( B(D) \) becomes a constant denoted by

\[
B_s \equiv \frac{\alpha}{1 + i} F(I^*). \tag{10}
\]

Since \( C(x_1) \) also does not depend on \( D \), the manager trades off only the first two terms in (8). The first–order condition then gives

\[
h'(\Delta^*) = \frac{\alpha}{\beta}, \tag{11}
\]

where \( \Delta^* = I^{FB} - I^* > 0 \) is the optimal level of underinvestment.\(^\text{13}\) This completes the existence argument. We first conjectured the existence of a separating equilibrium in which dividend policy is linear and investment is constant. We then showed that under investors’ beliefs that derive from such a policy it is optimal for the manager to distribute a dividend that keeps investment (and underinvestment) constant. Showing that this linear equilibrium is the unique separating equilibrium in this model is a bit more subtle. The proof is in Appendix A. The next proposition summarizes these results.

**Proposition 1** There exists a unique separating equilibrium. It satisfies:

1. Dividend policy is linear and given by \( \Lambda_s(x_1) = x_1 - (I^{FB} - \Delta^*) \), where \( \Delta^* \equiv h^{-1}((\alpha/\beta)) \).

2. Investment is fixed and given by \( I^* = I^{FB} - \Delta^* \).

3. The dividend reveals the true type of the manager. Given a dividend of \( D \), investors assign probability 1 to the event that the manager’s type is \( D + I^* \).

The separating equilibrium yields some intuitive comparative statics properties.

**Corollary 1** In the separating equilibrium, underinvestment increases with \( \alpha \) and decreases with \( \beta \).

The proof is immediate, since \( h'(\cdot) \) is an increasing function and hence so is \( h^{-1}((\cdot)) \). Intuitively, the manager balances the marginal benefit from a higher short-term price with the cost of underinvestment. A higher weight on short-term stock-based compensation (higher \( \alpha/\beta \)) induces him to underinvest more.

\(^{13}\)Since \( h'(0) = 0 \) and \( \lim_{\Delta \to I^{FB}} h'(\Delta) = \infty \), the solution to (11) satisfies \( \Delta^* \in (0, I^{FB}) \).
4 Partially Pooling Equilibria

The separating equilibrium serves as a benchmark for our analysis. An important feature of this equilibrium is that different types always pay different dividends. We will seek a smoothed-out version of the separating equilibrium: unless earnings are very low or very high, all manager types pay the same dividend. Such partial pooling of dividends is necessary to generate smoothing over time: if the type space is continuous and the last-year dividend is announced with a positive probability, then there must exist a pooling interval of types.

We obtain such partial pooling by constructing an equilibrium in which there is an interval \([a, b]\) such that all manager types \(x_1\) in this interval announce the same dividend (the “pooling dividend”). Outside of this interval, managers follow a separating strategy. Investors’ beliefs must be consistent with this strategy: conditional on observing the pooling dividend they update their beliefs using Bayes rule given the information that the manager’s type falls in the interval \([a, b]\). Thus, while the pooling dividend provides information to investors by narrowing the set of potential manager’s types, this information is not perfect.

To establish such an equilibrium we modify the linear separating equilibrium derived in the previous section. Specifically, let \(\delta \geq 0\) be an arbitrary underinvestment level. We construct a continuum of partially pooling equilibria parametrized by \(\delta\). We refer to an equilibrium corresponding to a given \(\delta\) as a \(\delta\)-equilibrium. The dividend policy in each such \(\delta\)-equilibrium is

\[
\Lambda_\delta(x_1) = \begin{cases} 
    b_\delta - (I^{FB} - \delta) & x_1 \in [a_\delta, b_\delta] \\
    \Lambda_\delta(x_1) = x_1 - (I^{FB} - \Delta^*) & \text{otherwise}.
\end{cases}
\]  

(12)

That is, in each \(\delta\)-equilibrium there is a non-empty interval \([a_\delta, b_\delta]\) (the “pooling interval”) such that all manager types in this interval announce the same dividend:

\[D_\delta \equiv b_\delta - (I^{FB} - \delta).\]  

(13)

The bounds of the interval \(a_\delta\) and \(b_\delta\) will be determined endogenously. The pooling dividend \(D_\delta\) is the dividend announced by type \(b_\delta\) (the highest type in the pooling interval) when this type underinvests by \(\delta\). All types outside the pooling interval follow the separating dividend policy. The investment made by a type \(x_1 \in [a_\delta, b_\delta]\) is then

\[I = x_1 - D_\delta = x_1 - (b_\delta - (I^{FB} - \delta)),\]

and the underinvestment is \(b_\delta - x_1 + \delta\). Figure 2 illustrates the partially pooling dividend policy for a given \(\delta\).
To understand the intuition behind this dividend policy, consider first the case \( \delta = 0 \); that is, following the partially pooling dividend policy, type \( b_\delta \) invests at the first-best level. In equilibrium, type \( x_1 = b_\delta \) must be indifferent between separating and pooling. If this type chooses the separating dividend he underinvests by \( \Delta^* \), and the market value of the firm reflects his true type. If, on the other hand, this type chooses the pooling dividend, he saves the cost of underinvestment since he gets to invest \( I_{FB} \), but then investors pool him with lower types in the interval \([a_\delta, b_\delta]\). In equilibrium, the costs and benefits exactly offset each other, rendering this type indifferent. Next, consider type \( x_1 = a_\delta \), who must also be indifferent between the two choices. If he chooses the separating dividend he underinvests by \( \Delta^* \), and the market value reflects his true type. If, on the other hand, he chooses the pooling dividend, the market pools him with higher types in the interval \([a_\delta, b_\delta]\), but he has to underinvest more. In equilibrium, the two effects again offset each other, rendering type \( a_\delta \) indifferent.

A similar intuition applies to \( \delta > 0 \). In a partially pooling equilibrium, type \( x_1 = b_\delta \) is pooled with lower types. To compensate for this, the partially pooling equilibrium’s underinvestment must be lower than in the separating equilibrium: \( \delta < \Delta^* \). Similarly, following a pooling dividend, type \( x_1 = a_\delta \) is pooled with higher
types. To render him indifferent, the underinvestment given the pooling dividend is higher, i.e., \( b_\delta - a_\delta + \delta > \Delta^* \). This shows that a necessary condition for a partially pooling \( \delta \)-equilibrium is

\[
\Delta^* - (b_\delta - a_\delta) < \delta < \Delta^*. \tag{14}
\]

Investors’ beliefs must be consistent with the partially pooling dividend policy (12). Given the normal prior, the posterior following a pooling dividend \( D_\delta \) follows a truncated normal distribution over \([a_\delta, b_\delta]\).\(^{14}\)

Our goal is to show that a continuum of partially pooling equilibria exists. We first identify additional necessary conditions for \( \Lambda_\delta(\cdot) \) to be an equilibrium.

The utility of type \( x_1 = a_\delta \) given a separating dividend is

\[
U_s = \alpha (a_\delta - I^{FB} + \Delta^*) - \beta h(\Delta^*) + B_\delta + C(x_1). \tag{15}
\]

The utility of this type given a pooling dividend is

\[
U_\delta = \alpha (b_\delta - I^{FB} + \delta) - \beta h(b_\delta - a_\delta + \delta) + B_\delta + C(x_1),
\]

where

\[
B_\delta = \frac{\alpha}{1 + \epsilon} E(F(I|D_\delta)). \tag{16}
\]

By equating \( U_s \) and \( U_\delta \) we obtain the indifference condition for type \( a_\delta \):

\[
\alpha (a_\delta - I^{FB} + \Delta^*) - \beta h(\Delta^*) + B_\delta = \alpha (b_\delta - I^{FB} + \delta) - \beta h(b_\delta - a_\delta + \delta) + B_\delta. \tag{17}
\]

Similarly, the indifference condition for type \( b_\delta \) is

\[
\alpha (b_\delta - I^{FB} + \Delta^*) - \beta h(\Delta^*) + B_\delta = \alpha (b_\delta - I^{FB} + \delta) - \beta h(\delta) + B_\delta. \tag{18}
\]

Subtracting (17) from (18) gives

\[
\alpha (b_\delta - a_\delta) = \beta [h(b_\delta - a_\delta + \delta) - h(\delta)]. \tag{19}
\]

This simple necessary condition has an intuitive interpretation. In the separating equilibrium there is no difference between the two types in terms of underinvestment. In the partially pooling \( \delta \)-equilibrium both types obtain the same price, as investors cannot distinguish between them. Since both types are indifferent between pooling

\[^{14}\text{This means that for all } q \in [a_\delta, b_\delta], \text{ the posterior beliefs satisfy } \Pr(\tilde{x}_1 \leq q|\tilde{x}_1 \in [a_\delta, b_\delta]) = \frac{G(q) - G(a_\delta)}{G(b_\delta) - G(a_\delta)}\]

14
and separating, the difference between the price benefit in the separating equilibrium, which is $\alpha(b_\delta - a_\delta)$, must equal the difference in the cost of underinvestment in the partially pooling equilibrium, which is $\beta [h(b_\delta - a_\delta + \delta) - h(\delta)]$. We have,

**Corollary 2** A necessary condition for (12) to be a partially pooling $\delta$-equilibrium dividend policy is that $b_\delta - a_\delta > 0$ and $b_\delta - a_\delta$ satisfies Equation (19).

Equation (19) has one obvious solution: $b_\delta - a_\delta = 0$, which corresponds to the separating equilibrium. The convexity of $h(\cdot)$ implies that there may exist at most one additional solution in which $b_\delta - a_\delta$ is strictly positive, and hence pooling occurs with positive probability. We claim that there exists a continuum of $\delta$ values for which such a non-trivial solution exists. Denote the size of the pooling interval by $y_\delta \equiv b_\delta - a_\delta$. Then,

**Lemma 3** There exists a $0 \leq \Delta < \Delta^*$ such that for all $\delta \in [\Delta, \Delta^*)$ the equation

$$\alpha y_\delta = \beta [h(y_\delta + \delta) - h(\delta)]$$

(20)

has a unique solution with $y_\delta > 0$. Moreover, for all $\delta \in [\Delta, \Delta^*)$, $\frac{\partial y_\delta}{\partial \delta} < -1$ and Condition (14) is satisfied.

Figure 3 illustrates the idea behind this lemma by presenting solutions to Eq. (20) for three cases: $\delta > \Delta^*$, $\delta = \Delta^*$, and $\delta < \Delta^*$. When $\delta = \Delta^*$, the slope of the LHS of Eq. (20) at $y_\delta = 0$ is exactly equal to the slope of the RHS. Thus, for $\delta = \Delta^*$, the curve $h(y_\delta + \delta) - h(\delta)$ is tangent to the line $\alpha y_\delta / \beta$ at $y_\delta = 0$, and $y_\delta = 0$ is the only intersection of the two. When $\delta > \Delta^*$, the slope of the curve at $y_\delta = 0$ is strictly larger than the slope of the line, which again implies that $y_\delta = 0$ is the only solution of (20). When $\delta < \Delta^*$ but is sufficiently close to $\Delta^*$ there is one additional, non-trivial solution to (20), represented by the intersection point. Note that for $\delta < \Delta^*$, as $\delta$ becomes smaller the intersection point moves to the right, implying a larger pooling interval.

Given Lemma 3, from now on we restrict attention to $\delta \in [\Delta, \Delta^*)$, for which we have shown that a nontrivial pooling interval exists. For any given such $\delta$ the size of the pooling interval, $y_\delta$, is uniquely determined by (20), and does not depend on the location of the interval. We next show that there always exists a location for a

---

15 In the figure we divided both sides by $\beta$, so that the straight line is $\alpha y_\delta / \beta$, whereas the three curves depict $h(y_\delta + \delta) - h(\delta)$ for different $\delta$ values.
pooling interval $[a_\delta, b_\delta]$ of size $y_\delta$, such that investors’ beliefs given $x \in [a_\delta, b_\delta]$, render both types $a_\delta$ and $b_\delta$ indifferent between separating and pooling.\footnote{The issue here is that investors’ beliefs conditional on a pooling dividend change with the location of the pooling interval. One has to make sure that there is a location that generates beliefs that render both types indifferent.} The strategy for finding such a location is to “search” from the left tail to the right tail of the distribution. We show that at the left tail, the beliefs are such that both types $a_\delta$ and $b_\delta$ strictly prefer pooling to separating, while at the right tail both strictly prefer separating to pooling. Given continuity, there is an intermediate location where both types are indifferent.

**Lemma 4** For each $\delta \in [\Delta, \Delta^*)$ there exists an interval $[a_\delta, b_\delta]$ of size $y_\delta$ given implicitly by (20), such that the indifference conditions (17) and (18) are satisfied.

The proof of the lemma is in Appendix B.

So far we have focused on the necessary conditions that types $a_\delta$ and $b_\delta$ be indifferent between separating and pooling. In equilibrium much more is needed: all types must (weakly) prefer their equilibrium dividend policy to any other dividend. The key to our equilibrium is that the indifference of types $a_\delta$ and $b_\delta$ together with the single-crossing property and out-of-equilibrium beliefs imply that no other type deviates. This establishes sufficiency and existence.

Figure 3: Determinants of the Size of the Smoothing Interval
Note that given a partially pooling dividend policy $\Lambda_\delta (\cdot)$, the dividend payment never lies in the range $(a_\delta - I^*, b_\delta - I^{FB} + \delta) \cup (b_\delta - I^{FB} + \delta, b_\delta - I^*)$. To completely specify the equilibrium, we must define the out-of-equilibrium beliefs associated with these zero-probability dividends. For simplicity, we assume first that in observing such a dividend, investors believe that the manager is “mistakenly” playing the separating dividend policy.\textsuperscript{17} See Appendix C for a generalization.\textsuperscript{18} We are now ready to state our main existence result.

**Proposition 2** For any $\delta \in [\Delta, \Delta^*)$ there exists an interval $[a_\delta, b_\delta]$ such that $\Lambda_\delta (\cdot)$ as given in (12) is an equilibrium. Furthermore,

1. The size of the pooling interval is given implicitly by (19).

2. Given a dividend of $D_\delta = b_\delta - (I^{FB} - \delta)$ (the pooling dividend), investors’ beliefs follow a truncated normal distribution over $[a_\delta, b_\delta]$. For any other dividend $D$, investors assign probability 1 to the event that the manager’s type is $D + I^*$.

The proof is in Appendix A. This completes the first step of our argument: a continuum of partially pooling equilibria exists in our model.

### 4.1 Example

To illustrate the partially pooling equilibria assume the production function takes the form: $F(I) = 2\sqrt{I}$. Set parameter values to $\alpha = 0.7$, $\beta = 1$, $i = 0$, $I_0 = 0.5$, and $\sigma_1 = 0.2$. It is easy to verify that the first-best investment is $I^{FB} = 1.0$. The investment in the separating equilibrium is $I^* = 0.346$, and the underinvestment is $\Delta^* = 0.654$. Table 1 presents the results for the partially pooling $\delta$-equilibria for several $\delta < \Delta^*$ values. For each $\delta$ we first calculated the size of the pooling interval, $y_\delta$, by numerically solving the implicit equation in (20). Then, using a grid search we found $a_\delta$ and $b_\delta$ such that $b_\delta - a_\delta = y_\delta$, and the indifference conditions are satisfied. The pooling dividend, $D_\delta$, is then given by (13). Using the parameters of the distribution function we also calculated the probability of pooling: $\Pr \{x_1 \in [a_\delta, b_\delta]\}$.\textsuperscript{17}

\textsuperscript{17}That is, given an out-of-equilibrium dividend $D$, investors assign probability 1 to the event that the manager’s type is $D + I^*$.

\textsuperscript{18}In Appendix C we characterize the set of out-of-equilibrium beliefs supporting the $\delta$-equilibria. We also show the existence of monotonic out-of-equilibrium beliefs that support the $\delta$-equilibria, and discuss the robustness of these beliefs vis-à-vis standard refinement concepts.
The case $\delta = 0.654 = \Delta^*$ corresponds to the separating equilibrium: the pooling interval vanishes. Smaller values of $\delta$ induce a pooling interval with positive length. For instance, when $\delta = 0.3$, the size of the pooling interval is 0.586, which accounts for 2.9 standard deviations of the earning distribution. Any realization of earnings between 1.033 and 1.619 induces a pooling dividend of 0.919. The probability that realized earnings fall in this interval is 0.819.

### 4.2 Dominance of Pooling

Partially pooling equilibria are a necessary building block of dynamic smoothing. Furthermore, the welfare properties of these equilibria offer a new rationale for dividend smoothing. We show below that pooling of dividends, and consequently dynamic smoothing thereof, improves welfare by lowering the extent of underinvestment.

Our first result concerns the preferences of the manager. It states that the manager strictly prefers any one of the $\delta$-equilibria to the separating equilibrium. This follows from the fact that no type $x_1 \in (a_\delta, b_\delta)$ has an incentive to deviate from the partially pooling dividend policy given the out-of-equilibrium beliefs that were set as the separating dividend policy.

**Lemma 5** Let $\delta \in [\Delta, \Delta^*)$, and consider a corresponding $\delta$-equilibrium. For every realized earnings $x_1$ the manager weakly prefers the partially pooling $\delta$-equilibrium to the separating equilibrium. Moreover, for $x_1 \in (a_\delta, b_\delta)$, the preference is strict.

The next result shows that not only do managers prefer the partially pooling $\delta$-equilibria, but that it also increases the value of the firm. The ex-ante firm value is just the average ex-post firm value, conditional on the manager's information. Since

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$y_\delta$</th>
<th>$a_\delta$</th>
<th>$b_\delta$</th>
<th>$D_\delta$</th>
<th>$\text{Pr} { x_1 \in [a_\delta, b_\delta] }$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.854</td>
<td>0.862</td>
<td>1.716</td>
<td>0.816</td>
<td>0.931</td>
</tr>
<tr>
<td>0.2</td>
<td>0.727</td>
<td>0.946</td>
<td>1.673</td>
<td>0.873</td>
<td>0.892</td>
</tr>
<tr>
<td>0.3</td>
<td>0.586</td>
<td>1.033</td>
<td>1.619</td>
<td>0.919</td>
<td>0.819</td>
</tr>
<tr>
<td>0.4</td>
<td>0.442</td>
<td>1.111</td>
<td>1.553</td>
<td>0.953</td>
<td>0.691</td>
</tr>
<tr>
<td>0.5</td>
<td>0.283</td>
<td>1.189</td>
<td>1.472</td>
<td>0.972</td>
<td>0.483</td>
</tr>
<tr>
<td>0.6</td>
<td>0.106</td>
<td>1.227</td>
<td>1.333</td>
<td>0.933</td>
<td>0.167</td>
</tr>
<tr>
<td>0.654</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Examples of the Partially Pooling Equilibrium
the manager is strictly better off in the partially pooling equilibrium, it must be that the ex-ante firm value is higher. But this in turn is possible only if investment moved closer to the first-best level. Thus, both the manager and the investors ex-ante prefer pooling to perfect separation. Formally,

**Proposition 3** Set $\delta \in [\Delta, \Delta^*)$, and consider a corresponding $\delta$-equilibrium. Then,

1. The expected intrinsic value is higher under the partially pooling $\delta$-equilibrium than under the separating equilibrium.

2. The expected underinvestment and expected dividends are lower under the partially pooling $\delta$-equilibrium than under the separating equilibrium.

3. The $\delta$-equilibrium generates higher expected investment and higher expected return on investment than the separating equilibrium.

4. Investors ex-ante prefer the $\delta$-equilibrium to the separating equilibrium.

5. The $\delta$-equilibrium Pareto-dominates the separating equilibrium.

Proof: In Appendix A.

Pooling enables the manager to retain some information rents. Compared to the separating equilibrium, the additional uncertainty introduced by the noisy pooling dividend does not pose a problem to fully diversified (risk-neutral) investors, since the firm is still correctly priced on average. Furthermore, in the $\delta$-equilibrium the investment is, on average, closer to its first-best level, increasing the firm value to the investors. Thus, both the manager and the investors benefit from the pooling. As any one of the partially pooling equilibria dominates the separating equilibrium, the latter does not survive the equilibrium selection criterion of Maskin and Tirole (1992).

This completes the second step of our argument. Any one of the partially pooling equilibria Pareto-dominates the fully revealing one, which provides motivation for why a partially pooling equilibrium may be played. Next we discuss how a partially pooling equilibrium in a single period translates into dividend smoothing over time.

### 5 Dynamic Dividend Smoothing

As discussed in the introduction, dividend smoothing manifests itself as the practice of keeping dividends constant over two or more consecutive periods. This is a stronger
notion of smoothing than the one advocated in Lintner (1956), which requires that the variation in dividends is lower than the variation in earnings. An immediate implication is that some pooling must be present. In this section we outline a simple dynamic version of the model and demonstrate how coordination on one of the partially pooling equilibria can result in dividend smoothing over time.

5.1 Simplified Dynamic Setting

Assume a finite number of periods \( t = 1, ..., T \). Earnings in each period are \( x_t = F(I_{t-1}) + \varepsilon_t \), where \( I_t \) is the investment at time \( t \) and \( \varepsilon_t \) is normally distributed with mean 0.\(^{19}\) At each time period the manager has to choose how to allocate the earnings \( x_t \) between a dividend payment \( D_t \) and investment \( I_t \). The initial investment and dividend \( I_0 \) and \( D_0 \) are given exogenously.

Solving the dynamic model presents two technical challenges:

1. Once earnings fall in the pooling interval in period \( t - 1 \), the distribution of \( x_t = F(I_{t-1}) + \varepsilon_t \) from the investors’ point of view is the sum of two random variables: one has a finite support and the other is normal. This sum is not normally distributed. This raises technical difficulties, as our existence proof makes use of the tail properties of the normal distribution.

2. Allowing the manager and the investors to take into account the effect of the dividend choice in period \( t \) on future earnings in all upcoming periods undermines the linearity of the separating equilibrium, making the related partially pooling equilibria intractable.

The first problem has a manageable solution. We show that while the distribution of earnings in periods 2, ..., \( T \) may not be normal, it has the same tail properties as the normal distribution. This enables us to use similar techniques to those used in the static setting to prove existence of equilibria in a dynamic setup. To overcome the second problem, we assume (similar to Stein (1989)) that both the manager and the investors are myopic: every period \( t \) they care only about the trade-off between the dividend paid in period \( t \) and the expected earnings in period \( t + 1 \).\(^{20}\) That is, the objective of the manager is to maximize a linear combination of the stock price at

\(^{19}\)The production function is assumed identical for all periods 1, ..., \( T \); however, it is easy to generalize the model to a time-varying production function.

\(^{20}\)Given this assumption, the model can also accommodate an infinite number of periods.
time \( t \) and the intrinsic value, under the assumption that the firm will be liquidated at time \( t+1 \). Formally, the manager maximizes in period \( t \)

\[
U_t(x_t, D_t) \equiv \alpha \left( D_t + \frac{1}{1+i} E(F(x_t - D_t) | D_1, D_2, ..., D_t) \right) + \beta \left( D_t + \frac{1}{1+i} F(x_t - D_t) \right).
\]

(21)

This assumption enables us to restore the tractability of the partially pooling equilibria.\(^{21}\)

It is important to note that this myopia does not transform the dynamic game into a sequence of separate “one shot” games. The reason is that investment is carried from one period to the next. If a pooling strategy is played in period \( t \), investors are uncertain about the amount of investment, and this affects their beliefs and pricing in period \( t+1, t+2 \), etc. This is reflected in (21), as investors condition their expectations in period \( t \) on the entire history of dividends up to this point in time.

One obvious equilibrium in this myopic, dynamic setting is to play the separating equilibrium outlined in Proposition 1 in each period. Next, we show the existence of a continuum of equilibria with partial pooling in this setting.

**Proposition 4** Assume that the manager and the investors are myopic (as explained above). There exist \( \Delta < \Delta^* \), a continuum of vectors \((\delta_1, ..., \delta_T) \in [\Delta, \Delta^*]^T\), and corresponding intervals \([a_{\delta_1}, b_{\delta_1}], ..., [a_{\delta_T}, b_{\delta_T}]\) such that the dynamic dividend policy

\[
\Lambda_{\delta_t}(x_t) = \begin{cases} 
  b_{\delta_t} - (I^{FB} - \delta_t) & \text{if } x_t \in [a_{\delta_t}, b_{\delta_t}] \\
  \Lambda_s(x_t) = x_t - (I^{FB} - \Delta^*) & \text{otherwise}
\end{cases}
\]

for \( t = 1, ..., T \)

with appropriate investors’ beliefs constitute an equilibrium in the dynamic game. Any of these equilibria Pareto-dominates the equilibrium in which the separating dividend policy is played for all \( t = 1, ..., T \).

Proof: In Appendix D.

**5.2 Coordination and Smoothing**

Though our dynamic setting shows the existence of a continuum of partially pooling equilibria, it is mute on how investors and the manager coordinate on just one out

\(^{21}\)We emphasize that this assumption is made for tractability reasons only. While we are not able to solve analytically for a partially pooling equilibrium in the full-fledged dynamic model, we do not see why the dynamic smoothing identified below would not work in such a setting.
of these equilibria. In fact, the game theoretic literature has very little to say about
equilibrium selection. The dominance of any partially pooling equilibrium over the
separating one, suggests that managers and investors will try to coordinate on one
of the partially pooling equilibria. However, we cannot rank the partially pooling
equilibria in terms of their efficiency.\footnote{We have shown (Lemma 3) that the size of the pooling interval decreases in $\delta$. The location of
the interval also varies with $\delta$, but in a non-monotone manner that we cannot track analytically.}

One way to generate coordination is to use the last-year dividend as a “focal
point.” That is, in each year $t = 1, \ldots, T$, the manager and the investors coordinate
on playing a partially pooling $\delta_t$-equilibrium in which the pooling dividend in year $t$
equals the dividend paid in year $t - 1$. Then, there is an interval of earnings $[a_{\delta_t}, b_{\delta_t}]$
such that for all earnings therein, the manager announces a dividend that equals the
last-year dividend. This gives rise to dividend smoothing over time. Only if earnings
are outside this interval is the dividend different from the last-year dividend, in which
case it fully reveals the earnings.

We illustrate this by an example using the myopic-dynamic setting developed
above. Assume the game is played for $T = 3$ years. The production function is
$F(I_{t-1}) = 2\sqrt{I_{t-1}}$ with parameter values as in Table 1. The initial investment is
$I_0 = 0.5$, and the initial dividend is $D_0 = 0.82$. Table 2 presents the results of a path
of realizations of $x_t$ and equilibrium play using coordination in dividends.

In year 1, $F(I_0) = 2\sqrt{0.5} = 1.41$. Hence, earnings in year 1 are distributed
normally with mean 1.41 and S.D. 0.2. We use these parameters to find a $\delta_1$ value,
such that the pooling dividend $D_{\delta_1}$ is equal to the dividend in the previous period,
namely, 0.82. The result is $\delta_1 = 0.11$. We then assume that the manager and the
investors coordinate on playing the partially pooling equilibrium that corresponds
to $\delta_1 = 0.11$. We numerically find the pooling interval in this case, which turns
out to be $[0.87, 1.71]$. Thus, the size of the pooling interval is $y_{\delta_1} = 0.84$, and the
probability of dividend smoothing is 0.93. The earnings realization in this case is
1.45, which falls inside the pooling interval. As a result, the dividend announced is
$D_1 = \Lambda_{\delta_1}(1.45) = 0.82$, which is equal to $D_0$.\footnote{Had the separating equilibrium been played, the dividend would have been increased to
$\Lambda_s(1.55) = 1.21$.}

The investment in year 1 is $I_1 = x_1 - D_1 = 1.55 - 0.82 = 0.73$, and the expected
return on investment is $F(I_1) = 1.7$. Consequently, the earnings in year 2, are
normally distributed with mean 1.7 and S.D. 0.2. The investors do not know at this

22
23
point what the investment was in period 1, since they only saw a pooling dividend. Hence, from the investors’ point of view, the year 2 earnings follow a distribution that is the sum of two random variables: one is drawn from a distribution over the pooling interval $[0.87, 1.71]$, and the other is normal noise with mean 0 and S.D. 0.2. This sum is calculated as the convolution of the two distributions (see Appendix D). As before, using a grid search we find a $\delta_2$ such that the pooling dividend is equal to 0.82. This turns out to be $\delta_2 = 0.25$. The new pooling interval is $[0.91, 1.57]$ with probability 0.71. The earnings realization is 1.50, which falls inside the pooling interval. As a result, the dividend is kept at $D_2 = 0.82$.

The process is repeated in period 3. The partially pooling equilibrium, for which the pooling dividend is equal to last year’s dividend, turns out to be the equilibrium associated with $\delta_3 = 0.52$. The pooling interval here is $[1.04, 1.29]$. The realized earnings this time are 1.55, which falls outside of the pooling interval. As a result, the dividend is increased to the separating level of $x_t = 1.55 - 0.35 = 1.2$.

This example illustrated how the partial pooling of dividends translates into smoothing over time. Dividend is smoothed in this example between the periods 0, 1, and 2 as earnings fall inside the corresponding pooling intervals (despite the growth in earnings). The dividend is first changed in period 3 after earnings have reached a level of 1.55, which falls outside of the pooling interval in that period.

### 5.3 Discussion

Dynamic smoothing of dividends implies that some manager types pool and pay out the last-year dividend. To model this we make a three-step argument. First, we show that in a Miller–Rock type model there exists a continuum of partially pooling equilibria. Second, we show that the partially pooling equilibria Pareto-dominate the fully revealing one. Thus, we argue that the pooling equilibria are likely to be observed, provided that the investors and the managers can coordinate on a particular
one. Finally, we argue that managers and investors can use the last-year dividend to coordinate on one of these equilibria. This coordination device is simple and, once established, does not require year-to-year adjustments. This combination predicts that dividends, once announced, persist over time, until the earnings change to the extent that they fall outside of the equilibrium pooling region. Then the dividend is cut or increased, and the process starts again.

Ultimately, the claim that dividend smoothing is a result of coordination should be empirically tested. While a serious empirical study of this question is beyond the scope of this paper, we present below some illustrative evidence that seems consistent with the notion of coordination. In Figure 1 we have shown that the unconditional probability of a firm not changing its dividend from one year to the next during 1966-2005 is over 25%. Are all firms equally likely to engage in this practice, or are some firms more prone to do so persistently, as would be expected from a coordination story? We calculated that 38% of all firms that ever paid dividends two years in a row between 1966 and 2005 never smoothed dividends even once in those 40 years. This already suggests that not all firms are playing the smoothing game. Furthermore, Figure 4 presents the probability of not changing the dividend per share from year $t$ to year $t+1$, unconditionally, as well as conditional on dividend change (or no change) from year $t-1$ to year $t$. Conditional on not changing their dividend in the previous year (thin line), around 60% of firms do not change their dividend the following year either. Conditional on changing the dividend in the previous year (dashed line), only slightly more than 10% keep the dividend constant in the following year. These simple statistics suggest that there exists a sub-sample of dividend-paying firms that consistently engage in dividend smoothing, while other firms do not. To us, this suggests that the former firms had adopted the last-year dividend as a coordination device.

6 Extensions

In this section we explore two extensions to the base model. First we study the structure of the manager’s objective function, and identify scenarios under which such an objective function can evolve endogenously. Then we explore how allowing the firm to retain some of the earnings affects our results. Throughout this section we use the static version of the model. Adapting the results to the dynamic-myopic
setup in Section 5 is straightforward.

6.1 Structure of the Manager’s Objective Function

Similar to the Miller-Rock model, our model relies on an exogenous managerial objective function, which assigns positive weights to both short-term and long-term stock price. Given the evolution of the mechanism design literature, it is reasonable to ask under which conditions such an objective function can emerge. First note that in the absence of additional frictions, the objective function (5) is sub-optimal. To see this, assume that the payment to the manager can be a function of the first-period stock price $V^M$, the first-period dividend $D_1$, and the second-period liquidating dividend $x_2$. The compensation contract is thus a function $\pi(V^M, D_1, x_2)$. Long-term investors, who are assumed to be the principal, maximize the expected intrinsic value

---

24 For example, if cash flows are private information and there is no investment, then we know from Townsend (1979) and Gale and Hellwig (1985) that the optimal mechanism is a debt. DeMarzo and Fishman (2007) show in a dynamic setting that with privately observed cash flows and investment, both equity and debt are optimally used in the agent’s compensation.
$EV^I$ net of expected payment to the manager:

\[
\max_{\pi, D_1} EV^I - E\pi(V^M, D_1, x_2) \\
\text{s.t.} \\
E\pi(V^M, D_1, x_2) \geq u_0 \quad (IR) \\
D_1 \in \arg\max_D E\pi(V^M, D, x_2) \quad (IC)
\]

That is, investors choose the contract $\pi$ and the dividend level paid for any realization of earning, subject to the constraints that (i) the expected payoff to the manager exceeds a reservation utility $u_0$; and (ii) the dividend is incentive-compatible.\textsuperscript{25} In this setting, giving the manager equity that vests in period 2 implements the first-best. Indeed, set $\pi = \beta V^I = \beta \left( D_1 + \frac{x_2}{1 + i} \right)$, where $\beta$ is chosen so that $IR$ holds with equality. Then the manager chooses the first-best investment and dividend levels, and the allocation is Pareto optimal. Put differently, granting the manager “restricted stock” by setting $\alpha = 0$ and $\beta > 0$ in (5) resolves the agency problem.\textsuperscript{26}

Despite its optimality in the above setting, the case $\alpha = 0$ does not conform with real-world compensation schemes. Typical stock-based compensation grants vest over a period of 3-4 years in equal installments. As a result, at any given point in time, managers are commonly exposed to both long-term and short-term stock-based incentives.\textsuperscript{27} The obvious question then is how one can reconcile the ubiquitous use of short-term stock-based incentives with the inefficiencies that they appear to create.

To answer this question one has to depart from the specific setup studied in this paper and consider additional frictions, such as managerial risk aversion, job-market considerations, and features of the market for corporate control. Below we discuss such setups under which the case $\alpha > 0$ is likely to evolve.

A first natural friction is managerial risk-aversion. Managers are likely to be more risk-averse in their firm-related decisions than well-diversified investors. Bhattacharyya and Cohn (2008) use this to show that short-term stock-based compensation should optimally be included in compensation contracts. They argue that

\textsuperscript{25}Note that in the $IR$ constraint the expectation is taken at the contracting stage before $\tilde{x}_1$ or $\tilde{x}_2$ is realized. The expectation in the $IC$ constraint is taken after $\tilde{x}_1$ but before $\tilde{x}_2$ is realized.

\textsuperscript{26}Note that this is not the unique optimal contract. For example, any contract of the form $c + \beta V^I$, where $c$ is a constant set to meet the reservation utility, implements first-best.

\textsuperscript{27}See for example Murphy (1999) and Bebchuk and Fried (2003).
short-term stock-based compensation mitigates the reluctance of risk-averse managers to pursue risky and profitable projects. Linking manager’s pay to short-term realizations provides managers with insurance against bad outcomes in the long run. This induces managers to take risky and profitable bets. Bhattacharyya and Cohn conclude that optimal compensation contracts include both short-term and long-term stock-based compensation.

Stein (1989) provides a different explanation. He argues that the firm may face early takeover or early funding requirements. As a result, even if the explicit managerial contract relies solely on long-term stock price, the manager’s objective function \textit{de facto} depends on the short-term price as well.

Career concerns can also play a role. Consider for example a setting in which managers differ in terms of their ability, where higher ability is associated with higher expected earnings. Then the stock price is informative about both investments and managerial ability. Thus, career-minded managers assign positive weights to both short- and long-term incentives. Additionally, outside options for the manager are likely to be correlated with the short-term stock price. Therefore, linking the manager’s pay to short-term price realizations may help retain the manager.

Based on the above arguments, we feel that $\alpha > 0$ is a reasonable assumption and a fair description of reality.

6.2 Retained Earnings

It is tempting to interpret (1) as implying that in the Miller-Rock model, dividends and investments are perfectly negatively correlated. Such a conclusion would contrast empirical regularities suggesting no such relation (e.g. Fama 1974). Note, however, that the model does not have any defined implications for the correlation between dividends and investments, since (1) does not imply that higher earnings are associated with lower investments.\footnote{Lemma 2 shows that dividends (weakly) increase in earnings. Investments, however, may increase, decrease, or stay constant when earnings increase. This prevents a simple signing of the correlation between dividends and investments.} As an example, in the separating equilibrium investments are constant and thereby uncorrelated with dividends. In the partially pooling equilibria, outside of the pooling interval investment is constant and dividends increase in earnings. Inside the pooling interval, investment increases in earnings while the dividend is constant. As a result, the correlation between investments and dividends
in these equilibria is quite low, in accordance with the evidence.

A natural question is whether the correlation structure between investments and dividends in the partially pooling equilibria is an artifact of the inability of managers in the model to retain earnings. To explore this issue we outline below a simple extension that allows for retained earnings.

Assume that the earnings in period 1 are allocated between dividends $D$, investments $I$, and retained cash/earnings $R$. That is, replace (1) with

$$x_1 = I + D + R.$$  \hspace{1cm} (23)

Assume that retained earnings are used by the firm for operating activities, working capital, and short-term investments. It is reasonable that the benefit from retained earnings $B(R)$ is a concave function with similar properties to those of the production function $F(\cdot)$. This implies that the first-best level of retained earnings $R^{FB}$ is given implicitly by

$$B'(R^{FB}) = 1 + i.$$  \hspace{1cm} (24)

If retained earnings were observable by investors, then the manager would always set them at the first-best level, and the results stay qualitatively the same as in the base model. A more interesting case is when retained earnings, similar to investments, are not perfectly observable by investors. Here, the dividend serves as a signal for both investments and retained earnings. Given this structure, for any earnings level, the manager has to choose the dividend, investment, and retained earnings. Since the manager has complete leeway in optimizing between investments and retained earnings, in equilibrium, the marginal productivity of investments equals the marginal benefit from retained earnings:

$$F'(I) = B'(R).$$  \hspace{1cm} (25)

This defines an implicit relation between investments and retained earnings. From (25) and the concavity of $F(\cdot)$ and $B(\cdot)$, it is immediate that investments and retained earnings are positively correlated. Furthermore, (25), (24), and (3) imply that underinvestment must be accompanied by a retained earnings level below first-best, while overinvestment implies retained earnings higher than first-best.

Using these observations, one can rewrite the objective function of the manager (8) as

$$U = \alpha D - \beta \hat{h}(\Delta_T) + B(D) + C(x_1),$$
where $\Delta_T \equiv (I^{FB} - I) + (R^{FB} - R)$ is the sum of the underinvestment and deviation from the first-best retained earnings, and $\hat{h} (\cdot)$ is the real cost of this deviation. Moreover, (25) implies that $\hat{h} (\cdot)$ has the same loss-function properties as $h (\cdot)$. From here, the derivation of the separating equilibrium and partially pooling equilibria is similar to the base case of the model.

This simple extension has several implications. First, similar to the base-case model, in the separating equilibrium the sum of investments and retained earnings must be constant, and hence by (25), both are constant and uncorrelated with dividends. In the partially pooling equilibria, investments and retained earnings are fixed outside the pooling interval. Inside the pooling interval, both investments and retained earnings are positively correlated with earnings. Overall, in this model both investments and retained earnings exhibit a weak correlation with dividends.

7 Empirical Predictions

The main novel implication of our model is that dividend smoothing can arise endogenously in a world where earnings, return on investment, and managerial compensation are all continuous and smooth. In addition, our model has several empirical predictions on the relations between dividend smoothing, investments, firm profitability, ownership mix, and managerial incentives.\(^{29}\)

A first set of empirical implications comes from analyzing the determinants of the size of the pooling interval given implicitly in (20).

**Lemma 6** For any given $\delta \in [\Delta, \Delta^*]: \frac{\partial y_i}{\partial \alpha} > 0$, and $\frac{\partial y_i}{\partial \beta} < 0$.

The proof is immediate, since increasing $\alpha$ and decreasing $\beta$ raises the slope of the straight line in Figure 3, moving the intersection point to the right. Note, however, that an increase in the size of the interval does not automatically imply a higher probability of pooling, since the location of the interval changes as well. We have checked numerically (using Cobb–Douglass production functions) that the movement of the interval is of second order relative to the increase in size. Thus, an increase in $\frac{\partial y_i}{\partial \beta}$ is likely to be associated with an increase in the probability of dividend smoothing.

\(^{29}\)The specific choice of model (Miller and Rock) does not seem to be crucial for the smoothing argument. While we have not done the analysis, similar partially pooling equilibria are also likely to emerge in other static models of dividend signaling. The empirical predictions below rely on our setting only.
The ratio $\frac{\alpha}{\beta}$ is a measure of managerial myopia. It can be given two interpretations. First, $\frac{\alpha}{\beta}$ is a measure of the extent of short-term stock-based compensation. Thus, we predict that a higher extent of short-term stock-based compensation is associated with more dividend smoothing. Additionally, Stein (1989) argues that managers are prone to short-term incentives in firms that are likely to be acquired and firms that need to raise capital in the short term. Our model suggest that such firms are likely to exhibit a larger extent of dividend smoothing. A second interpretation of $\frac{\alpha}{\beta}$ is based on the premise that the manager represents the interests of both short-term and long-term investors. Using this interpretation, firms with more short-term investors (or firms in which short-term investors have more influence) tend to smooth dividends more. In particular, this suggests that dividend smoothing is less likely in private firms.

Next, note that a second way to move the intersection point in Figure 3 to the right is to flatten the convex curve. This curve determines the real cost of underinvestment, and is analogous to the investment opportunity curve. A flatter curve means that the marginal cost of underinvestment increases at a slower rate. Thus, better investment opportunities (a flatter curve) are associated with more dividend smoothing.

Finally, Proposition 3 suggests that both investments and the return on investments are higher on average when firms smooth dividends. This has both cross-sectional and time-series implications: firms that smooth dividends are expected to invest more and to show higher profitability in following years. Similarly, periods of dividend smoothing should be associated with higher investments, and followed by high profitability.

8 Conclusion

Dividend smoothing manifests itself as the practice of firms keeping their dividends constant over several years, until earnings change significantly. An implication is that firms follow a partially pooling dividend policy in which the same dividend is paid over an interval of earnings realizations. This behavior is not reflected in the perfectly separating equilibria studied in the classic dividend signaling models.

30 Note that in a setting that allows for an endogenous managerial objective function (see Section 6.1), both managerial incentives and dividend smoothing/investment decisions are determined endogenously. Thus, testing this prediction is not easy and requires finding a good instrument for compensation.

31 This is the original interpretation of Miller and Rock (1985).
We show that while such a unique separating equilibrium exists in our setting, the model has a multitude of partially pooling equilibria. These equilibria have full revelation of earnings for low and high outcomes, but for all “intermediate outcomes” the same dividend is announced. Thus, for all earnings that fall within a designated interval the manager chooses exactly the same dividend. Investors anticipate this behavior and price the firm correctly given this dividend policy. We show the existence of a continuum of such partially pooling equilibria.

Both the manager and the investors prefer any one of the partially pooling equilibria to the standard separating equilibrium. This suggests that they coordinate on playing one of these equilibria. We argue that the last-year dividend can serve as a simple and robust coordination mechanism allowing them to focus on just one of the partially pooling equilibria. Thus, they play a partially pooling equilibrium in which this year’s pooling dividend is equal to last year’s dividend. As a result, unless earnings fall outside of the pooling interval, the dividend is not changed. This gives rise to dividend smoothing over time.

References


32


Appendix A - Proofs

Proof of Lemma 1: Fix investors’ beliefs and let \( D_H > D_L \). From (8), and recalling that \( \Delta = I_F - (x_1 - D) \), we have

\[
U(x_1, D_H) - U(x_1, D_L) = \alpha (D_H - D_L) - \beta (h(I_F - x_1 + D) - h(I_F - x_1 + D_L)) + B(D_H) - B(D_L).
\]

Differentiating by \( x_1 \) gives

\[
\frac{\partial}{\partial x_1} (U(x_1, D_H) - U(x_1, D_L)) = \beta (h'(I_F - x_1 + D) - h'(I_F - x_1 + D_L)),
\]

which is positive by the convexity of \( h(\cdot) \), as required.

Proof of Proposition 1: In the text we have shown existence and derived the properties of the linear separating equilibrium. Below we show uniqueness. We first establish two properties of any separating equilibrium.
Claim 1: In any separating equilibrium, dividend policy is differentiable everywhere.

Proof of Claim 1: Consider any separating equilibrium. Suppose that given this equilibrium a type $x_1$ optimally pays a dividend of $D^*$. Let $D > D^*$. By the optimality of $D^*$ and using (8):

$$\alpha D^* - \beta h(I_{FB} - (x_1 - D^*)) + B(D^*) \geq \alpha (D) - \beta h(I_{FB} - (x_1 - D)) + B(D).$$

Rearranging yields

$$\alpha (D - D^*) - \beta (h(I_{FB} - (x_1 - D)) - h(I_{FB} - (x_1 - D^*))) \leq B(D^*) - B(D).$$

By the differentiability of $h$ we have

$$h(I_{FB} - (x_1 - D)) - h(I_{FB} - (x_1 - D^*)) = h'(\xi_D)(D - D^*),$$

where $I_{FB} - (x_1 - D^*) < \xi_D < I_{FB} - (x_1 - D)$, and $\lim_{D\to D^*} \xi_D = I_{FB} - (x_1 - D)$.

It follows that

$$\alpha (D - D^*) - \beta h'(\xi_D)(D - D^*) \leq B(D^*) - B(D),$$

or equivalently

$$\beta h'(\xi_D) - \alpha \geq \frac{B(D) - B(D^*)}{D - D^*}. \tag{26}$$

Consider now $D < D^*$. As before, we can write

$$\alpha (D - D^*) - \beta h'(\xi_D)(D - D^*) \leq B(D^*) - B(D),$$

where $I_{FB} - (x_1 - D) < \xi_D < I_{FB} - (x_1 - D^*)$, and $\lim_{D\to D^*} \xi_D = I_{FB} - (x_1 - D)$.

Equivalently,

$$\beta h'(\xi_D) - \alpha \leq \frac{B(D) - B(D^*)}{D - D^*}. \tag{27}$$

Combining (26) and (27) yields

$$\beta h'(\xi_D) - \alpha \leq \frac{B(D) - B(D^*)}{D - D^*} \leq \beta h'(\xi_D) - \alpha.$$

By the continuity of $h'(\cdot)$, both the RHS and LHS converge to the same limit, implying that $B(D)$ is differentiable at $D^*$. As this argument applies to any $D^*$, it follows that $B(D)$ is differentiable everywhere. But, in a separating equilibrium with dividend policy $\Lambda(\cdot)$, $B(D) = \frac{\alpha}{1 + \lambda} F(\Lambda^{-1}(D))$. Thus, the differentiability of $B(\cdot)$ implies that dividend policy is also differentiable.
Claim 2: In any separating equilibrium, investment is bounded between 0 and $I^{FB}$.

That is, there can never be overinvestment in a separating equilibrium.

Proof of Claim 2: Consider a separating equilibrium with dividend policy $\Lambda(\cdot)$.

Denote $\varphi = \Lambda^{-1}$. Assume on the contrary that for some $x_1$, $\Lambda(x_1) = D^*$, and $I^{FB} < x_1 - D^*$. By (6), the equilibrium payoff to a manager paying a dividend $D^*$ is (using $\varphi(D^*) = x_1$)

$$U(D^*) = \alpha \left( D^* + \frac{F(\varphi(D^*) - D^*)}{1+i} \right) + \beta \left( D + \frac{F(x_1 - D^*)}{1+i} \right).$$

By Claim 1 we can differentiate to obtain

$$U'(D^*) = \alpha \left( \frac{F'(\varphi(D^*) - D^*) (\varphi'(D^*) - 1)}{1+i} \right) + \beta \left( \frac{F'(x_1 - D^*)}{1+i} \right).$$

By the concavity of $F(\cdot)$ and using $x_1 - D^* > I^{FB}$ and (3) we obtain that Term $B$ is strictly positive. Also, by Lemma 2, $\varphi'(D^*) - 1 \geq -1$, and hence again by (3) and the concavity of $F(\cdot)$, Term $A$ is positive. It follows that $U'(D^*)$ is strictly positive, contradicting that $D^*$ is optimal.

We now turn to the uniqueness proof. By Claim 1 and using (8), we can write the first order condition of the manager:

$$\alpha - \beta h'(I^{FB} - x_1 + D) + B'(D) = 0.$$

In a separating equilibrium, $B(D) = \frac{\alpha}{1+i} F(x_1 - D)$. Hence:

$$\alpha - \beta h'(I^{FB} - F^{-1}(\frac{1+i}{\alpha} B(D))) + B'(D) = 0.$$

This differential equation describes $B(D)$ in equilibrium. By the differentiability of $h'(\cdot)$ and $F(\cdot)$ we have that $B(\cdot)$ is twice differentiable. Implicitly differentiating by $D$ yields:

$$B''(D) = -\beta h'' \left( I^{FB} - F^{-1} \left( \frac{1+i}{\alpha} B(D) \right) \right) \frac{\frac{1+i}{\alpha} B'(D)}{F'(F^{-1}(\frac{1+i}{\alpha} B(D)))}. \quad (28)$$

As $h'' > 0$ and $F' > 0$, $B(D)$ is either (i) decreasing and convex everywhere; or (ii) increasing and concave everywhere. From Claim 2, investment is bounded between 0 and $I^{FB}$; hence, the range of $\Lambda(\cdot)$ is the entire real line. It follows that $B(D)$ has the entire real line as its domain but is bounded between 0 and $\frac{\alpha}{1+i} F(I^{FB})$. However, a
Proof of Lemma 3: We will show that the straight line \( \varphi_1(y) = \frac{a_2}{y} \) and the curve \( \varphi_2(y) = h(y + \delta) - h(\delta) \) must cross at \( y > 0 \) for some \( \delta \) values in a left neighborhood of \( \Delta^* \).

The slope of \( \varphi_2(\cdot) \) is \( \alpha/\beta \) while the slope of \( \varphi_2(\cdot) \) is \( \varphi'_2(y) = h'(y + \delta) \). Consider the case \( \delta = \Delta^* \). In this case \( \varphi'_2(y) = h'(y + \Delta^*) \), and in particular, \( \varphi'_2(0) = h'(\Delta^*) = \frac{a_2}{y} \) (by 11). Since \( \varphi_1(0) = \varphi_2(0) \) this means that for \( \delta = \Delta^* \), \( \varphi_2(\cdot) \) is tangential to \( \varphi_1(\cdot) \) at 0. Moreover, since \( h(\cdot) \) is convex, \( y = 0 \) is the only intersection point of the two curves. Hence, in this case \( \varphi_2(\cdot) \) lies everywhere above \( \varphi_1(\cdot) \).

Consider now the case \( \delta = \Delta^* - \varepsilon \), where \( \varepsilon > 0 \) can be arbitrarily small. Then,

\[
\varphi'_2(0) = h'(\Delta^* - \varepsilon) < h'(\Delta^*),
\]

due to the convexity of \( h(\cdot) \). Since \( \varphi_1(\cdot) \) and \( \varphi_2(\cdot) \) still coincide at \( y = 0 \), there is a right neighborhood of \( y = 0 \) for which \( \varphi_2(y) < \varphi_1(y) \). Also, because for \( \delta = \Delta^* \), \( \varphi_2(\cdot) \) lies everywhere above \( \varphi_1(\cdot) \) and strictly so for \( y > 0 \), and since \( \varepsilon > 0 \) can be arbitrarily small, there must by a \( y > 0 \) for which \( \varphi_2(y) > \varphi_1(y) \). Consequently, by continuity of both \( \varphi_1(\cdot) \) and \( \varphi_2(\cdot) \), the two curves must meet for some \( y > 0 \). The crossing point is unique since \( h(\cdot) \) is strictly convex.

By the implicit function theorem,

\[
\frac{\partial y_8}{\partial \delta} = \frac{\beta (h'(y_8 + \delta) - h'(\delta))}{\alpha - \beta h'(y_8 + \delta)} = -\frac{h'(y_8 + \delta) - h'(\delta)}{h'(y_8 + \delta) - h'(\Delta^*)} < -\frac{h'(y_8 + \delta) - h'(\Delta^*)}{h'(y_8 + \delta) - h'(\Delta^*)} = -1,
\]

where the first equality follows from implicitly differentiating (20), the second equality from (11), and the inequality from the convexity of \( h(\cdot) \).

Since \( y_8 = b_8 - a_8 \), it remains to show that for \( \delta < \Delta^* \), \( -y_8 + \Delta^* < \delta \). Define \( R(\delta) \equiv \delta + y_8 - \Delta^* \). Since \( y_8 \) vanishes for \( \delta = \Delta^* \) we have \( R(\Delta^*) = 0 \). Moreover,

\[
\frac{\partial R(\delta)}{\partial \delta} = 1 + \frac{\partial y_8}{\partial \delta} < 0,
\]

showing that \( R(\delta) \) is strictly positive for \( \delta < \Delta^* \), as required.
Proof of Proposition 2: Let $\delta \in [\Delta, \Delta^*)$. From Lemma 4 there exists a non-empty interval $[a_\delta, b_\delta]$ such that the indifference conditions (17) and (18) are satisfied. We first show that following any earnings $x_1 \in (a_\delta, b_\delta)$, the manager strictly prefers to announce the pooling dividend, $D_\delta = b_\delta - I^{FB} + \delta$, to any other out-of-equilibrium dividend, $D \in (a_\delta - I^{FB} + \Delta^*, b_\delta - I^{FB} + \delta) \cup (b_\delta - I^{FB} + \delta, b_\delta - I^{FB} + \Delta^*)$.

Consider $D' \in (a_\delta - I^{FB} + \Delta^*, b_\delta - I^{FB} + \delta)$. Following earnings $a_\delta$, the manager strictly prefers the pooling dividend to $D'$. By the single-crossing property, following earnings $x'_1 > a_\delta$ the manager strictly prefers the pooling dividend to $D'$, as required.

Consider now $D'' \in (b - I^{FB} + \delta, b - I^{FB} + \Delta^*)$. We assume (by contradiction) that there exists realized earnings $x''_1 \in (a_\delta, b_\delta)$ for which the manager strictly prefers the dividend $D''$ to $D_\delta$. According to the single-crossing property, it must be that also following earnings $b_\delta$ (which are higher than $x''_1$) the manager strictly prefers the dividend $D''$ to $D_\delta$. This contradicts the fact that following earnings $b_\delta$ the manager prefers $D_\delta$ to any other dividend.

The claim that given earnings $x_1 \notin (a_\delta, b_\delta)$, type $x_1$ prefers the separating dividend to the pooling dividend follows from a similar application of the SCP.

Proof of Lemma 5: For $x_1 < a_\delta$ and $x_1 > b_\delta$ the partially-pooling and separating dividend policies are identical. For $x_1 = a_\delta$ and $x_1 = b_\delta$ the manager is indifferent between the two alternatives. Thus, consider $x_1 \in (a_\delta, b_\delta)$. The partially-pooling dividend policy is to pay a dividend of $D_\delta = \Lambda_\delta(x_1) = b_\delta - (I^{FB} - \delta)$. The separating dividend policy is to pay a dividend equal to $D_s = \Lambda_s(x_1) = x_1 - (I^{FB} - \Delta^*)$.

Consider the following two cases:

Case 1: $D_\delta \geq D_s$. We know that type $a_\delta$ prefers paying $D_\delta$ to $\Lambda_s(x_1)$, by the SCP, and since $D_\delta > D_s$, it follows that type $x_1$ (which is higher than $a_\delta$) strictly prefers the pooling dividend over the separating dividend policy.

Case 2: $D_\delta < D_s$. Again, we know that type $b_\delta$ prefers paying $D_\delta$ to $\Lambda_s(x_1)$. By the SCP, and since $D_\delta > D_s$ it follows that type $x_1$ (which is lower than $b_\delta$) strictly prefers the pooling dividend to the separating dividend policy.

Proof of Proposition 3. Set $\delta \in [\Delta, \Delta^*)$, and consider the corresponding $\delta$-
equilibrium.

**Claim 1.** From Lemma 5 we know that the manager prefers the \( \delta \)-equilibrium to the separating equilibrium for every \( x_1 \), and strictly so on a positive length interval. It follows that the manager also strictly prefers the \( \delta \)-equilibrium to the separating equilibrium in expectation (ex-ante). That is,

\[
E_{\delta} U = E_{\delta} (\alpha V^M) + E_{\delta} (\beta V^I) > E_s U = E_s (\alpha V^M) + E_s (\beta V^I),
\]

(29)

where \( E_{\delta} \) and \( E_s \) stand for expectation under the partially pooling and separating equilibria, respectively. The ex-ante expected market value equals the intrinsic value of the firm. That is, \( E_{\delta} (V^M) = E_{\delta} (V^I) \) and \( E_s (V^M) = E_s (V^I) \). Hence, (29) can be written as

\[
(\alpha + \beta) E_{\delta} (V^I) > (\alpha + \beta) E_s (V^I),
\]

which implies that \( E_{\delta} (V^I) > E_s (V^I) \). Namely, the expected intrinsic value of the firm is higher in the partially pooling equilibrium.

**Claims 2 and 3.** Rearranging terms in (17) gives

\[
\alpha (\Delta^* - \delta) - \beta (h(\Delta^*) - h(\delta)) = B_{\delta} - B_s.
\]

By Lemma 10 (Appendix B), the LHS is strictly positive. It follows that \( B_{\delta} - B_s > 0 \). From the definitions of \( B_s \) and \( B_{\delta} \) (Eq. (10) and (16)) we have

\[
0 < B_{\delta} - B_s = \frac{\alpha}{1 + \frac{1}{I}} E (F(I|D_\delta)) - F(I^*).
\]

This implies that \( E (F(I|D_\delta)) > F(I^*) \). Now, the monotonicity and concavity of \( F(\cdot) \) imply that also \( E (I|D_\delta) > I^* \).

\[33\] Thus, expected investment and expected return on investment are higher in the \( \delta \)-equilibrium, while expected underinvestment and expected dividends are lower.

**Claim 4.** Since the expected value of the firm is higher under the partially pooling equilibrium, and the investors price the stock correctly on average, they ex-ante prefer the \( \delta \)-equilibrium.

**Claim 5.** Since both the manager and the investors ex-ante prefer the \( \delta \)-equilibrium, it Pareto-dominates the separating equilibrium. ■

\[33\] To see this, assume on the contrary that \( E (I_1|D_\delta) \leq I^* \). Then, by the monotonicity of \( F(\cdot) \), we have \( E (F(I|D_\delta)) \leq F(I^*) \). By Jensen’s inequality using the concavity of \( F(\cdot) \), \( E (F(I|D_\delta)) \leq F(I^*) \) — a contradiction.
Appendix B - Proof of Lemma 4

We first prove preliminary results in Lemmas 7, 8, 9, and 10 below.

**Lemma 7** For all \( y > 0 \) the following hold:

\[
\lim_{a \to \infty} E \left( x_1 - a \mid x_1 \in [a, a + y] \right) = 0
\]

\[
\lim_{a \to \infty} E \left( x_1 - (a + y) \mid x_1 \in [a, a + y] \right) = 0.
\]

**Proof of Lemma 7:** The proof follows from standard properties of the normal distribution. See Lemma 16 (in Appendix D) for a more general proof that establishes this result for a wide class of distributions that includes the normal. ■

**Lemma 8** Let \( \delta \in [\Delta, \Delta^*] \), and let \( y_\delta > 0 \) be the uniquely determined size of the pooling interval. Also, let \( D_\delta = a + y_\delta - (I^{FB} - \delta) \) be the pooling dividend corresponding to \( a \) and \( \delta \). Define

\[
L_1(a) \equiv E \left( F \left( x_1 - D_\delta \right) - F(a - D_\delta) \mid x_1 \in [a, a + y_\delta] \right)
\]

\[
L_2(a) \equiv E \left( F \left( x_1 - D_\delta \right) - F(a + y_\delta - D_\delta) \mid x_1 \in [a, a + y_\delta] \right).
\]

Then, \( \lim_{a \to \infty} L_1(a) = 0 \) and \( \lim_{a \to \infty} L_2(a) = 0 \).

**Proof of Lemma 8:** Using the mean value theorem, we can write

\[
L_1(a) = E \left( F'(\xi_a) \left( x_1 - a \right) \mid x_1 \in [a, a + y_\delta] \right),
\]

where

\[
I^{FB} - \delta - y_\delta < \xi_a < x_1 - a - y_\delta + I^{FB} - \delta.
\]

Since \( x_1 - a < y_\delta \) we have

\[
I^{FB} - y_\delta < \xi_a < I^{FB}.
\]

That is, regardless of \( a \), \( \xi_a \) is bounded in the interval \([I^{FB} - y_\delta, I^{FB}]\). Since \( F'(\cdot) \) is continuous it follows that \( F'(\xi_a) \) is uniformly bounded. So, there are \( \overline{m}, \underline{m} > 0 \) such that \( \underline{m} \leq F'(\xi_a) \leq \overline{m} \) for all \( a \in \mathbb{R} \). It follows that

\[
\overline{m} E(x_1 - a) \mid x_1 \in [a, a + y_\delta] \leq E \left( F'(\xi_a) \left( x_1 - a \right) \mid x_1 \in [a, a + y_\delta] \right) \leq \overline{m} E(x_1 - a) \mid x_1 \in [a, a + y_\delta].
\]

But, from Lemma 7, both the LHS and the RHS of this inequality tend to 0 as \( a \to \infty \). This shows that \( \lim_{a \to \infty} L_1(a) = 0 \). The second limit is proved similarly. ■

41
Lemma 9 Set $\delta \in [\Delta, \Delta^*),$ and let $y_{\delta} > 0$ be the uniquely determined size of the pooling interval. Also, let $D_{\delta} = a + y_{\delta} - (I^{FB} - \delta)$ be the pooling dividend. For any $a \in \mathbb{R}$ define

$$J(a) \equiv E \left( F(x_1 - D_{\delta}) - F(I^*) | x_1 \in [a, a + y_{\delta}] \right).$$

Then

$$\lim_{a \to \infty} J(a) = F \left( I^{FB} - \delta - y_{\delta} \right) - F \left( I^{FB} - \Delta^* \right) < 0$$

$$\lim_{a \to -\infty} J(a) = F \left( I^{FB} - \delta \right) - F \left( I^{FB} - \Delta^* \right) > 0.$$

Proof of Lemma 9:

\[
\begin{align*}
\lim_{a \to \infty} J(a) & = \lim_{a \to \infty} E \left( F(x_1 - D_{\delta}) - F(I^*) | x_1 \in [a, a + y_{\delta}] \right) \\
& = \lim_{a \to \infty} E \left( F(x_1 - D_{\delta}) - F(a - D_{\delta}) | x_1 \in [a, a + y_{\delta}] \right) \\
& \quad + \lim_{a \to \infty} E \left( F(a - D_{\delta}) - F(I^*) | x_1 \in [a, a + y_{\delta}] \right) \\
& = \lim_{a \to \infty} E \left( F(a - D_{\delta}) - F(I^*) | x_1 \in [a, a + y_{\delta}] \right) \quad \text{(by Lemma 8)} \\
& = F \left( I^{FB} - \delta - y_{\delta} \right) - F \left( I^{FB} - \Delta^* \right) \quad \text{(using $D_{\delta} = a + y_{\delta} - (I^{FB} - \delta)$).}
\end{align*}
\]

From Lemma 3, $\Delta^* < \delta + y_{\delta}.$ This implies that $F \left( I^{FB} - \delta - y_{\delta} \right) - F \left( I^{FB} - \Delta^* \right) < 0,$ as required. The other limit is proved similarly. ■

Lemma 10 For all $\delta < \Delta^*, \alpha (\Delta^* - \delta) - \beta (h(\Delta^*) - h(\delta)) > 0.$

Proof of Lemma 10: Let $\delta < \Delta^*.$ We have

$$\alpha (\Delta^* - \delta) - \beta (h(\Delta^*) - h(\delta)) = \alpha \Delta^* - \beta h(\Delta^*) - (\alpha \delta - \beta h(\delta)).$$

Recall that $\Delta^*$ is the underinvestment level in the separating equilibrium. That is, $\Delta^* = \arg \max_{\Delta} (\alpha \Delta - \beta h(\Delta)).$ This implies that $\alpha \Delta^* - \beta h(\Delta^*) > \alpha \delta - \beta h(\delta),$ as required. ■

We turn now to the proof of Lemma 4 itself.

Set $\delta \in [\Delta, \Delta^*)$ and let $y_{\delta} > 0$ be the unique positive solution to Eq. (20), which exists by Lemma 3. We will show that there exists an interval $[a_\delta, b_\delta]$ such that $b_\delta = a_\delta + y_{\delta}$ and the indifference conditions (17) and (18) are satisfied.

Consider the indifference condition (18). Let $H(a)$ be the difference between the RHS and the LHS of (18):

$$H(a) = B_\delta - B_* - \alpha (\Delta^* - \delta) + \beta (h(\Delta^*) - h(\delta)).$$

42
Showing that (18) is satisfied is equivalent to showing that there is an \( a \) such that \( H(a) = 0 \). Using the definitions of \( B_\delta \) and \( B_s \) and Lemmas 9 and 10, we have

\[
\lim_{a \to -\infty} H(a) = -\alpha (\Delta^* - \delta) + \beta (h(\Delta^*) - h(\delta)) + \frac{\alpha}{1 + i} \lim_{a \to -\infty} E(F(x_1 - D_\delta) - F(I^*)|x_1 \in [a, a + y_\delta])
\]

\[
= -\alpha (\Delta^* - \delta) + \beta (h(\Delta^*) - h(\delta)) + \frac{\alpha}{1 + i} \lim_{a \to -\infty} J(a) < 0.
\]

Similarly,

\[
\lim_{a \to -\infty} H(a) = -\alpha (\Delta^* - \delta) + \beta (h(\Delta^*) - h(\delta)) + \frac{\alpha}{1 + i} \lim_{a \to -\infty} J(a)
\]

\[
= -\alpha (\Delta^* - \delta) + \beta (h(\Delta^*) - h(\delta)) + \frac{\alpha}{1 + i} (F(I^{FB} - \delta) - F(I^{FB} - \Delta^*))
\]

\[
= \beta h(\Delta^*) + \alpha \left( \frac{1}{1 + i} (F(I^{FB}) - F(I^{FB} - \Delta^*)) - \Delta^* \right)
\]

\[
- \beta h(\delta) - \alpha \left( \frac{1}{1 + i} (F(I^{FB}) - F(I^{FB} - \delta)) - \delta \right).
\]

From the definition of \( h(\cdot) \) (Eq. (7)) it follows that Term 1 is equal to \( h(\Delta^*) \), and Term 2 is equal to \( h(\delta) \). Hence,

\[
\lim_{a \to -\infty} H(a) = (\alpha + \beta) (h(\Delta^*) - h(\delta)) > 0,
\]

where the inequality follows since \( 0 \leq \delta < \Delta^* \) and \( h(\cdot) \) is strictly increasing for positive underinvestment levels.

We have shown that \( H(a) \) changes signs when \( a \) moves from the left tail to the right tail of the distribution. From continuity, there exists an \( a \) such that \( H(a) = 0 \), which implies that the indifference condition (18) is satisfied. Choose one such \( a \) and denote it by \( a_\delta \). Set \( b_\delta = a_\delta + y_\delta \). Since both (18) and (19) are satisfied, and since (17) is the difference between (18) and (19), it follows that also (17) is satisfied.

\section*{Appendix C - Out-of-Equilibrium Beliefs}

The game-theoretic literature on equilibrium refinements offers a multitude of concepts to limit the freedom of the modeler in choosing “reasonable” out-of-equilibrium
beliefs in signaling games. One prevalent criterion for refinement is the “Intuitive Criterion” of Cho and Kreps (1987). It is straightforward that all the partially pooling equilibria survive this criterion. In fact, the Intuitive Criterion works best in models with just two types, and does not impose any restriction on out-of-equilibrium beliefs in our continuous type framework. Other, stronger refinements are D1 (Cho and Kreps, 1987), Universal Divinity (Banks and Sobel, 1987), and Never-a-Weak-Best-Response (Kohlberg and Mertens, 1986). Cho and Sobel (1990) show that these three are equivalent when single-crossing and other standard properties are satisfied (as is the case here). Furthermore, they show that these three refinements rule out any kind of pooling in equilibrium. Hence, these criteria cannot be used to refine our partially-pooling equilibria, unless one would like to rule them out up-front. By contrast, the equilibrium selection criterion of Maskin and Tirole (1992) eliminates the separating equilibrium, since it is Pareto dominated by any of the partially pooling equilibria.

Our approach in checking the robustness of the out-of-equilibrium beliefs is similar to Harrington (1987). His idea is that out-of-equilibrium beliefs are “reasonable” if the expected type given the signal is increasing in the signal. In our setting, this means that higher dividends should be associated with higher types (on average) both on and off the equilibrium path. Formally,

**Definition 2** We say that investors’ beliefs are monotone if for all \( D \) (both on and off the equilibrium path), \( E(x_1|D) \) is non-decreasing in \( D \).

While deriving the partially pooling \( \delta \)-equilibria we assumed for simplicity that if investors observe an out-of-equilibrium dividend then they assign a probability 1 to the event that the manager is playing the benchmark linear equilibrium. This set of out-of-equilibrium beliefs, while simple and sufficient to support the \( \delta \)-equilibrium, is not monotone, since out-of-equilibrium dividends that just slightly exceed \( D_\delta \) entail an expected type lower than the conditional type given \( D_\delta \). In the rest of this appendix we show that each \( \delta \)-equilibrium can be supported by monotone out-of-equilibrium beliefs, maintaining all of the results in the paper.

**Definition 3** Let \( \delta \in [\Delta, \Delta^*] \). Consider an out-of-equilibrium dividend \( \hat{D} \). We say that beliefs given \( \hat{D} \) are concentrated, if there is a type \( \hat{x} \) such that the beliefs given \( \hat{D} \) assign probability 1 to \( \hat{x} \).
We will show that there is a concentrated set of out-of-equilibrium beliefs that is also monotone.

Recall that all out of equilibrium dividends satisfy \( \hat{D} \in (a_\delta - I^*, b_\delta - I^{FB} + \delta) \cup (b_\delta - I^{FB} + \delta, b_\delta - I^*) \). The single-crossing property suggests that if \( a_\delta \) is indifferent between its pooling dividend \( D_\delta = b_\delta - I^{FB} + \delta \) and an out-of-equilibrium dividend \( \hat{D} \in (a_\delta - I^*, b_\delta - I^{FB} + \delta) \), then all types in the pooling interval \((a_\delta, b_\delta)\) will strictly prefer the pooling dividend. Similarly, if \( b_\delta \) is indifferent between the pooling dividend and an out-of-equilibrium dividend \( \hat{D} \in (b_\delta - I^{FB} + \delta, b_\delta - I^*) \) then all types in the pooling interval strictly prefer the pooling dividend. Thus, a natural candidate for a monotone and concentrated set of out-of-equilibrium beliefs is those beliefs that render types \( a_\delta \) and \( b_\delta \) indifferent between the pooling divided and the appropriate out-of-equilibrium dividend.

Let \( c_\delta \) be implicitly defined by

\[
F(c_\delta - D_\delta) = E(F(x - D_\delta) | D_\delta).
\]

That is, if investors assign a probability 1 to the event that the manager’s type is \( c_\delta \), then given a dividend of \( D_\delta \) the market price would be the same as the market price in the pooling dividend case. Clearly, \( c_\delta \in (a_\delta, b_\delta) \). The uniqueness of \( c_\delta \) is a simple application of the implicit function theorem. Consider an out-of-equilibrium dividend \( \hat{D} \) and a concentrated belief \( \hat{x} \in (a_\delta, b_\delta) \). The market price given these two is

\[
V^M(\hat{x}, \hat{D}) = \hat{D} + \frac{1}{1 + \hat{r}}F(\hat{x} - \hat{D}).
\]

Consider now an out-of-equilibrium dividend \( \hat{D} \in (a_\delta - I^*, b_\delta - I^{FB} + \delta) \). Compare the utility of a type \( a_\delta \) manager who issues a dividend \( \hat{D} \) to the utility of such a manager if he issues the separating dividend. The former naturally depends on the out-of-equilibrium beliefs. If beliefs are such that the market price is identical to the separating price, then the manager obviously prefers the separating dividend to \( \hat{D} \) because the market price is identical, while the intrinsic value is lower given \( \hat{D} \). If, on the other hand, beliefs are such that the market price is equal to \( V^M(D_\delta) \), then the type \( a_\delta \) manager strictly prefers \( \hat{D} \) to the separating dividend. This follows since the market price is identical now to the pooling dividend price, while the intrinsic value is higher, and we know that a type \( a_\delta \) is indifferent between separation and pooling. By continuity, there is a market price \( \hat{P} \) between the separating price of \( a_\delta \) and \( V^M(D_\delta) \).
that renders type $a_\delta$ exactly indifferent between announcing its separating dividend and announcing $\check{D}$. In particular,

$$\Lambda_s (a_\delta) + \frac{1}{1+i} F (a_\delta - \Lambda_s (a_\delta)) < \check{V}_1 < D_\delta + \frac{1}{1+i} E (F (x - D_\delta|D_\delta)). \quad (31)$$

We claim now that there is also a unique set of concentrated beliefs that corresponds to this market price. Namely, there is a unique $\check{x} \in (a_\delta, c_\delta)$ such that

$$\check{V}_1 = V_1 (\check{x}, \check{D}) = \check{D} + \frac{1}{1+i} F (\check{x} - \check{D}).$$

First note that for $\check{x} = a_\delta$,

$$\check{D} + \frac{1}{1+i} F (\check{x} - \check{D}) = \check{D} + \frac{1}{1+i} F (a_\delta - \check{D}) < \Lambda_s (a_\delta) + \frac{1}{1+i} F (a_\delta - \Lambda_s (a_\delta)), \quad (32)$$

where the inequality follows since $F' > 1+i$ on the relevant range.

Also, if $\check{x} = c_\delta$ then

$$\check{D} + \frac{1}{1+i} F (\check{x} - \check{D}) = \check{D} + \frac{1}{1+i} F (c_\delta - \check{D}) > D_\delta + \frac{1}{1+i} F (c_\delta - D_\delta) \quad (33)$$

where the inequality follows again by $F' > 1+i$, and the last equality from (30).

Combining (31), (32), and (33) we obtain that by varying $\check{x}$, $V_1 (\check{x}, \check{D})$ can take all possible values between the separating price of $a_\delta$ and the price corresponding to the pooling dividend $D_\delta$. Furthermore, since $F (\cdot)$ is strictly increasing, there is a unique $\check{x} = \check{x} (\check{D})$ that corresponds to the market price $\check{V}_1$.

A parallel argument applies to out-of-equilibrium dividends $\check{D} \in (b_\delta - I^{FB} + \delta, b_\delta - I^*)$.

We have proved:

**Lemma 11** The following holds:

1. For each out-of-equilibrium dividend $\check{D} \in (a_\delta - I^*, b_\delta - I^{FB} + \delta)$ there is a unique $\check{x} = \check{x} (\check{D})$ such that if given $\check{D}$, beliefs are that the manager’s type is $\check{x}$ with probability 1, then type $a_\delta$ is indifferent between announcing its separating dividend and $\check{D}$.

2. For each out-of-equilibrium dividend $\check{D} \in (b_\delta - I^{FB} + \delta, b_\delta - I^*)$ there is a unique $\check{x} = \check{x} (\check{D})$ such that if given $\check{D}$, beliefs are that the manager’s type is $\check{x}$ with probability 1, then type $b_\delta$ is indifferent between announcing its separating dividend and $\check{D}$. 

46
From the derivation above, it is easy to see that \( \hat{x}(D) \) is continuous at \( D_\delta = b_\delta - I^{FB} + \delta \) and strictly increasing everywhere. We have proved the following result:

**Proposition 5** Let \( \delta \in [\Delta, \Delta^*). \) There is a unique set of concentrated out-of-equilibrium beliefs supporting the \( \delta \)-equilibrium, such that:

1. For each out-of-equilibrium dividend \( \hat{D} \in (a_\delta - I, b_\delta - I^{FB} + \delta) \), type \( a_\delta \) is indifferent between announcing \( \hat{D} \) and announcing its separating dividend \( \Lambda_\delta(a_\delta) \).

2. For each out-of-equilibrium dividend \( \hat{D} \in (b_\delta - I^{FB} + \delta, a_\delta - I^*) \), type \( b_\delta \) is indifferent between announcing \( \hat{D} \) and announcing its separating dividend \( \Lambda_\delta(b_\delta) \).

Furthermore, this set of beliefs is continuous and monotone increasing.

We can now replace the out-of-equilibrium beliefs used in Section 4 with the monotone out-of-equilibrium beliefs derived in Proposition 5. The existence result is then intact as are all of the other results in the paper.

**Appendix D - Derivations for the Dynamic Setup**

**Preliminary Results**

We start by developing results that enable us to calculate the conditional expectations in (21). In the one period model, the prior distribution was normal, and the posterior conditional on pooling was truncated normal. This is no longer the case in the dynamic model. Instead, if a partially pooling equilibrium was played in period \( t-1 \) and a pooling dividend was announced, then earnings in period \( t \) (from the point of view of the uninformed investors) follow a sum of a random variable with a continuous distribution over a bounded support and a normal noise term. The following results study the tail properties of such distributions, which are essential for proving existence of equilibrium in the dynamic setup.

Let \( y > 0 \). Let \( \tilde{w} \) be a random variable with support in the form of a compact interval \([a, a + y]\) and a continuous density \( \kappa(w) \) for \( w \in [a, a + y] \). Let \( \tilde{\varepsilon} \) be normally distributed random noise. For brevity, assume that \( \tilde{\varepsilon} \) is standard normal, with density \( \phi(\cdot) \). We are interested in the distribution of \( \tilde{z} \equiv \tilde{w} + \tilde{\varepsilon}. \) Clearly \( \tilde{z} \) has continuous density and its support is the entire real line; however, \( \tilde{z} \) is not normally distributed.
The density of \( \tilde{z} \), \( h(z) \), is given by the convolution

\[
h(z) = \int_{a}^{a+y} \kappa(w) \phi(z - w) \, dw.
\]  

(34)

We first argue that while \( h(\cdot) \) is not normal, it approaches the normal density in both the left and the right tails. Formally,

**Lemma 12** The following hold:

1. \( \lim_{z \to \infty} |h(z) - \phi(z)| = 0 \).
2. \( \lim_{z \to -\infty} |h(z) - \phi(z)| = 0 \).

**Proof of Lemma 12:** Note that for all \( z \), \( \phi(z) = \int_{a}^{a+y} \kappa(w) \phi(z - w) \, dw. \) Hence,

\[
h(z) - \phi(z) = \int_{a}^{a+y} \kappa(w) (\phi(z - w) - \phi(z)) \, dw.
\]

Denote \( \tau_{zw} \equiv \min \{ z - w, z \} \) and \( \bar{\tau}_{zw} \equiv \max \{ z - w, z \} \). By the mean value theorem, \( \phi(z - w) - \phi(z) = -w\phi'(\xi_w) \) for some \( \xi_w \in (\tau_{zw}, \bar{\tau}_{zw}) \). By the triangle inequality for integrals,

\[
|h(z) - \phi(z)| = \left| \int_{a}^{a+y} \kappa(w) w\phi'(\xi_w) \, dw \right| \leq \int_{a}^{a+y} \kappa(w) |w| \left| \phi'(\xi_w) \right| \, dw.
\]

Now, recall that the normal density satisfies \( \phi'(z) < 0 \) and \( \phi''(z) > 0 \) for all \( z \) sufficiently large, and \( \phi'(z) > 0 \) and \( \phi''(z) < 0 \) for all \( z \) sufficiently small. For all \( z \) sufficiently large \( \phi''(z) > 0 \), hence \( |\phi'(\xi_w)| < |\phi'(\bar{\tau}_{zw})| \). Consequently, for all \( z \) sufficiently large

\[
|h(z) - \phi(z)| < \int_{a}^{a+y} \kappa(w) |w| \left| \phi'(\bar{\tau}_{zw}) \right| \, dw.
\]

Note that \( \lim_{z \to \infty} \phi'(\bar{\tau}_{zw}) = 0 \). Hence, by Lebesgue’s dominated convergence theorem, \( \lim_{z \to \infty} |h(z) - \phi(z)| = 0 \). In a similar manner, \( \lim_{z \to -\infty} |h(z) - \phi(z)| = 0 \).

The following lemma shows a standard property of the normal density.

**Lemma 13** For all \( t > 0 \), \( \lim_{c \to \infty} \frac{\phi(c+t)}{\phi(c)} = 0 \) and \( \lim_{c \to -\infty} \frac{\phi(c+t)}{\phi(c)} = \infty \).
The proof is immediate from the functional form of the normal density. Next we show that the same property holds for \( h(\cdot) \). First, for each two real numbers \( c, w \) denote

\[
\rho_c(w) = \frac{\kappa(w) \phi(c - w)}{\int_a^{a+y} \kappa(u) \phi(c - u) \, du}.
\]

Note that \( \rho_c(w) \) is a probability measure over \([a, a + y]\).

**Lemma 14** For all \( t > 0 \), \( \lim_{c \to \infty} \frac{h(c + t)}{h(c)} = 0 \) and \( \lim_{c \to -\infty} \frac{h(c + t)}{h(c)} = \infty \).

**Proof of Lemma 14:** Using (34),

\[
\frac{h(c + t)}{h(c)} = \frac{\int_a^{a+y} \kappa(w) \phi(c + t - w) \, dw}{\int_a^{a+y} \kappa(w) \phi(c - w) \, dw} = \frac{\int_a^{a+y} \kappa(w) \phi(c - w) \frac{\phi(c + t - w)}{\phi(c - w)} \, dw}{\int_a^{a+y} \kappa(w) \phi(c - w) \, dw} = \int_a^{a+y} \frac{\rho_c(w) \phi(c + t - w)}{\phi(c - w)} \, dw.
\]

Note that \( \frac{\phi(c + t - w)}{\phi(c - w)} \) is increasing in \( w \), implying that \( \max_{a \leq w \leq a + y} \frac{\phi(c + t - w)}{\phi(c - w)} = \frac{\phi(c + t - a - y)}{\phi(c - a - y)} \). Consequently,

\[
\frac{h(c + t)}{h(c)} \leq \int_a^{a+y} \rho_c(w) \phi(c + t - a - y) \phi(c - a - y) \, dw = \frac{\phi(c + t - a - y)\phi(c - a - y)}{\phi(c - a - y)}.
\]

It follows that

\[
0 \leq \lim_{c \to \infty} \frac{h(c + t)}{h(c)} \leq \lim_{c \to \infty} \frac{\phi(c + t - a - y)}{\phi(c - a - y)} = 0,
\]

where the last equality follows from Lemma 13. Thus, \( \lim_{c \to \infty} \frac{h(c + t)}{h(c)} = 0 \). In a similar way it can be shown that \( \lim_{c \to -\infty} \frac{h(c + t)}{h(c)} = \infty \). \( \blacksquare \)

The next lemma establishes an additional useful property of the normal density.

**Lemma 15** Let \( t > 0 \). \( \lim_{c \to \infty} \int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz = 0 \) and \( \lim_{c \to -\infty} \int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz = \infty \).

**Proof of Proposition 15:** For any \( \varepsilon \in (0, t) \) and \( c \) sufficiently large we have

\[
\int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz = \int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz + \int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz \\
\leq \varepsilon + \frac{\phi(c + \varepsilon)}{\phi(c)} (t - \varepsilon) \quad (\text{since} \quad \frac{\phi(z)}{\phi(c)} \leq 1 \text{ on } [c, c + \varepsilon]).
\]

Hence,

\[
0 \leq \lim_{c \to \infty} \int_c^{c+t} \frac{\phi(z)}{\phi(c)} \, dz \leq \varepsilon + \lim_{c \to \infty} \frac{\phi(c + \varepsilon)}{\phi(c)} (t - \varepsilon) = \varepsilon \quad (\text{by Lemma 13}).
\]
But, this is true for all $\varepsilon > 0$ and hence $\lim_{c \to \infty} \int_c^{c+t} \frac{h(z)}{h(c)} \, dz = 0$. A similar argument applies to the other limit. ■

We can now establish the next lemma.

**Lemma 16** Assume that $h(z)$ is given by (34). For all $t > 0$, $\lim_{c \to \infty} E(z - c|z \in [c, c + t]) = 0$ and $\lim_{c \to \infty} E(z - c | z \in [c, c + t]) = 0$.

**Proof of Lemma 16:**

$$\lim_{c \to \infty} E(z - c|z \in [c, c + t]) = \lim_{c \to \infty} \frac{\int_c^{c+t} (z - c) h(z) \, dz}{\int_c^{c+t} h(z) \, dz}.$$  

By Lemma (12) both the numerator and the denominator tend to 0. By L’Hopital’s rule:

$$\lim_{c \to \infty} E(z - c|z \in [c, c + t]) = \lim_{c \to \infty} t \cdot \frac{h(c + t) - \int_c^{c+t} h(z) \, dz}{h(c) - h(c + t)} = t \lim_{c \to \infty} \frac{1}{1 - \frac{h(c)}{h(c + t)}} - \lim_{c \to \infty} \frac{\int_c^{c+t} h(z) \, dz}{h(c) - h(c + t)}.$$

The first limit is 0 by Lemma 14. Consider the second limit. The denominator tends to $-1$ by Lemma 14. We claim that the numerator tends to zero. Indeed,

$$\int_c^{c+t} \frac{h(z)}{h(c)} \, dz = \int_c^{c+t} \frac{\int_a^{a+y} \kappa(w) \phi(z - w) \, dw}{\int_a^{a+y} \kappa(w) \phi(c - w) \, dw} \, dz = \int_c^{c+t} \frac{\int_a^{a+y} \phi(c - w) \kappa(w) \phi(z - w)}{\int_a^{a+y} \phi(c - w) \kappa(w) \phi(c - w) \, du} \, du \, dz.$$

However, by Lemma 15, $\lim_{c \to \infty} \int_c^{c+t} \phi(z-a-y) \, dz = 0$, implying that $\lim_{c \to \infty} \int_c^{c+t} \frac{h(z)}{h(c)} \, dz = 0$. It follows that $\lim_{c \to \infty} E(z - c|z \in [c, c + t]) = 0$ as required. The other limit is proved in a similar manner. ■

**Proof of Proposition 4**

At each period $t = 1, \ldots, T$ the manager maximizes (21). Hence, the objective function of a manager in period $t$ is identical to that of the static model, except that he takes into account that investors condition the market price on all prior dividends. We will show existence of equilibrium by induction. The case $t = 1$ coincides with the static case. Using Proposition 2, choose some $\delta_1 < \Delta^*$ and consider the corresponding $\delta_1$-partially pooling equilibrium with a pooling interval $[a_{\delta_1}, b_{\delta_1}]$.  

50
Consider now the case $t = 2$. If a pooling dividend was paid in period 1, then earnings in period 2 follow a distribution that is the sum of a random variable with support $[a_{\delta_1}, b_{\delta_1}]$ and $\varepsilon_1$. The density of this sum can be written using the convolution in (34). Lemma 16 then shows that the statement of Lemma 7 applies to this distribution. It follows that the results of Lemma 4 hold for this distribution and hence so do the results of Proposition 2. That is, there exists $\delta_2 < \Delta^*$ and a corresponding $\delta_2$-equilibrium with a pooling interval $[a_{\delta_2}, b_{\delta_2}]$ for the second period.

We now continue this process for all $t = 3, \ldots, T$. In each period we get a distribution of earnings that is a sum of a random variable with a bounded support and a normal noise, and the same procedure applies. This process generates a sequence $\delta_1, \delta_2, \ldots, \delta_T$ and corresponding pooling intervals. The myopia of the manager and investors guarantees that the sequence of short-term equilibria established in each period is an equilibrium in the dynamic game. ■