Dividends as Reference Points: A Behavioral Signaling Approach

Malcolm Baker
Harvard Business School and NBER

Brock Mendel
Harvard University

Jeffrey Wurgler
NYU Stern School of Business and NBER

We outline a dividend signaling model that features investors who are averse to dividend cuts. Managers with strong unobservable cash earnings pay high dividends but retain enough to be likely not to fall short next period. The model is consistent with a Lintner partial-adjustment model, modal dividend changes of zero, stronger market reactions to dividend cuts than increases, comparatively infrequent and irregular repurchases, and a mechanism that does not depend on public destruction of value, which managers reject in surveys. New tests involve stronger reactions to changes from longer-maintained dividend levels and reference point currencies of American Depository Receipt dividends. (*JEL* D82, G35)

Managers share a number of common views about their dividend policies, as shown in the survey by Brav et al. (2005). They strive to avoid reducing dividends per share (of the 384 managers surveyed, 93.8% agreed); they try to maintain a smooth dividend stream (89.6%); and they are reluctant to make changes that might have to be reversed (77.9%). They follow such policies because they believe that there are negative consequences to reducing dividends (88.1%), which they believe convey information to investors (80%). While caution is merited in interpreting survey responses, the Brav et al. (2005)

results are further consistent with Lintner's (1956) own survey and interviews, his partial-adjustment model, and a large empirical literature demonstrating a significant response to dividend announcements.

While managers view dividends as some sort of signal to investors, they also cast doubt on the mechanisms of standard dividend signaling models. For example, the proposition that dividends are used to show that their firm can bear costs such as borrowing external funds or passing up investment was summarily rejected (4.4% agreement, the lowest in the entire survey). The idea of signaling through costly taxes did not receive much more support (16.6%). Again, while we might not expect managers to admit public destruction of value even in an anonymous survey, these findings suggest there is more to the story than the economic mechanisms driving well-known signaling models such as Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), Allen, Bernardo, and Welch (2000), and Guttman, Kadan, and Kandel (2010).

In this paper we use loss aversion, a feature of the prospect theory value function of Kahneman and Tversky (1979), to motivate a behavioral signaling model. A loss-averse value function has a kink at the reference point whereby marginal utility is discontinuously higher in the domain of losses. Loss aversion is supported by a considerable literature in psychology, finance, and economics, as we briefly review later.

The essence of our stylized model is that investors evaluate current dividends against a psychological reference point established by past dividends. Because investors are particularly disappointed when dividends are cut, dividends can credibly signal information about earnings. The model is inherently multiperiod, which leads to more natural explanations for the survey results above and other facts about dividend policy such as the Lintner partial-adjustment model, which emerges in equilibrium, and which static signaling models cannot address. While it is difficult to measure investor utility functions per se outside the laboratory, we perform some novel tests that get at the core intuitions of the approach.

To provide a bit more detail, the model uses reference point preferences as the mechanism for costly signaling. The manager's utility function reflects a preference both for a high stock price today and for avoiding a dividend cut in the future. In the first period, the manager inherits an exogenous reference level dividend, and receives private information about earnings. The manager balances the desire to signal current earnings by paying higher dividends with the potential cost of not being able to meet or exceed a new and higher reference point through the combination of savings from the first period and random second-period earnings. In equilibrium, managers that cannot meet the inherited dividend level pay out everything in the first period, as the marginal cost of missing the reference point is high; managers with intermediate first-period earnings pool to pay the reference dividend; and managers with strong first-period earnings pay out a fraction that raises the reference level for the future

but, given their savings and expected second-period earnings, to a level they are relatively confident that they can maintain.

This simple model is consistent with several facts about dividend policy that cannot be handled in static models. First, the modal dividend change is zero. In a rational continuous setting, there is no special significance to paying the same dividend as last period.[1] Second, for reasonable parameter values, firms with high or stable earnings engage in a partial-adjustment policy that resembles the Lintner model. Third, firms are punished more for dividend cuts than they are for symmetric raises, and so avoid raising the dividend to a level that will be difficult to sustain. Fourth, the approach offers an explanation for why repurchases are less frequent than dividends despite their tax advantage: unlike dividends per share, the parameters of a hypothetical "regular" repurchase program cannot be specified in salient and repeatable numbers. Fifth, the approach is intuitively compatible with the widespread practice of paying dividends in round numbers. Finally, the mechanism of the model is novel and not inconsistent with the available survey evidence. Strong types do not publicly burn money with certainty, but rather they implicitly burn expected utility by risking falling short the next period; for reasonable parameter values, actual utility burning by strong firms does not usually occur in equilibrium.

After confirming the result that investors respond asymmetrically to dividend increases and cuts, we proceed to our main test. The general intuition of signaling around a movable reference point suggests the comparative static that the signaling power of dividends, in terms of market reactions to changes, is stronger when reference point effects are stronger. While we do not model memory and salience, we hypothesize that the repetition of a particular dividend level tends to ingrain a reference point. Consistent with this hypothesis, we find that the asymmetric response to dividend changes is more pronounced when the new dividend breaks a multi-year streak of quarterly dividends per share.

We also conduct a brief placebo test involving American Depository Receipts (ADRs). In this sample, there is no clustering and nothing special about the market's reaction around zero dividend changes in terms of U.S. dollars. Rather, the reference point is denominated in foreign currency. Investors have settled on a particular reference point currency even though the dividend's economic value to some of them, and perhaps the firm's ability to pay that value, fluctuates with the exchange rate.

Our approach complements other papers that connect dividends and reference points or prospect theory. Shefrin and Statman (1984) argue that dividends improve the utility of investors with prospect theory value functions if they also mentally account (Thaler 1999) for dividends and capital gains and losses separately. However, in their work, dividends serve no signaling

---

[1] Guttman, Kadan, and Kandel (2010) show that an extension of the Miller-Rock model can generate sticky dividends, but point out that such equilibria are just a subset of multiple pooling equilibria. In the reference point approach, by contrast, the reason to focus on sticky dividend levels is straightforward.

purpose, and the model generates no particular relationship to a Lintner policy. Another recent contribution is Lambrecht and Myers (2012). In their model, managers maximize the present value of the utility of rents that they extract from profits. They smooth dividends and prefer a Lintner policy because they have habit formation preferences and rents move with dividends given the budget constraint. We discuss these and related papers. More generally, our paper adds to the literature on behavioral corporate finance surveyed in Baker and Wurgler (2012).

The first section reviews the relevant literature on reference-dependent utility. The second section describes the model. The third discusses its compatibility with known facts of dividend policy, as well as some extensions of those facts. The fourth conducts a novel test and a placebo test. The fifth section concludes.

1. Background: Reference-Dependence and Reference Points

In the time since Markowitz (1952) and Kahneman and Tversky (1979) proposed theories of choice based on utility that depends not only on the level of economic states, but on changes, the literatures on empirical choice behavior and the psychological analysis of value have advanced considerably, as have their applications to economics and finance. We first review this literature more generally and then apply it to dividends.

1.1 Reference-dependent utility and loss aversion

We focus on two central features of the prospect theory value function: that utility depends on changes in states relative to a reference point, and that losses bring more pain than symmetric gains bring pleasure. Our applications to dividends do not require a full review of prospect theory, which as a whole is a theory of choice under uncertainty.

Tversky and Kahneman (1991) review the classic literature on loss aversion. Kahneman and Tversky (1979) introduced loss aversion to reflect then-known patterns in choice behavior. The subsequent literature suggests its relevance in a wide range of applications. One implication of loss aversion is what Thaler (1980) termed the endowment effect. Kahneman, Knetch, and Thaler (1990) found that the value of an item increases when it is considered already in one's endowment. A literature has developed on differences between the willingness to pay for a small improvement versus willingness to accept a small loss, another reflection of loss aversion. (These literatures suggest the ballpark figure that losses matter slightly more than twice as much as gains, a figure that we employ in our numerical example.) Finally, a related phenomenon is status quo bias. Samuelson and Zeckhauser (1988) documented a preference for the status quo even when costs of change are small relative to potential benefits, such as in choices about medical plans.

1.2 Reference points

Several aspects of reference points in the context of dividend policy deserve discussion. If "gains" and "losses" matter, how are they defined? In other words, what is the reference point and how is it formed? Can it change? What determines its strength? Can there be multiple reference points? The literature on loss aversion does not provide general answers to these questions. The relevant reference point depends on the setting. In static choice situations, it is often obvious. For example, in the applications and experiments above, the reference point is the decision maker's current position. But in many circumstances, "current position" is not always so well defined. In Abel (1990), for example, the reference point for utility includes others' current consumption levels.

A more complicated situation arises when the decision maker has some control over the framing of an outcome. Thaler (1999) reviews the concept of mental accounting, in which the decision maker may, for example, choose to define reference points and segregate outcomes so as to strategically maximize his happiness under a value function defined over gains and losses.

Shefrin and Statman (1984) apply these ideas to explain why investors like dividends, although their perspective is very different than ours. Shefrin and Statman argue that investors prefer to mentally divide returns into capital gains and dividends and consider each separately. Their explanation employs a third feature of the prospect theory value function—its concavity in gains and convexity in losses. Dividends allow investors to flexibly repackage what would otherwise be a large capital loss into a slightly larger capital loss and a dividend. If the capital loss is large, then a slightly larger loss causes little extra pain, while the dividend can be accounted for as a gain relative to a reference point of no dividend and thus a return to the value function where marginal utility is high. Likewise, if there is a large positive return, making the capital gain slightly smaller does not decrease utility much, while the ability to treat the dividend as a separate gain allows for an additional, disproportionate utility increase.

Reference points can also differ in their temporal character. In dynamic situations with uncertainty, the reference point is even harder to generalize about. It may involve the future, not just the present. In Koszegi and Rabin (2006, 2009), agents are loss averse over changes in beliefs about future outcomes such as consumption. Here, expectations about the future make up the reference point. For example, utility might depend in part on the prospect of a raise.

Past circumstances can also supply powerful reference points. Genesove and Mayer (2001) find that people resist selling their homes below the purchase price. Shefrin and Statman (1985) find that the purchase price of a security serves as a reference point. Odean (1998) confirms this, and also suggests, like Arkes, Hirshleifer, Jiang, and Lim (2008), that such reference points can change over time, albeit sluggishly. Baker and Xuan (2015) argue that the stock price

that a new CEO inherits is an important reference point for raising new equity. The idea of one's prior consumption as a reference point for the utility of current consumption is represented through internal habit formation preferences as in Constantinides (1990).

In settings where the past supplies the reference point, its power may depend on the strength of the associated memory. Most of the literature does not incorporate the role of memory, however. A probability distribution is not memorable, and a rational expectation about the future is going to be continuous and somewhat indeterminate. The particulars of past consumption levels may not be memorable. In general, factors that increase the strength of a memory include repetition and rehearsal (Atkinson and Shiffrin 1968), elaboration (Palmere et al. 1983), distinctiveness (Eysenck and Eysenck 1980), salience, associated effort (Tyler et al. 1979), or emotional association. For individual numbers, ease of recall matters. Phone companies sell phone numbers that include round numbers or repeated digits at a premium.

A stock's 52-week high provides an interesting example of a memorable number that for some purposes forms a reference point. The shareholder may have a positive association with that level. It is a specific and salient number. It can be constant (repeated and rehearsed) for up to 52 weeks, but also varies over time. Heath, Huddart, and Lang (1999) find that employee exercise of stock options doubles when the stock price tops its 52-week high. Recent peak prices are important for the pricing and deal success of mergers and acquisitions (Baker, Pan, and Wurgler 2012).

1.3 Past dividends as reference points

This discussion shows that theory alone cannot identify "the" reference point. In this paper we hypothesize that in the context of dividend policy, past dividends are reference points against which current dividends are judged. Our hypothesis touches on many of the concepts discussed above. The reference point we hypothesize is based on past experience, as in the disposition effect of Shefrin and Statman (1985). It is also dynamic, as in internal habit formation. Fluctuations in the dividend upset expectations about future dividends. Baker, Nagel, and Wurgler (2007) find that many investors consume the full amount of their dividends, drawing attention further to their level.

Dividends are also packaged to be memorable. They are formally and explicitly announced at discrete and regular intervals, which encourages the formation of memory. The same level is often repeated for many quarters in a row, further encouraging memory. Anecdotally, they tend to be announced with special fanfare and management commentary upon initiation or increases. Conversely, dividend decreases are downplayed, accompanied by excuses or explanations that frame the action as repositioning for growth. Dividend cuts are disproportionately announced on Fridays and often after the market close (Damodaran 1989). As we shall see, dividends cluster at round numbers, and changes are commonly in round-number intervals or designed to meet or exceed

a round-number threshold. The memorability of prior dividends is central to our theory—it increases their power as reference points and, consequently, current dividends as signals.

2. A Model of Signaling with Dividends as Reference Points

We present a stylized dividend signaling model with reference dependence. The model uses nonstandard investor preferences, not public destruction of firm value through investment distortions or taxes, to provide the costly signaling mechanism.

There are two key ingredients. First, a reference point appears in a representative investor's objective function. There is a kink in utility, so that a drop in dividends just below the reference point is discontinuously more painful than a small increase in dividends is pleasurable. Second, in common with all signaling models, the manager cares about the current estimate of firm value as well as the long-term welfare of shareholders.

Reference points shape dividend policy in several ways. On one hand, to the extent that today's dividend is the reference point against which future dividend payments will be judged, the manager would like to restrain current dividends, saving some resources for the next period to make up for a possible shortfall in future income. On the other hand, setting aside effects on future shareholder welfare, the manager would like to pay a dividend today that exceeds the current reference point. Moreover, because the manager also cares about the current estimate of firm value, he might also increase dividends beyond the current reference point to signal private information about the firm's ability to pay. This sort of signaling mechanism works because firms with limited resources are unwilling to incur the expected future costs of missing an endogenous reference point.

2.1 Setup
The model has two periods: $t = 1$ and 2. There are two players: a manager and an investor with reference-dependent preferences. In the first period, the investor arrives with an exogenous reference point for dividends d_0. In some ways, this is a single snapshot in a multiperiod model. While the inherited reference point could in principle be endogenized, we believe the technical costs would be large compared with the benefits in extra realism or intuition. The manager also receives private information about cash earnings ε_1 and pays a dividend d_1 in the first period. This dividend forms a new reference point for the second, liquidating dividend d_2. Today's dividend d_1 relative to d_0 tells the investor something about the manager's private information and hence the value of the firm. In the second period, the manager simply pays d_2.

2.1.1 Manager utility. The manager cares about what the investor thinks about ε_1 today. He also cares about the investor's long-run utility. This objective

function is in the spirit of standard signaling models like Leland and Pyle (1977), Miller and Rock (1985), or Stein (1989), which use weighted averages of the dividend-adjusted stock market price and the investor's long-run utility. Our simplified objective function is:

$$V \equiv E_m[\lambda E_i[\varepsilon_1] + (1-\lambda)u(d_1, d_2|d_0)], \tag{1}$$

where d_1 and d_2 are the period-specific dividends of the firm, u is the investor's utility function, given an exogenous initial reference point of d_0, and E_m and E_i are the expectations operators for the manager and the investor, respectively.[2]

2.1.2 Investor utility. The interesting aspect of this signaling model is that the investor has a kink in his preferences for dividends d_1 and d_2. The first kink is around an exogenous reference point for first-period dividends d_0, and the second kink is around an endogenous reference point for second-period dividends:

$$u(d_1, d_2|d_0) = d_1 + b(d_1 - d_0)\mathbf{1}_{d_1 < d_0} + \beta d_2 + \beta b(d_2 - d_1)\mathbf{1}_{d_2 < d_1}. \tag{2}$$

Put simply, the investor cares about fundamental value, or total dividend payments, but with a twist. The level of the reference point comes from historical firm dividend policy, and b is greater than zero so that the marginal utility of dividends falls discontinuously at the reference point; the pain of coming up short is high. This utility function reflects loss aversion with a kink at a reference point. We leave out the complexity of curvature.[3] $\beta \leqslant 1$ is a discount rate. The second-period reference point equals first-period dividends d_1 by assumption. In reality, the reference point and the intensity of the reference point b may be determined by a long history of levels and changes in dividend policy. The fact that each dividend payment forms a separate reference point also requires a practice of narrow framing. This is not a reference point applied to total ending wealth, but rather a reference point applied much more narrowly both across stocks and time, in the spirit of Barberis, Huang, and Thaler (2006).

---

[2] As in the standard signaling models, we do not microfound λ. The usual argument for this general class of utility functions is that the adjusted stock price, separate from fundamentals, has a direct impact on the manager's welfare through compensation or corporate control or an indirect impact through the interests of short-term investors. Also, rather than compute a stock price, we put the investor's expectation of ε_1 directly into the manager's objective. This is an innocuous assumption, because in equilibrium the stock price will be a linear transformation of the expectation of ε_1.

Note that there is no investment decision, so the firm is a simple payout machine. This is entirely for tractability and to focus on the reference-dependent preferences of investors. If investment were not observable, it could in principle be another dimension of the signaling problem as in Bebchuk and Stole (1993). If investment is observable with a convex cost of suboptimal investment, it can interfere with the specific form of the equilibrium. We conjecture that no firm would miss the reference point by a small margin. In the event that dividends fall below the reference point, there would be a gap, where dividends would fall below current profits by an amount that accommodated at least some investment.

[3] We conjecture that adding convexity in the domain of losses in the utility function helps to explain why firms that cut dividends tend to do so by a larger amount. That realistic and empirically relevant element of investor preferences is left out of the theoretical development.

It is worth noting at this point that many of the features of the equilibrium described below do not require both a $\beta > 1$ and a $\lambda < 0$. Either one will suffice to generate a managerial incentive to pay dividends sooner rather than later. The important trade-off is between maximizing current dividends and avoiding future cuts. Either parameter restriction is sufficient to generate the incentive to maximize current dividends.

2.1.3 Budget constraint. There is no new equity or debt available to finance the payment of dividends and no excess cash balances available in the first period. The most the manager can pay in the first period is ε_1, and the most he can pay in the second period is ε_2 plus any savings from the previous period. This leads to the following constraints:

$$0 \leqslant d_1 \leqslant \varepsilon_1 \text{ and } d_2 = \varepsilon_1 + \varepsilon_2 - d_1. \tag{3}$$

These follow from the assumptions of no new financing and a manager that pays out all earnings.

2.1.4 Information. In the absence of a signaling motive, the first-best would have the manager saving any resources above the first-period reference point, both to lower the reference point created for the second period and to save resources to meet this lower reference point in the event of low second-period earnings. We now introduce the signaling problem. For simplicity, the manager has no control over the cash earnings of the firm. Note that this is a bit different from a traditional signaling model where the manager must destroy firm value to impress the capital markets.[4] There is also no agency problem of the sort examined in Lambrecht and Myers (2012), where the manager can keep cash for himself. This setup is, at least in spirit, more consistent with what managers say in surveys about their dividend policy.

The fundamental value of the firm appears in two installments,

$$\varepsilon_1 + \varepsilon_2. \tag{4}$$

Think of these as cash earnings that are not observable to the investor. This is obviously an extreme assumption of asymmetric information. It is worth noting the key elements of the assumption, which might each seem more reasonable. First, the manager must have some informational advantage over the investor in learning ε_1. Otherwise, there is no signaling problem. Second, the payment of the observable dividend must form the investor's reference point, not the firm's reported financials, such as earnings per share or cash balances. Otherwise, the manager has no lever to signal his information about ε_1. For simplicity,

---

[4] In any model that requires costly external finance as a production input, dividends destroy value through the budget constraint. We wish to distinguish standard signaling models where this is the driving mechanism, as opposed to a generic side effect.

we assume that the second-period cash earnings have a uniform distribution, $\varepsilon_2 \sim U[0,2]$.

We have considered extensions of the model where the source of the information asymmetry is over ε_2, a quantity that would not appear explicitly in any financial statements. This assumption produces similar results, although the effects of the budget constraint described in the next paragraph can change. The simpler model that we examine here has a mechanical link between type in terms of firm quality and current resources. Another positive feature of this setup is that it builds nicely on the empirical evidence of Benartzi, Michaely, and Thaler (1997) and Grullon, Michaely, and Swaminathan (2002), who show that dividends are not useful signals of future earnings growth.

2.1.5 Beliefs. The investors' beliefs are determined in equilibrium by a rational expectations condition. This can in some cases induce unreasonable equilibria. For instance, suppose that investors believe ε_1 equals the mean of the prior distribution if $d_1 = 0$, and believe that $\varepsilon_1 = -10^{100}$ otherwise. In this situation, managers will set $d_1 = 0$ regardless of ε_1. Because $d_1 \neq 0$ is never observed in equilibrium, the investors' beliefs satisfy a rational expectations equilibrium. To rule out these perverse equilibria, we restrict attention to Trembling Hand Perfect Equilibria. The relevant restriction on beliefs is that the expectation of ε_1 is greater than or equal to the manager's choice of d_1, which follows directly from the budget constraint. Without loss of generality, we make a further and intuitive restriction on the investor's out-of-equilibrium beliefs. If there is a level of dividends d_1 equal to d that is not observed in equilibrium, then we can modify this equilibrium so that the investor's conditional expectation of ε_1 is simply equal to d. Because no manager pays d in equilibrium, this does not cause any deviations.

2.2 Solving the Model

An equilibrium is characterized by the manager's choice of d_1 and the investor's beliefs over ε_1 conditional on d_1. The manager takes these beliefs as given when solving his maximization problem. To simplify notation, we define $F(d_1, d_0)$ as the investor's expectation of ε_1 when the manager pays d_1. Substituting the investor's utility function from Equation (2) into the manager's objective in Equation (1), applying the law of iterated expectations, using the assumption of a uniform distribution for ε_2, and simplifying leads to:

$$V \equiv \lambda F(d_1, d_0) + (1-\lambda)\Big(1 + \varepsilon_1 + (1-\beta)d_1 + b(d_1 - d_0)\mathbf{1}_{d_1 < d_0}$$

$$- b\beta\Big(d_1 - \tfrac{\varepsilon_1}{2}\Big)^2 \mathbf{1}_{d_1 > \frac{\varepsilon_1}{2}}\Big). \tag{5}$$

The first term is weighted by λ and measures the impact of the dividend decision on the investor's expectation over current profits, and the second term is weighted by $(1-\lambda)$ and measures the impact of the dividend decision on the

investor's long-run utility. We relabel this second term as $V_0(d_1, d_0)$. Note that V_0 is differentiable in d_1 everywhere except at d_0 and $\frac{\varepsilon_1}{2}$. The first derivative of V_0 breaks down into four regions as follows:

$$
V_0' = \begin{cases} 1 - \beta + b & d_1 < d_0, \frac{\varepsilon_1}{2} \\ 1 - \beta + b - 2\beta b\left(d_1 - \frac{\varepsilon_1}{2}\right) & \frac{\varepsilon_1}{2} < d_1 < d_0 \\ 1 - \beta & d_0 < d_1 < \frac{\varepsilon_1}{2} \\ 1 - \beta - 2\beta b\left(d_1 - \frac{\varepsilon_1}{2}\right) & d_0, \frac{\varepsilon_1}{2} < d_1 \end{cases} \quad \text{where} \quad (6)
$$

It is also useful to note that V_0 is supermodular. That is, marginal utility is weakly increasing in ε_1, or $\frac{\partial^2 V_0}{\partial d_1 \partial \varepsilon_1} \geqslant 0$.

2.2.1 Characterizing equilibria.
We define an equilibrium as a mapping from manager types ε_1 to dividend d_1, or $d_1(\varepsilon_1)$. If there exists a level of dividends such that $d(\varepsilon_1) = \overline{d}$ for all $x_1 \leqslant \varepsilon_1 \leqslant x_2$ and $x_1 \neq x_2$, call $\overline{d}$ a pool.

2.2.2 Out-of-equilibrium beliefs.
We also assume without loss of generality that any out-of-equilibrium beliefs are simply $F(d_1, d_0) = d_1$. To show that this is without loss of generality, suppose that there exists some value $d_1 = d$ such that d is not paid by any type of manager in equilibrium. Consider a modification of this equilibrium in which the investor's conditional beliefs $F(d, d_0)$ are (weakly) lowered to d. This makes paying d (weakly) less attractive to any type of manager. Because no manager pays d in equilibrium, this does not induce any deviations, and this new set of beliefs is also an equilibrium.

This is enough to narrow down the possible equilibria somewhat. Using the fact that a manager's marginal utility of dividends is increasing in first-period profits ε_1, we can show that any equilibrium must feature dividends that rise with first-period profits ε_1, provided that there is at least some discounting. The proof of Proposition 1 is in Appendix A.

Proposition 1. Dividends are weakly monotonic in profits. The equilibrium level of dividends $d_1 = d_1(\varepsilon_1)$ is a non-decreasing function of ε_1.

We make a simplifying assumption, restricting the range of model parameters. With this restriction, we guarantee that marginal utility in Equation (6) is only negative when d_1 is high relative to both d_0 and $\frac{\varepsilon_1}{2}$. This will be true as long as d_0 is not too large and b is not too small. Intuitively, this assumption means that the effect of the reference point is not trivial, and it serves two purposes. It guarantees that the marginal utility shown in the $\frac{\varepsilon_1}{2} < d_1 < d_1$ portion of Equation (6) is positive when d_1 is less than d_0. Second, it ensures that the signaling game does not break down for high-profit managers. If the cost of signaling is too low, there is no effective way to separate high-profit from lower-profit managers.

Assumption 1. d_0 is not too large and b is not too small. The exogenous parameters of the model must satisfy $d_0 \leqslant \frac{1-\beta+b}{\beta b}$.

2.2.3 Separating behavior. The first step is to derive the two regions of separating behavior. Because the level of dividends paid as a function of profits $d_1(\varepsilon_1)$ is weakly monotonic given Proposition 1, it must also be differentiable almost everywhere. So, focusing for the moment on areas where the dividend function is differentiable, we can write the derivative of the investor belief function $F'(d_1,d_0)$ with respect to d_1 as $F'(d_1,d_0)=d_1'(\varepsilon_1)^{-1}$. Differentiating the manager's objective in Equation (7) with respect to dividends paid d_1 leads to two possibilities. The first is that the manager is constrained at a corner solution by the budget constraint, so that $d_1(\varepsilon_1)$ simply equals ε_1, and the second is that the following interior first-order condition holds:

$$0=\lambda F'(d_1,d_0)+(1-\lambda)\left((1-\beta)+b\mathbf{1}_{d_1<d_0}-2b\beta\left(d_1-\tfrac{\varepsilon_1}{2}\right)\mathbf{1}_{d_1>\frac{\varepsilon_1}{2}}\right), \quad (7)$$

Rearranging terms and substituting for F' gives the following differential equation:

$$2b\beta\left(d_1(\varepsilon_1)-\tfrac{\varepsilon_1}{2}\right)\mathbf{1}_{d_1>\frac{\varepsilon_1}{2}}=\frac{\lambda}{1-\lambda}\frac{1}{d_1'(\varepsilon_1)}+(1-\beta)+b\mathbf{1}_{d_1<d_0}. \quad (8)$$

So, any separating points in the equilibrium must either have dividends binding at the budget constraint or Equation (7) must hold at $d_1(\varepsilon_1)>\frac{\varepsilon_1}{2}$. Lemmas 1 and 2 define the two regions of separation.

Lemma 1. Separating behavior below d_0 takes the form $d_1(\varepsilon_1)=\varepsilon_1$. Formally, suppose that $d_1(\varepsilon_1)$ is strictly monotonic in the neighborhood of x with $d_1(\varepsilon_1)>d_0$, then $d_1(\varepsilon_1)=\varepsilon_1$.

This follows from Assumption 1, which guarantees that the marginal utility in Equation (6) is positive everywhere in the range where $d_1<d_0$.

Lemma 2. Separating behavior above d_0 can take the form $d_1(\varepsilon_1)=\frac{\varepsilon_1}{2}+\alpha$.

This follows directly from differential Equation (8), evaluated in the range where $d_1>d_0$. A solution to this equation is:

$$d_1(\varepsilon_1)=\frac{\varepsilon_1}{2}+\frac{1}{b\beta}\left(\frac{\lambda}{1-\lambda}+\frac{1-\beta}{2}\right). \quad (9)$$

In other words, managers pay out a fixed fraction of profits plus a constant to differentiate themselves, at a cost in terms of establishing a higher reference point that might not be met.

2.2.4 Pooling behavior. In general, there will be a pool in equilibrium. If there were no pool, then by definition, there would be separating behavior everywhere. Proposition 1 and separating behavior everywhere imply that dividends are strictly monotonic. Lemma 1 says that separating behavior below takes the form $d_1(\varepsilon_1) = \varepsilon_1$ when $d_1 < 0$. Strict monotonicity means that $d_1(d_0) = d_0$. Strict monotonicity, the budget constraint in Equation (3), and Lemma 2 mean that the right-hand side of Equation (9) must equal d_0 when evaluated at $\varepsilon_1 = d_0$. This pins down a knife-edge set of parameter values where there is no pool in equilibrium. In particular, we must have d_0 precisely equal to $\frac{2}{b\beta}\left(\frac{\lambda}{1-\lambda} + \frac{1-\beta}{2}\right)$. In all other cases, there will be a pool in equilibrium.

Proposition 2. There is a pool in equilibrium unless $d_0 = \frac{2}{b\beta}\left(\frac{\lambda}{1-\lambda} + \frac{1-\beta}{2}\right)$.

2.3 Characterizing the Equilibrium

Now, armed with the knowledge that there is generally a pool in equilibrium, we can attempt to characterize a reasonable set of pooling equilibria. It is worth noting that this is not a unique equilibrium. It is possible to have multiple pools and pools at levels other than d_0. However, a particularly natural equilibrium, given the salience of the reference point, involves three ranges of dividend policies. There is a high payout ratio of 100% or $d_1(\varepsilon_1) = \varepsilon_1$ for firms with the extra motivation to clear the initial reference point of d_0. Next, managers cluster at d_0 once this motivation drops out of Equation (6). Finally, there is a lower payout ratio characterized by Equation (9) for firms well above the initial reference point, who nonetheless pay higher dividends than the reference point to separate themselves from each other and from the pool at d_0. The form of the partial pooling equilibrium is not dissimilar to that in Guttman et al. (2010), though the nature of the signaling mechanism is quite different, and perhaps more importantly, in our approach it is clear why dividends pool at the prior period's level as opposed to some arbitrary level. We describe a specific equilibrium in more mathematical detail using Propositions 1 and 2 and a solution in Equation (9).

Proposition 3. There exists an equilibrium dividend policy in which

$$d_1 = \varepsilon_1 \quad for \ \varepsilon_1 < d_0$$
$$d_1 = d_0 \quad for \ d_0 < \varepsilon_1 < x_1$$
$$d_1 = \tfrac{1}{2}\varepsilon_1 + \tfrac{1}{b\beta}\left(\frac{\lambda}{(1-\lambda)} + \frac{(1-\beta)}{2}\right) \equiv \tfrac{1}{2}\varepsilon_1 + \alpha \ \ for \ \varepsilon_1 > x_1$$

with associated equilibrium beliefs of:

$$E_i[\varepsilon_1 \,|\, d_1] = d_1 \ for \ d_1 < d_0$$
$$E_i[\varepsilon_1 \,|\, d_1] = \tfrac{1}{2}(d_0 + x_1) \ \ for \ \ d_1 = d_0$$

$$E_i[\varepsilon_1 \mid d_1] = d_1 \ \text{for } d_0 < \ d_1 < \ x_1.$$
$$E_i[\varepsilon_1 \mid d_1] = 2(d_1 - \alpha) \ \text{for } d_1 \geqslant d_0.$$

Where x_1 solves a quadratic or a linear equation depending on the size of the pool:

$$b\beta\left(d_0 - \tfrac{x_1}{2}\right)^2 + \left(1 - \beta + \tfrac{\lambda}{1-\lambda}\right)\tfrac{x_1}{2} - \left((1-\beta)(d_0 - \alpha) + b\beta\alpha^2 + \tfrac{\lambda}{1-\lambda}\tfrac{d_0}{2}\right) = 0$$
$$\text{when } x_1 < 2d_0, \text{ or } x_1 = \frac{\tfrac{\lambda}{1-\lambda}d_0 + 2(1-\beta)(d_0 - \alpha) + 2b\beta\alpha^2}{\tfrac{\lambda}{1-\lambda} + (1-\beta)} \text{ when } x_1 \geqslant 2d_0.$$

The derivation of is x_1 in Appendix A. Incentive compatibility here requires a manager with $\varepsilon_1 < d_0$ to be willing to pay $d_1 = \varepsilon_1$. More importantly, a manager with $\varepsilon_1 = x_1$ must be indifferent between paying d_0 and paying $\tfrac{1}{2}x_1 + \tfrac{1}{b\beta}\left(\tfrac{\lambda}{(1-\lambda)} + \tfrac{(1-\beta)}{2}\right)$, which we define as $\tfrac{1}{2}x_1 + \alpha$ to simplify notation and show that in this range the manager is paying a fixed amount α above the dividend level that would sustain second-period dividends at the new reference point with no risk. For this to hold, the signaling benefit of shifting the investor's expectations from $\tfrac{1}{2}(d_0 + x_1)$ to x_1 must equal the cost differential of evaluating Equation (7) at d_0 and at $\tfrac{1}{2}x_1 + \alpha$. This leads to the simple quadratic or linear equation in the proposition. The specific formula for the top of the pool is relatively uninformative, except below we will argue that in general the size of the pool, meaning the size of x_1, will be increasing in the strength of the reference point b, provided that the parameters allow for an equilibrium of this type.

For example, suppose the manager cares equally about stock price and utility—that is, $\lambda = 0.5$. If d_0 is 1, b is 2.5, and β is 1, then Proposition 3 indicates that the equilibrium cutoff x_1 is 1.6. For ε_1 above 1.6, the first-period dividend is $\tfrac{1}{2}\varepsilon_1 + 0.4$. This exactly trades off the marginal signaling benefit per unit of dividends of 2.0, using investor beliefs implicit in Proposition 1, against the second-period marginal cost—that is, the first derivative of Equation (7), of $5 \times \left(d_1 - \tfrac{1}{2}\varepsilon_1\right) = 5 \times (0.4) = 2.0$. For ε_1 between 1.0 and 1.6, the first-period dividend is simply d_0, or 1.0. At ε_1 equal to 1.6, where there should be indifference, the signaling benefit of separating from this pool is 1.6 minus the average of $x_1 = 1.6$ and $d_0 = 1.0$, or 0.3. The cost from Equation (7) is $2.5 \times (0.4^2 - 0.2^2) = 0.3$. This cost is decreasing in ε_1, so there is no incentive for any of the managers clustered at $d_1 = 1.0$ to raise the first-period dividend. For ε_1 below 1.0, the first-period dividend is ε_1. Here, the manager is limited by the budget constraint. The marginal first-period benefit per unit of dividends of 2.5 plus the marginal signaling benefit per unit of dividends of 1.0, using investor beliefs implicit in Proposition 1, totals 3.5. This exceeds the second-period marginal cost just below d_0 of $5 \times (1 - 0.5) = 2.5$. So, dividends are at a corner solution of $d_1 = \varepsilon_1$ from Equation (3).

The intuition is straightforward. There is a powerful incentive to try to reach the existing reference point of d_0 both to satisfy the kinked investor utility and to raise investor beliefs discontinuously from d_0 to $\frac{1}{2}(d_0+x_1)$. There is a steep rise in dividends per unit of cash earnings, or a 100% payout ratio, below the reference point. Then, there is clustering at d_0. Even firms that could pay somewhat more choose not to, because of the costs of setting a high reference point for the second period. Eventually, there is a jump in dividends once cash earnings become sufficiently high. At that point, though, dividend policy is still fairly conservative, with managers saving a large fraction of dividends in reserve for the second period. In other words, there is a more gradual rise in dividends per unit of cash earnings. See Panel A of Figure 1.

Another way to see this is by plotting the histogram of dividend changes in Panels B and C of Figure 1. It simplifies graphical illustration to assume that first-period cash earnings ε_1 are normally distributed, with a mean of $\frac{1}{2}(d_0+x_1)$.[5] There is a spike in the distribution at the reference point in Panel B. Moreover, even when we remove this spike in Panel C, there is still a jump in the distribution moving from the range just to the left of the reference point to the range just to the right. The distribution of dividend changes otherwise has a lower and longer tail of larger dividend cuts to the left of the reference point.

Finally, we plot the "market reaction" to dividend announcements in Panel D of Figure 1. This is measured as the percentage change in expected utility in Equation (2) from before the announcement. Note that because utility takes a linear form, it can be thought of as a stock price, or the amount of cash the investor would give up prior to the dividend payment, in order to hold the firm's shares. An equivalent picture would emerge by simply plotting the change in the rational expectation of realized profits ε_1 because this is the source of the change in expected utility. The interesting behavior is in the narrow range around the reference point. The drop in utility per unit of dividends is steeper to the left of the reference point than to the right. Missing the reference point by just a tiny amount leads to a drop of $\frac{1}{2}(x_1-d_0)$ in the investor's expectation. But it takes a discontinuous increase in dividends of $\frac{1}{2}x_1+\alpha-d_0$ to achieve the equivalent increase in expectations. As a result, there is a kink in the reaction at exactly d_0.[6]

Next, we turn to comparative statics. In particular, we are interested in how these patterns change with a change in the cost b of falling below the reference point.

---

[5] Also, a sum of uniform variables converges quickly to a normal distribution, so if one conceives ε_1 as such a sum, then the relationship to the uniform distribution of ε_2 in the model is not entirely unnatural.

[6] It is important to note that this asymmetry is only in the neighborhood of zero. The slopes of changes in expected utility take a different pattern for differences that are away from zero. In the range of dividend cuts, the slope is in expected profits and utility is flatter, when compared with the slope in the range of dividend increases.
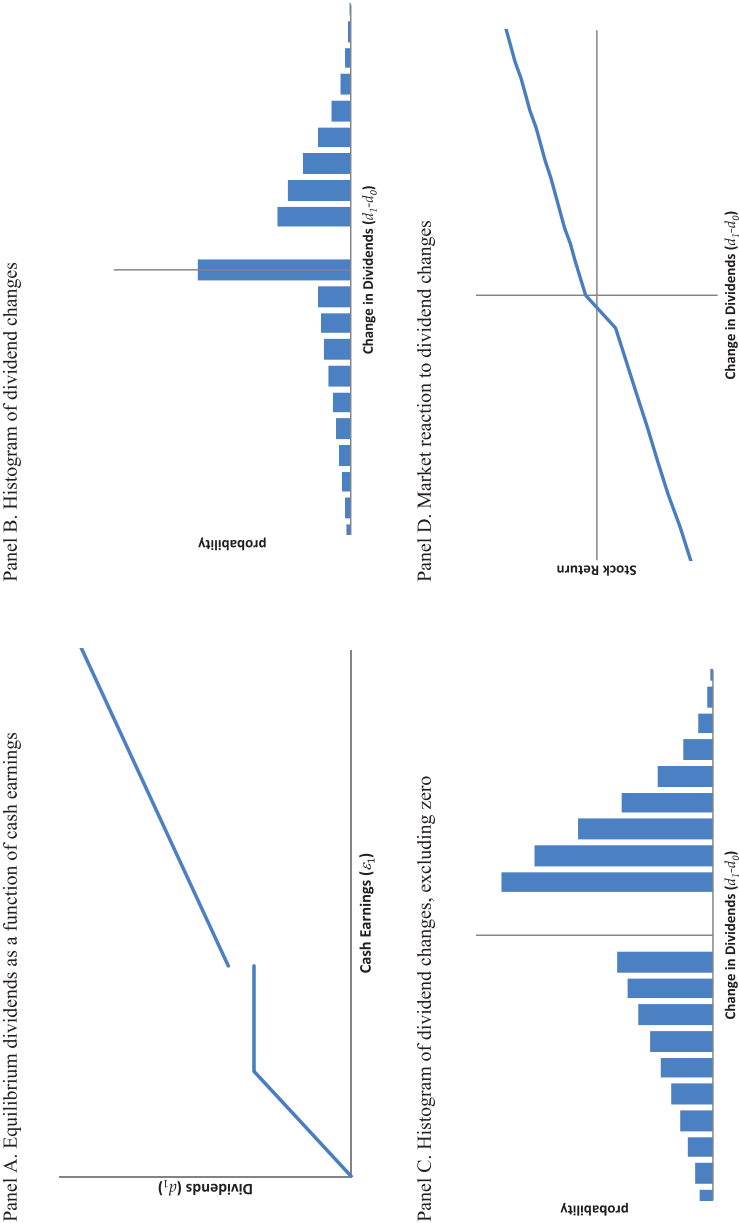
Panel A. Equilibrium dividends as a function of cash earnings

Panel B. Histogram of dividend changes

Panel C. Histogram of dividend changes, excluding zero

Panel D. Market reaction to dividend changes

Figure 1
Equilibrium dividend policy
We plot the equilibrium in Proposition 1 for parameter values $d_0 = 1$, $b = 2.1$, and $\lambda = 0.5$. Panel A shows the relationship between first-period dividends d_1 and cash earnings ε_1. Panel B plots the same data as a histogram. For this plot, we need the distribution of cash earnings ε_1. We assume a normal distribution with a mean equal to the average of d_0 and $\varepsilon*$, and a standard deviation of 0.15. Panel C reproduces Panel B, excluding dividend changes of zero. Panel D shows the market reaction to dividend changes. This is the percentage change in expected utility from Equation (2).

Proposition 4. In the equilibrium described in Proposition 3, x_1 and the market reaction to $d_1 < d_0$ are increasing in b.

The proof for the linear case is in Appendix A. Put another way, Proposition 4 says that there is more clustering of dividends at the reference point d_0 as the intensity of reference point preferences increases. As a result, the market reacts more negatively to a narrow miss. More information is revealed.

Again, we show this by example in Figure 2. The comparison is between b equal to 2.1 and b equal to 2.3. The exact b is not important. A similar equilibrium can be sustained at higher d_0 and lower b; these parameters make for clear pictures (as $b = 2.5$ in the proof yields round numbers). As we have no prior on the level of the reference point, this confirms that the assumptions required to support equilibrium here are not unreasonable ones. In each case, we re-center the ex-ante distribution of ε_1 at a mean of $\frac{1}{2}(d_0 + x_1)$ and repeat the plots from Figure 1. Note the effects of a higher b and hence x_1. There is more clustering and a larger spike in the distribution of dividend changes at d_0. The market reaction is more negative when dividends fall just short, and the reaction is flatter above it, so that the kink at zero is more pronounced.

2.4 Reference-dependent behavior

Similar results are obtained if we replace reference-dependent preferences with reference-dependent behavior. For example, we have analyzed a setup in which investors sell their shares to risk-averse arbitrageurs with a probability that rises as the dividend falls below a reference point. The possibility of a dividend-induced dislocation in share price creates the same three incentives for the manager: to restrain dividends to lower the hurdle for the future, to clear today's reference point and avoid the associated share price hit, and to increase dividends and tomorrow's reference point to signal firm quality. Such a model delivers additional predictions about volume, which we test below. Otherwise, however, it adds complexity and is somewhat further removed from the psychology of reference dependence. In practice, both types of reference dependence may exist.

3. The Model and Existing Evidence

We consider a number of stylized facts here in light of the model. We argue that it captures most of these facts at least as well as existing signaling approaches, such as Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), and Allen, Bernardo, and Welch (2000). To keep the discussion finite we will not make comparisons to models based on agency problems, catering motives, clientele effects, or other non-signaling mechanisms.
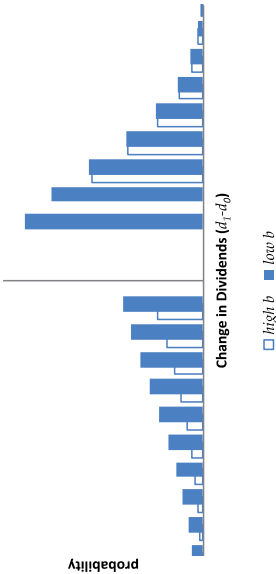
Panel A. Equilibrium dividends as a function of cash earnings
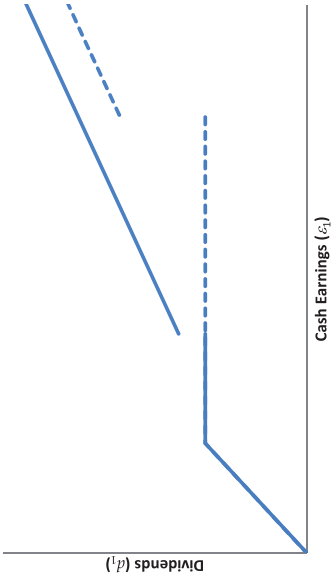
Panel B. Histogram of dividend changes
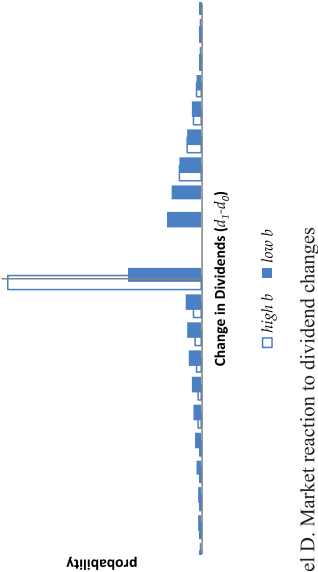
Panel C. Histogram of dividend changes, excluding zero

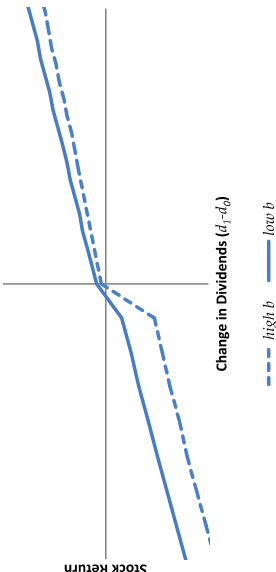Panel D. Market reaction to dividend changes

Figure 2
Equilibrium dividend policy: Weak and strong reference points
We plot the equilibrium in Proposition 1 for parameter values $d_0 = 1$ and $\lambda = 0.5$. We repeat Figure 1, comparing the results for $b=2.1$ and $b=2.3$.

3.1 Surveys

Dividend policy is an explicit choice of executives and the board. Although survey results should always be treated cautiously, the concern that managers' behavior may be guided not by their own hands but by an unseen higher market force that they cannot articulate is harder to justify here. Thus, we view the fact that our model is consistent with what managers say about dividend policy, or at least not directly inconsistent with it, as an important success.

The strongest survey results fit well with our setup. For example, as noted in the Introduction, the Brav et al. (2005) survey of 384 executives revealed strong agreement that shareholders will react negatively to cuts in the dividend, whereas the reward for increases is modest. Executives believe that dividends convey information. As a result, they strive to keep a stable dividend policy. These are straightforward predictions of the model. It is intrinsically dynamic, and the stability of dividends is a central feature. Once a reference point is established, even weak firms will strive to minimize the difference between it and current dividends.

While standard signaling theories also predict that lower dividends are associated with lower market values, executives reject them as based on unrealistic foundations. As noted in the Introduction, executives say that they do not use dividends to show that their firm can withstand the costs and scrutiny associated with raising external capital, or to show that their firm can pass up good projects and still perform well. Only a small minority of executives endorsed signaling through taxes; Brav et al. (2005, 521) summarize taxes as of "second-order importance."

Brav et al. (2005) followed up their survey with in-depth interviews of 23 executives. They noted that "not a single interviewed executive told us that his or her firm had ever thought of increasing payout as a costly means of separating itself from competitors" (522–23). Our model does not explicitly rely on public destruction of value to create a credible signal, so in that sense it would not conflict with this aspect of the evidence.

Finally, standard signaling theories do not naturally focus on dividends per share. It is natural to do so in the reference point context. Standard theories also predict a continuous market reaction. There is nothing special about stability or the historical level of payouts, such that falling short of this level would produce a discontinuous reaction. Dividend policy is defined in more "economic" terms in standard models, such as dividend yield or payout ratio, which are less salient to investors. Dividend policy measured in these units would not make natural reference points, explaining why stability of dividends per share is the usual target.

3.2 The Lintner model

Miller (1986) suggested that the Lintner (1956) model was a behavioral model, as opposed to originating from a traditional maximization problem. Given reasonable parameter values, our model does have dividends

following a partial-adjustment policy, and smoothed relative to earnings (Fama and Babiak 1968).

In particular, the Lintner model takes last period's dividend as the reference level against which current dividends are determined. In our model, firms confident of their ability to sustain high dividends pay out slightly above half of current earnings (exactly half in the case of extreme reference point preferences). They adopt this payout ratio because they do not want to set a reference level that may be too high for themselves next period, but do wish to separate themselves from a pool of firms with intermediate prospects that keep dividends extremely smooth (indeed, constant). On average, then, dividends are increasing in earnings but less than one-for-one, and firms are focused on changes relative to the reference level set by lagged dividends. All of this adds up to Lintner-like empirical predictions.

The models of Bhattacharya (1979), Miller and Rock (1985), and John and Williams (1985) are static and focus on levels, not changes. The model of Allen, Bernardo, and Welch (2000) is also presented in terms of levels, though they outline a possible multiperiod extension that would be compatible with smoothing. The model of Kumar (1988) leads to smoothing to the extent that firm productivity does not vary much over time.

The cross-section of smoothing behavior is also reasonably consistent with our model. Michaely and Roberts (2012) study dividend policy in public versus private firms. The presumption is that private firms are less subject to asymmetric information problems. Thus, there is less expectation that their dividend policy would adhere to the predictions of our signaling model. In particular, Michaely and Roberts find that private firms engage in much less smoothing. In our framework, the interpretation would be that private firms have less need for precautionary savings to be sure to meet the prior period's dividend. If they pay a lower dividend, little is revealed to insiders. Leary and Michaely (2011) find another robust pattern in the cross-section, that high-earnings firms are more likely to smooth dividends. This is exactly what the medium and high "types" in our stylized model will do.

3.3 Repurchases vs. dividends
As Grullon and Michaely (2002) point out, some popular signaling models, including Bhattacharya (1979), Miller and Rock (1985), and the further investigation of Miller and Rock by Guttman, Kadan, and Kandel (2010), imply that dividends and repurchases are perfect substitutes. An enduring puzzle is why repurchases are considerably less regular.

The reference point approach suggests a reason why dividends are better signals: there is only one number to remember. There are, in contrast, at least two parameters of a repurchase program: the purchase price, or price range for a Dutch auction, and the number of shares sought. Of these, only the price is likely to be salient, and clearly a firm cannot commit to repurchasing shares at the same constant price quarter after quarter. Moreover, each investor receives
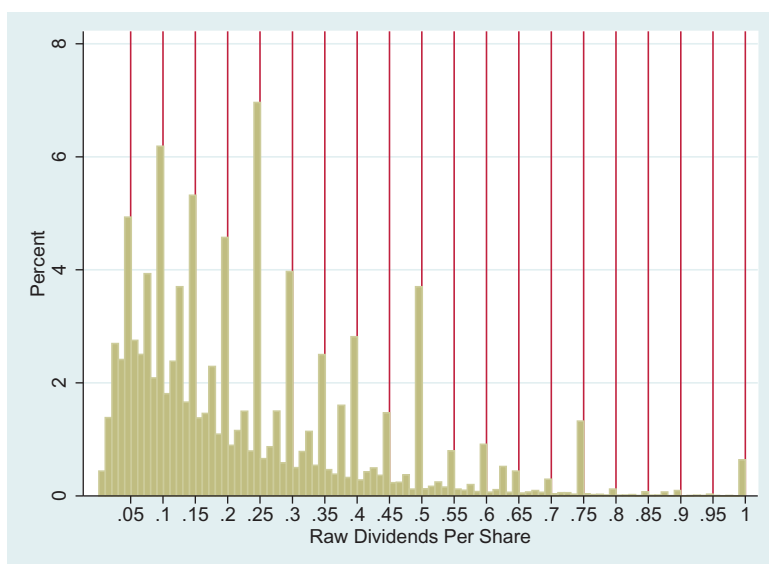
Figure 3
Quarterly dividends per share
Histogram of announced dividends per share. The sample includes all records from the CRSP event file with a distribution code (DISTCD) of 1232 (ordinary taxable cash dividends, paid quarterly) with a share code (SHRCD) of 10 or 11 (ordinary common shares, excluding companies incorporated outside the United States, Americus Trust Components, closed-end funds, and REITs) and non-missing data on the amount of the dividend and the declaration date.

a dividend, but need not engage in a repurchase; only those selling need to pay any attention at all. In short, repurchase programs are hard to specify in salient and memorable terms, even when they are financially equivalent to dividends.

3.4 Salience and round numbers

In the remainder of this section we confirm some additional patterns in U.S. dividends and extend them in directions suggested by our theoretical setup. See Appendix B for a detailed description of our data and variable construction.

A core idea of the model is that the signaling power of dividends is proportional to their use as reference points against which future dividends are judged. We do not model memory, but it is inherent in the concept of a reference point. Firms that signal in our model do not want to hide their dividend; they want to make it salient or memorable. The psychological evidence and introspection indicates that round numbers are easier to remember. In contrast, one-period signaling theories are silent on this because investors need not remember anything.

Consistent with the notion of salient reference points, dividends are indeed paid in round numbers. Aerts, Van Campenhout, and Van Canegham (2008) use Benford's law to show that "0" is an unusually common second digit, but a histogram makes the use of round numbers obvious. Figure 3 shows that

the most common dividend per share is a quarter, followed by a dime, fifteen cents, a nickel, and twenty cents. There are notable spikes at multiples of five cents. Non-round values are rare. Apparently, many boards think of setting and communicating dividend policy in an easily recalled dollar amount, rather than the more "economic" alternative of setting a specific total payout to then divide equally among shareholders. (We make some use of this fact in our empirical specifications below.)

A related pattern is the tendency for increases to reach a round number threshold—the next round number in dividends per share—presumably contributing to the salience of the new level. For example, the next $0.10 threshold for a firm paying $0.11 is $0.20, the next $0.05 threshold is $0.15, and the next $0.025 threshold is $0.125. In unreported results, we find that the modal increase is exactly to the next threshold. For example, for the firm that is currently paying $0.11, we find that paying $0.20 is much more likely than paying $0.19 or $0.21. The same pattern occurs for $0.05 and $0.025 thresholds. These results might also be loosely interpreted in terms of the gap in the distribution in Panel B of Figure 1. To be appreciated for raising the dividend, firms must do so in more than a trivial way. (We note that the gap in the distribution in Panel B of Figure 1 can be made arbitrarily small with appropriate parameter values, and thus be made to look just like the increases plotted in Figure 4.)

3.5 Clustering at zero changes

The pooling behavior in the model is consistent with the great reluctance to change dividends. Figure 4 shows dividend changes when no split occurs between dividend payments. As in Panels A and B of Figure 1, and further summarized in Table 1, there is a very large mass at exactly zero, as also noted by Guttman, Kadan, and Kandel (2010).
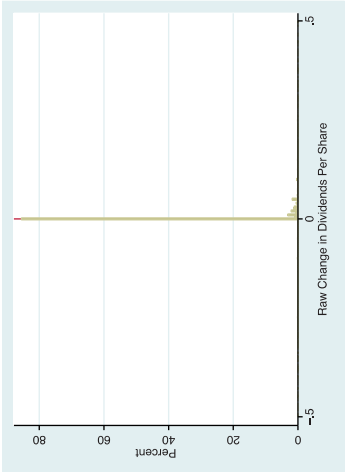
The exact mass at zero changes depends on the assumption about how often firms actually review their dividend policy. We assume a quarterly cycle to be consistent with prior work and because every firm in our sample could change the dividend at any quarter. In practice, however, many firms revisit their payout policy annually. The clustering at zero depicted in the histogram thus overstates the true frequency of actual zero-change decisions.

Complicating matters further, a case study in the subsequent section shows how a firm can change its frequency not just of active review but also of actual payment. Unfortunately, one cannot know the policy for any given firm, let alone all firms in CRSP. Our "streaks" analysis in the next section demonstrates that both quarterly and annual reviews are used in practice. As we discuss there, our quarterly assumption—while leading to some overstatement of clustering at zero changes in the histogram—also leads to a bias "against" the new tests.
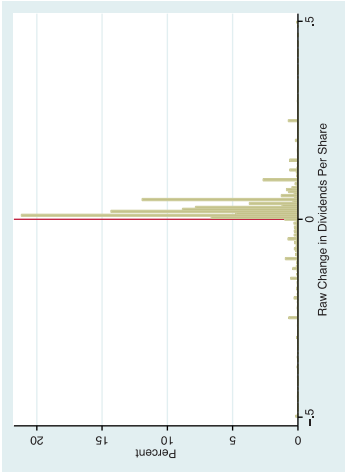
Panel A. Changes in dividends

Panel B. Excluding zero

Panel C. Increases in dividends

Panel D. Decreases in dividends
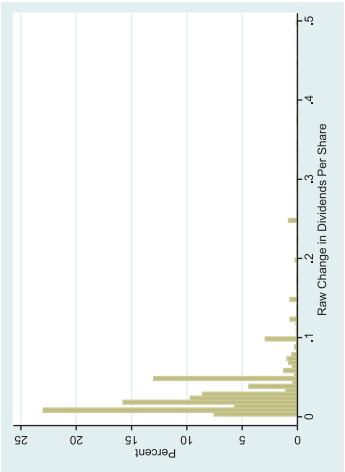
Figure 4
Changes in quarterly dividends per share
Histogram of changes in quarterly dividends per share. Panel A shows the distribution of changes in dividends per share, excluding zero, Panel C shows the distribution of dividends-per-share increases, and Panel D shows the distribution of dividend-per-share decreases. The sample from Table 1 is further limited to changes over a three-month window with no stock splits.

Table 1
Summary statistics

|	N	Mean	Median	SD	5	25	75	95
Dividends per share	391,865	0.245	0.195	0.214	0.038	0.100	0.326	0.625
Change in dividends per share	327,816	0.003	0.000	0.030	0.000	0.000	0.000	0.030
Constant dividend streak	383,809	6.670	4	8.264	1	2	8	23
Three-day announcement return	327,026	0.002	0.000	0.040	−0.053	−0.017	0.019	0.063

The sample includes all records from the CRSP event file with a distribution code (DISTCD) of 1232 (ordinary taxable cash dividends, paid quarterly) with a share code (SHRCD) of 10 or 11 (ordinary common shares, excluding companies incorporated outside the United States, Americus Trust Components, closed-end funds, and REITs) and non-missing data on dividends per share and the declaration date. Change in dividends per share is computed only over three-month windows (with an allowance for the fact that gaps between dividend payments are frequently two or four calendar months) with no stock splits. Constant dividend streak is the number of past periods where dividends per share remained unchanged. Three-day announcement return is computed as the sum in the three days surrounding the dividend declaration date of the difference between the total return and the CRSP value-weighted stock index return.

3.6 Asymmetric announcement returns

Even if executives disavow standard signaling models, shareholders clearly care about dividends—particularly cuts. That dividend cuts would be received especially poorly is a prediction of the model, but not of standard signaling models. The main effect is that cutting a dividend, even slightly, is fully revealing and betrays the firm as one that cannot afford even the reference dividend.

Aharony and Swary (1980) examine cases in which dividend announcements occur separately from earnings announcements. The average cumulative abnormal return in a 21-day window surrounding a dividend decrease was on the order of five percentage points. The average cumulative abnormal return surrounding a dividend increase was closer to one percentage point. Fama, Fisher, Jensen, and Roll (1969), Charest (1978), Eades, Hess, and Kim (1985), Lang and Litzenberger (1989), and Yoon and Starks (1995) find similar conclusions. These results are consistent with an asymmetric announcement effect as predicted by the model.[7]

In our sample, a dividend declaration is met with a 20-basis-point abnormal return in our sample, as reported in Table 1. The median abnormal return is zero. This is a sample of firms that did not omit a dividend, so a slightly positive average is not surprising. More importantly, Figure 5 shows the market reaction to changes in dividends per share. We split the sample into increments of $0.05 in Panel A or $0.025 in Panel B around zero change. We round down to the nearest threshold, so that a dividend increase of $0.01 is included in the zero dividend change group, and a dividend cut of $0.01 is included in the $0.025 cut group. Next, we compute the median three-day abnormal return for each group. The pattern in both panels is similar. Dividend cuts are greeted with

---

[7] Many of these papers simply follow "dividend increases" and "dividend decreases." The extent to which such an exercise can actually document an asymmetry is ambiguous, given that the average decrease is larger than the average increase, at least in dividend per share (DPS) terms. Our specifications involve following "big" and "small" changes in DPS as well.

Panel A. Changes in dividends per share grouped to the nearest 0.05
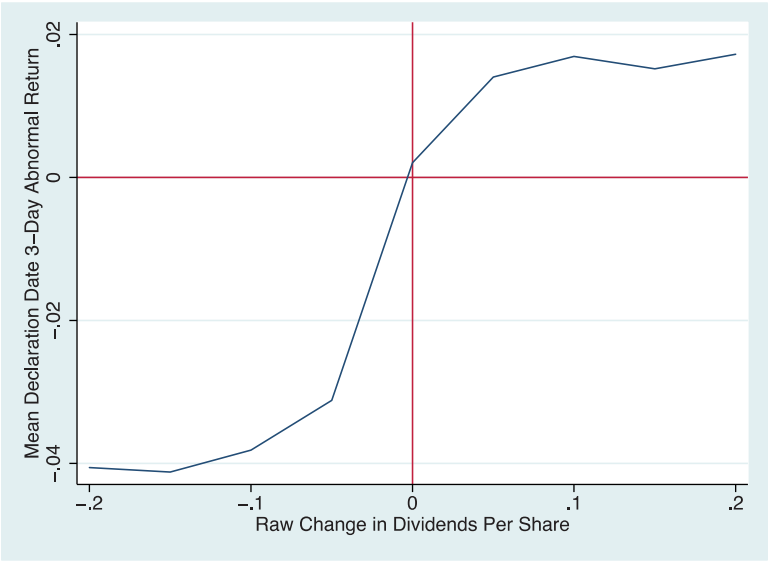
Panel B. Changes in dividends per share grouped to the nearest 0.025
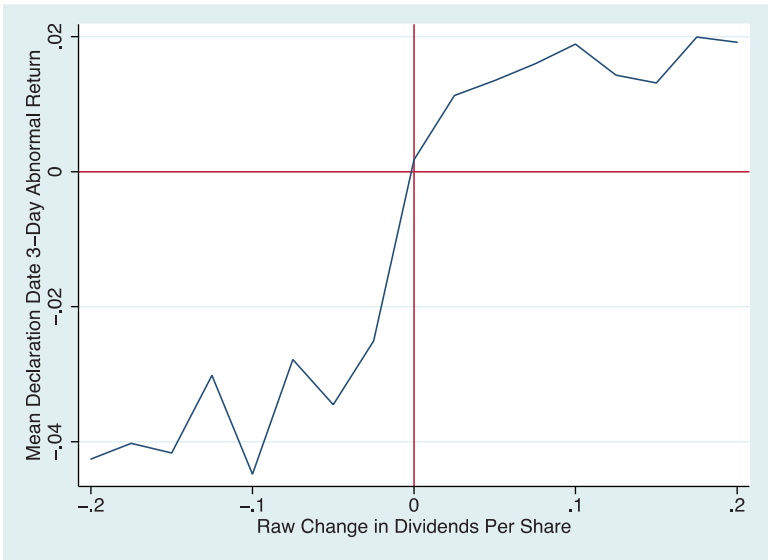
Figure 5
Market reaction to changes in quarterly dividends per share
Average three-day abnormal return by change in quarterly dividends per share. Panel A groups changes in dividends per share into groups of 0.05, while Panel B groups changes in dividends per share into groups of 0.025. The groups are formed by rounding the changes in dividends per share down to the nearest threshold. Three-day abnormal returns are computed as the sum in the three days surrounding the dividend declaration date of the difference between the total return and the CRSP value-weighted stock index return. The sample from Table 1 is further limited to changes over a three-month window with no stock splits.

Table 2
Market reaction to changes in quarterly dividends per share

	1		2		3	
				[T-statistic]		**[T-statistic]**
	Coefficient	**[T-statistic]**	**Coefficient**	**(p-value)**	**Coefficient**	**(p-value)**
b	10.95	[22.38]				
$b_{-\infty,-0.2}$			−4.83	[−3.27]	−4.83	[−3.28]
$b_{-0.2,-0.1}$			−0.93	[−0.21]	−0.79	[−0.18]
$b_{-0.1,-0.05}$			11.03	[1.28]	8.96	[1.12]
$b_{-0.05,-0.025}$			52.13	[2.40]		
$b_{-0.025,0}$			75.77	[4.66]		
$b_{-0.05,0}$					65.93	[11.93]
$b_{0,0.05}$					27.96	[39.80]
$b_{0,0.025}$			35.78	[26.73]		
$b_{0.025,0.05}$			12.77	[5.63]		
$b_{0.05,0.1}$			5.41	[2.46]	0.81	[0.38]
$b_{0.1,0.2}$			−4.43	[−2.02]	−3.94	[−1.79]
$b_{0.2,\infty}$			−0.09	[−0.09]	−0.12	[−0.12]
N	327,026		327,026		327,026	
R^2	0.0066		0.0136		0.0135	
$b_{-0.1,0}-b_{0,0.1}$			84.97	(0.000)	46.12	(0.000)

Piecewise linear regressions of average three-day abnormal return on change in quarterly dividends per share. Three-day abnormal returns are computed as the sum in the three days surrounding the dividend declaration date of the difference between the total return and the CRSP value-weighted stock index return. The sample from Table 1 is further limited to changes over a three-month window with no stock splits. Heteroscedasticity-robust standard errors are reported in brackets.

a larger negative return than dividend increases of the same magnitude. The difference is roughly a factor of two. In fact, the whole response curve is quite similar to the pattern predicted by the model in Panel D of Figures 1 and 2.[8]

We examine the patterns in Figure 5 econometrically in Table 2, where we estimate piecewise linear regressions of the market reaction on the change in dividends per share. We are particularly interested in the shift in slope just below and above zero. The first regression is a simple linear regression. Each $0.01 change in dividends leads to an 11-basis-point reaction. We use robust standard errors in all tests to control for the possibility that, for example, there may be positive correlation among high-beta stocks announcing in the same period.

This obscures a highly nonlinear relationship where changes around zero are much more important than larger movements. The second, piecewise linear regression shows that small cuts in dividends up to $0.025 are greeted with a market reaction of 76 basis points for each $0.01 change. Small increases in dividends up to $0.025 are greeted with a market reaction of 36 basis points, or approximately half the slope that we observe in cuts. There are similarly large differences in the next increments, though the reaction per $0.01 of dividend change drops off quickly. As a summary test, we compare the sum of the three

---

[8] The apparent concavity is not a prediction of our simple model. We focus instead on the effects in the neighborhood of zero. It is worth noting that larger changes in dividends are comparatively rare and may reflect a set of circumstances that are outside of the model, where firms make larger adjustments in regular dividends. They may also reflect firms that simply have higher share prices, and so a $0.10 change in dividends per share is not as newsworthy, inducing a downward bias as we increase dollar dividend changes.

coefficients between –$0.10 and zero to those between zero and +$0.10. The slope for dividend cuts is larger both economically and statistically. We repeat the analysis with a coarser estimation of slopes, combining the slope between cuts or increases of less than $0.025 with those that are between $0.025 and $0.05, with similar conclusions. Finally, in unreported results, we have limited the sample to firm-quarters of actual dividend changes—that is, excluding zero changes. The asymmetries are at least as strong in this case, demonstrating that the misclassification of firms on an annual decision cycle does not affect inferences.

An interesting question is: why focus on dividends per share (DPS) changes as opposed to more "economically relevant" dividend measures? (We noted above that most papers in the announcement effects literature simply analyze "dividend increases" and "dividend decreases," which are one step less informative than our tests.) One reason is simply that this is how managers and investors think about dividend policy, and to be empirically accurate it helps to be behaviorally accurate. Brav et al. (2005) find that managers most commonly think in DPS terms, less commonly in payout ratio or growth terms, and almost never in dividend yield terms. The prevalence of round numbers in DPS—and exhibit clustering in DPS terms—confirms this.

It is still useful to explore the limits of our specification. We have examined restricting the sample to firms in the share price range of $10 and $50, hence cutting off the extreme price deciles. In combination with our winsorization of DPS at $2, this also keeps yields within at least a loose economic bound. After this restriction, the point estimate of the size of the asymmetry diminishes, but the F-test remains highly significant. The asymmetry results remain strong if we include changes in yield, dividend growth, or a dummy variable for a cut. These exercises control for more traditional measures of dividend change and allow the tests to focus on DPS, the most plausible reference point unit.[9]

3.7 Asymmetric announcement volume

In Table 3, we extend the announcement returns analysis by replacing it with abnormal volume as the dependent variable. This offers some additional evidence that dividend cuts get extra attention—the stronger market reaction comes with somewhat stronger volume. Both increases and decreases are associated with higher than normal volume. The negative coefficients below zero and the positive ones above zero in the piecewise linear regressions suggest a V-shaped pattern around zero dividend change. The coefficients are slightly

---

[9] There are some practical issues with measures other than raw DPS as well. Small dividend cuts are rare— firms with stock prices of $100 do not miss their dividend by $0.01, limiting our flexibility in scaling "small" dividend cuts and our power to empirically distinguish a small and a big yield change. Dividend growth, the most theoretically satisfying definition of a change in terms of a straightforward connection to returns via the Gordon growth model, has other empirical problems—for example, it is easy to get growth rates of 500% in moving from an economically tiny dividend to a small one. Changes in payout ratios become noisy as earnings dip to near zero or become negative.

Table 3
Market reaction to changes in quarterly dividends per share: Abnormal volume

	1		*2*		*3*	
	Coefficient	[*T*-statistic]	Coefficient	[*T*-statistic] (*p*-value)	Coefficient	[*T*-statistic] (*p*-value)
b	29.73	[5.45]				
$b_{-\infty,-0.2}$			95.84	[4.93]	95.88	[4.93]
$b_{-0.2,-0.1}$			77.50	[1.52]	76.86	[1.51]
$b_{-0.1,-0.05}$			−374.41	[−3.58]	−365.25	[−3.74]
$b_{-0.05,-0.025}$			−576.20	[−1.94]		
$b_{-0.025,0}$			−686.12	[−2.94]		
$b_{-0.05,0}$					−637.73	[−9.46]
$b_{0,0.05}$					522.41	[33.72]
$b_{0,0.025}$			605.63	[20.10]		
$b_{0.025,0.05}$			360.93	[7.18]		
$b_{0.05,0.1}$			79.66	[1.89]	30.74	[0.77]
$b_{0.1,0.2}$			−63.96	[−1.55]	−58.70	[−1.43]
$b_{0.2,\infty}$			−7.90	[−0.50]	−8.19	[−0.52]
N	288,740		288,740		288,740	
R^2	0.0001		0.0069		0.0069	
$b_{-0.1,0}-b_{0,0.1}$			590.51	(0.000)	449.84	(0.000)

Piecewise linear regressions of average three-day abnormal volume on change in quarterly dividends per share. Three-day abnormal volume is computed as the log difference between the average daily volume in the three days surrounding the dividend declaration date and the average daily volume in the 90 calendar days preceding the announcement. The sample from Table 1 is further limited to changes over a three-month window with no stock splits. Heteroscedasticity-robust standard errors are reported in brackets.

larger in absolute value for cuts than increases, however. In the range from zero to a cut of $0.025, every $0.01 cut in dividends is associated with an increase in volume of 686 basis points, or almost 7% more than normal volume. Similar dividend increases are also associated with higher volume, but the rate is somewhat smaller at 606 basis points. The difference is not large in economic terms, compared with the difference in returns, but the joint test of the differential sensitivity of volume to dividend changes above and below zero is statistically strong. Also, Figure 4 showed that dividend decreases are larger, on average, bringing the total increase in volume to a more economically significant level.

4. New Evidence

In addition to explaining known stylized facts at least as well as standard signaling models, whether taken individually or collectively, the behavioral signaling approach suggests a testable comparative static. The core idea is that the power of dividends as a signal is proportional to their use as reference points—in the model, this is parameter b. Splitting the sample on b should illustrate cross-sectional differences in announcement effect asymmetries, the degree of clustering, and so on.

How to measure the intensity of a dividend reference point for any given firm-quarter? Several proxies come to mind, including firm size, number of analysts, or retail ownership, for example. These proxies, unfortunately, invite type II error, because they pick up both b and, simultaneously, information
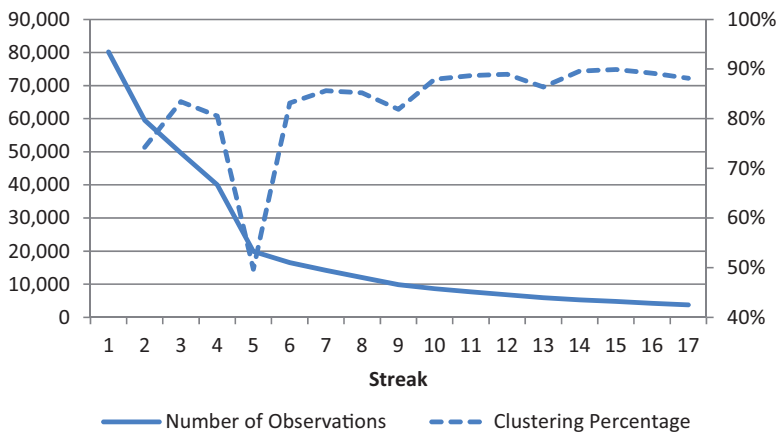
Figure 6
Clustering following quarterly dividend-per-share streaks
The number of observations and the clustering percentage for streaks of up to four years. We divide the sample of dividend announcements into those where the dividend changed from the previous period, a streak of 0, where the dividend remained the same after being changed the previous period, a streak of 1, and so on. The clustering percentage at a streak of 1 is the percentage of firms that do not change their dividend after having changed their dividend in the previous period. The clustering percentage at a streak of 2 is the percentage of firms that do not change their dividend after having maintained their dividend at the same level for the previous two periods. And so on.

asymmetry. Consider a high-analyst firm, for example. It will indeed attract more attention and scrutiny of dividends. However, this also means that firms with these characteristics have less to signal in the first place. There is less information asymmetry over ε_1 precisely because of the attention. We could at best determine whether the reference point effect dominates the need to signal, not whether the reference point effect exists. A proxy for b thus needs to be chosen carefully.

4.1 A direct test: Dividend streaks
Our proxy for b involves dividend streaks. If memory is an important part of reference point formation, then repeated dividends of the same amount would presumably be stronger reference points. The basic idea is that long streaks constitute stronger reference points, so there should be more clustering and the patterns in announcement effects should be more pronounced as the streak lengthens. Importantly, the repetition of a given dividend does not have any obvious correlation with investor attention or a fundamental degree of uncertainty.

We plot clustering around successive streaks in Figure 6, examining streaks of length 1 through 16 separately. A large number of firms reevaluate their dividend policy annually, so there is a drop in clustering after a streak of 4, 8, 12, and 16. Controlling for this seasonality shows a strict monotonicity, however; an increase is apparent in any year-over-year comparison—for example, in
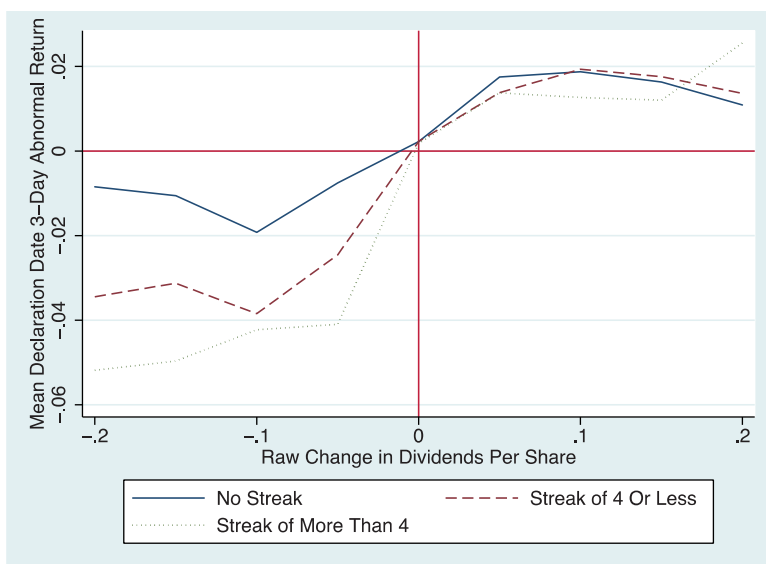
Figure 7
Market reaction following quarterly dividend-per-share streaks
Average three-day abnormal return by change in quarterly dividends per share. Changes in dividends per share are sorted into groups of 0.05, by rounding the raw dividend per share down to the nearest threshold. We partition the sample into situations where the dividend was changed in the previous period (no streak), where the dividend was not changed in the previous period, but it was changed within the last four periods (streak $\leqslant 4$), and where the dividend was not changed in the previous four periods (streak > 4). Three-day abnormal returns are computed as the sum in the three days surrounding the dividend declaration date of the difference between the total return and the CRSP value-weighted stock index return. The sample from Table 1 is further limited to changes over a three-month window with no stock splits.

the continuance probabilities of 2, 6, 10, and 14 quarters, or of 3, 7, 11, and 15 quarters. Long streaks are also likelier to continue. After a streak lasting a few years, the probability of continuing from the past quarter nears 90%. These patterns are consistent with the prediction that clustering becomes more pronounced as the reference point strengthens.

We next consider the market reaction. We partition the sample into three categories: decisions following a change in the prior quarter; decisions following no change for up to four quarters, to negate the periodicity apparent in reviews of dividend policy; and dividend decisions following no change for more than four quarters. The results are in Figure 7 and Table 4. The results are as predicted. The no-streak sample has essentially no difference between the effect of a decrease and increase around zero, while the patterns are stronger as the streaks get longer.

For example, in the long-streak sample, the market reaction to dividend cuts is stronger than gains, and also stronger than the unconditional coefficients in Table 2 at 106 basis points per $0.01 change in dividends just below zero, versus 76 basis points in Table 2. The market reaction to a dividend increase

Table 4
Market reaction following quarterly dividend-per-share streaks

	No streak		Streak ⩽4		Streak >4	
	Coefficient	[*T*-statistic]	Coefficient	[Diff. *T*-statistic]	Coefficient	[Diff. *T*-statistic]
$b_{-\infty,-0.2}$	−7.81	[−4.40]	−3.82	[1.79]	−1.68	[3.40]
$b_{-0.2,-0.1}$	18.32	[2.26]	−0.32	[−1.90]	−7.22	[−3.14]
$b_{-0.1,-0.05}$	−1.62	[−0.08]	24.03	[0.97]	1.27	[0.12]
$b_{-0.05,-0.025}$	4.03	[0.09]	9.45	[0.10]	73.21	[1.24]
$b_{-0.025,0}$	31.96	[1.22]	79.37	[1.29]	106.15	[1.86]
$b_{0,0.025}$	44.15	[8.51]	29.48	[−2.70]	54.03	[1.65]
$b_{0.025,0.05}$	12.43	[1.32]	15.25	[0.29]	2.30	[−0.97]
$b_{0.05,0.1}$	3.33	[0.31]	11.29	[0.71]	−5.73	[−0.79]
$b_{0.1,0.2}$	−4.09	[−0.51]	−5.61	[−0.18]	−1.94	[0.25]
$b_{0.2,\infty}$	0.72	[0.31]	−1.86	[−1.00]	1.41	[0.23]
N					327,026	
R^2					0.0147	
$b_{-0.1,0}-b_{0,0.1}$	−25.54	(0.206)				
$b_{-0.1,0}-b_{0,0.1}$						
No streak $b_{-0.1,0}-b_{0,0.1}$			87.90	(0.002)	161.09	(0.000)

Average three-day abnormal return by change in quarterly dividends per share. We partition the sample into situations where the dividend was changed in the previous period (no streak), where the dividend was not changed in the previous period, but it was changed within the last four periods (streak ⩽ 4), and where the dividend was not changed in the previous four periods (streak > 4). *T*−statistics and *p*-values are on differences from the no-streak sample. Three-day abnormal returns are computed as the sum in the three days surrounding the dividend declaration date of the difference between the total return and the CRSP value-weighted stock index return. The sample from Table 1 is further limited to changes over a three-month window with no stock splits. Heteroscedasticity-robust standard errors are reported in brackets.

is also larger at 54 versus 36 basis points in Table 2, despite being half of the reaction to a dividend cut in Table 4. These results support the prediction that stronger reference points generate stronger market reactions to changes.[10]

4.2 A placebo test: ADRs
The dividends streaks test is our main new test of the model. We supplement this direct test, where we find an effect where we expect it, with a placebo test, which shows that we do not find an effect where we do not expect it.

Our placebo test involves dividends on American Depository Receipts (ADRs). An ADR allows U.S. investors to purchase shares in a company that is incorporated abroad and listed on a foreign exchange, but without executing a transaction on a foreign exchange in a different currency. Because of foreign exchange volatility, the dividend policy of a firm with a U.S. ADR is by definition unable to create a reference point in two different currencies at once. What this means is that the anchor of past dividends can be relevant only in one currency, not both. As a result, we expect to see clustering in local currency terms, but not in dollars. Moreover, we expect to see no market reaction to a dividend "miss" in dollar terms. This might seem somewhat trivial, but it

---

[10] Despite their strength, the effects here are diluted by our assumption that all firms evaluate their dividend quarterly. For an "annual" firm that just changed its dividend, we place it in the second category for the next three quarters even though it belongs more in the first category (no streak), and so on.

reveals that investors are focused on a nominal reference point, rather than a specific economic outcome measured in real terms.

The case of BP-Amoco shown in Figure 8 provides a fascinating demonstration of how the reference point is set to appeal to the relevant investor base. In December of 1998, British Petroleum acquired Amoco to form BP-Amoco. BP was listed on the London Stock Exchange but also traded through an ADR. Panel A shows that prior to the merger, Amoco had increased dividends by $0.025 each year for the prior four years. BP had increased dividends by £0.0125 semiannually for the previous two years. Not surprisingly, the dollar dividend on the ADR was hardly so regular.

The merger required some reconciliation between these two different but equally rigid policies. The reconciliation was for BP to now fix dividend increases in dollar terms. Moreover, for the several years following the merger, the rate of increase in BP dividends exactly matched Amoco's old rate of increase, amounting to $0.025 each year. The common British policy of semiannual payment, however, was kept. Dividend policy in the transition was thus managed carefully so as not to upset dollar-dividend reference points that had been created for Amoco shareholders over years, presumably since they now owned a large fraction of BP shares.

Turning toward a large sample test, the suggestion from the BP-Amoco case is that zero change in dollar dividends has no special significance for ADRs. We use an ADR sample that is described in Appendix C. In other words, investors do not care about dividend cuts per se, but about a cut from a mutually agreed-upon reference point. Because reference points cannot hold simultaneously in two currencies, ADR dividends can in most cases freely cross the zero-change boundary. We find that the market reaction is similarly unremarkable in this range.

Figure 9 shows the dividend policy of ADRs measured in both dollars and local currency. Dividend changes in U.S. dollars are centered on zero change, but the mass point at zero in Panel A is very far from what we saw in Table 1 for the CRSP sample. Moreover, the asymmetry between dividend cuts and increases is barely apparent in Panel B when we eliminate zeros from the sample. By contrast, when measured in local currency, there is a much clearer delineation at zero. Non-zero dividend changes are comparatively rare in Panel C, and when we exclude zero changes in Panel D, a preference for small increases over decreases is readily apparent. It is noteworthy that these effects are less pronounced in the parents-of-ADRs sample than in the CRSP sample of Figure 4. Part of this is because we broadened the sample as much as possible (after imposing various data screens), perhaps at the cost of including some special or liquidating dividends. Part of this may be because the Datastream data is lower quality, inducing an attenuation effect. In any event, it is plain that dividend policy is more often decided in local currency than dollar terms.

We examine the market reaction to these changes in Table 6. The first observation is that the relationship between dividend changes and the market
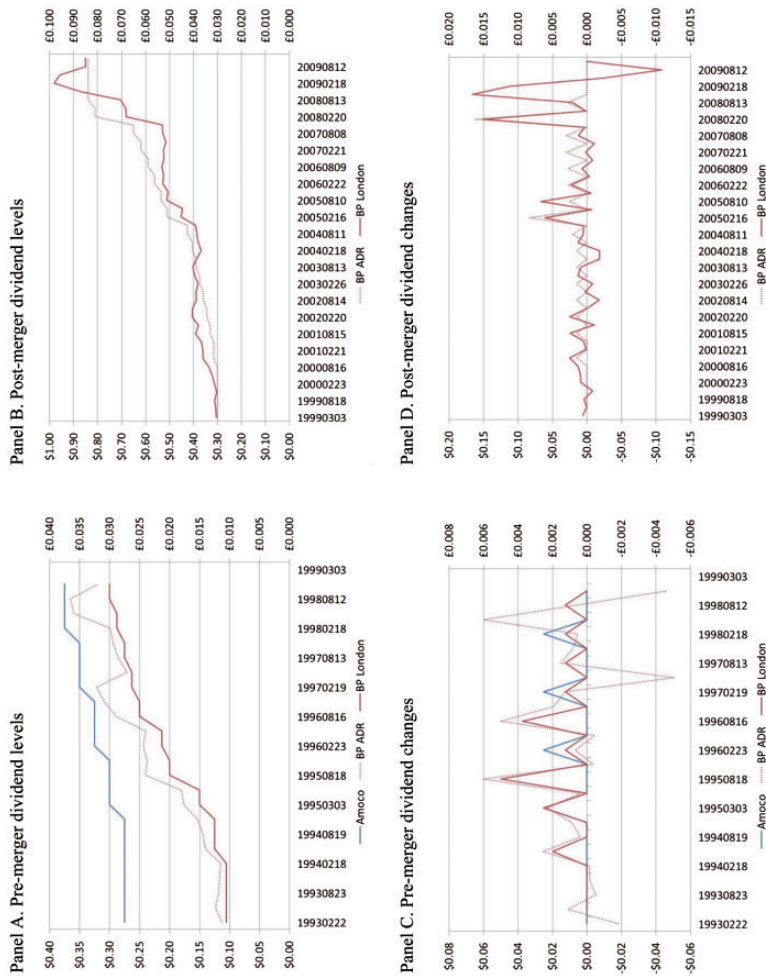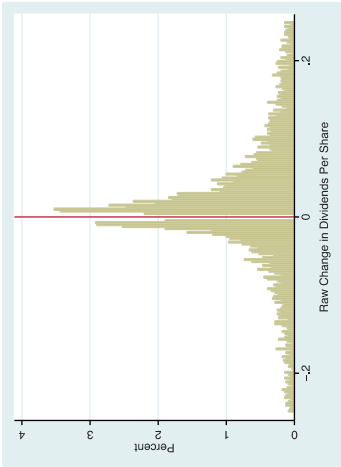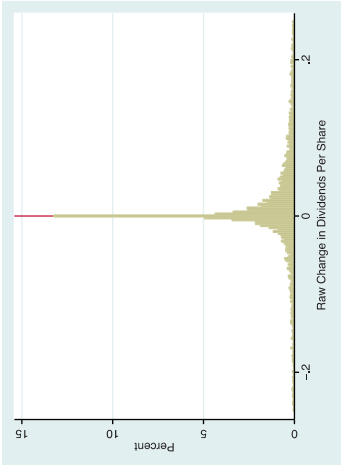
Figure 8
BP-Amoco dividend policy
Split-adjusted dividends per share for BP, Amoco, and the merged company. BP and Amoco merged in December of 1998 forming BP-Amoco. Panels A and C show dividend levels and changes prior to the merger. Panels B and D show dividends after the merger.

Panel A. ADRs, changes in dividends per share

Panel B. ADRs, changes in dividends per share, excluding zero

Panel C. Parents, changes in dividends per share

Panel D. Parents, changes in dividends per share, excluding zero

Figure 9
Changes in quarterly dividends for ADRs and parents
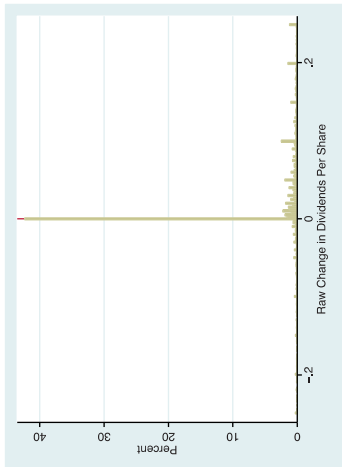Histogram of changes in quarterly dividends per share. Simple changes in consecutive dividends per share (DD) from Datastream for the following Datastream types: QTR, HYR, YR, FIN, INT. Panels A and C show the distribution of changes in dividends per share, and Panels B and D show the distribution excluding zero. Panels A and B show ADRs, and Panels C and D show parent companies.

Table 5
Summary statistics

	N	Mean	Median	SD	5	25	75	95
Panel A. U.S. ADRs								
Dividends per share, split adjusted	10,634	0.438	0.210	2.504	0.018	0.080	0.471	1.354
Change in dividends per share	8,478	0.022	0.002	1.108	−0.228	−0.012	0.044	0.301
Five-day announcement return	3,574	0.006	0.000	0.060	−0.088	−0.023	0.031	0.116
Panel B. Parents								
Dividends per share, split adjusted	31,828	42.307	1.199	878.1	0.0	0.2	5.5	34.0
Change in dividends per share	8,968	−5.042	0.005	1973.1	−2.1	0.0	0.2	5.0
Change in dividends per share, all	28,877	0.094	0.000	1144.0	−4.4	0.0	0.2	5.7

The sample includes all ADRs from Datastream and their matched parents with non-missing data on dividends per share and without an apparent practice of paying different levels of annual versus interim dividends. The sample includes all ADRs from Datastream with a trading history between 1990 and 2009 on a U.S. exchange in U.S.$. We limit the sample to the following Datastream dividend types (DT): QTR, HYR, YR, FIN, INT. We compute clean changes that require the consecutive dividends to be of the same type (DT). These data are merged onto the Datastream returns data. For ADRs, we compute five-day returns surrounding announcement dates (DECQ1, DECQ2, DECQ3, or DECQ4 from Worldscope via Datastream) where we have a clean change in dividends per share. Returns are winsorized at the 1st and 99th percentiles. We look back up to three months to find a matching announcement date, and we use announcement dates from ADR-parent pairs to enlarge the sample. The returns use the Datastream return index (RI) up to five weekdays surrounding the announcement date, with no adjustment for market movements.

Table 6
Market reaction to changes in quarterly dividends per share for ADRs

	1		*2*		*3*	
	Coefficient	[*T*-statistic]	Coefficient	[*T*-statistic] (*p*-value)	Coefficient	[*T*-statistic] (*p*-value)
b	0.21	[1.20]				
$b_{-\infty,-0.2}$			0.32	[1.27]	−0.33	[−1.27]
$b_{-0.2,-0.1}$			−3.26	[−0.39]	−3.53	[−0.43]
$b_{-0.1,-0.05}$			−6.33	[−0.31]	−3.92	[0.22]
$b_{-0.05,-0.025}$			20.68	[0.62]		
$b_{-0.025,0}$			7.25	[0.33]		
$b_{-0.05,0}$					12.43	[1.12]
$b_{0,0.05}$					15.01	[1.84]
$b_{0,0.025}$			15.76	[0.89]		
$b_{0.025,0.05}$			15.30	[0.66]		
$b_{0.05,0.1}$			−24.39	[−1.86]	−24.22	[−2.12]
$b_{0.1,0.2}$			7.94	[1.44]	7.92	[1.46]
$b_{0.2,\infty}$			−0.00	[0.06]	0.00	[0.06]
N		3,574		3,574		3,574
R^2		0.0006		0.0034		0.0033
$b_{-0.1,0}-b_{0,0.1}$			14.93	(0.54)	17.72	(0.28)

Piecewise linear regressions of average five-day announcement return on clean changes in quarterly dividends per share. The sample is described in Panel A of Table 5. We compute five-day returns surrounding announcement dates (DECQ1, DECQ2, DECQ3, or DECQ4 from Worldscope via Datastream) where we have a clean change in dividends per share. We look back up to three months to find a matching announcement date, and we use announcement dates from ADR-parent pairs. The returns are based on the Datastream return index (RI) up to five weekdays surrounding the announcement date. Heteroscedasticity-robust standard errors are reported in brackets.

reaction is everywhere less economically and statistically significant. The R^2 drops from 0.0136 to 0.0034. Moreover, there is no asymmetry in the neighborhood of zero. The marginal reaction to small cuts is on average about the same as the reaction to small increases. Together, these results suggest that

neither ADR boards nor investors view past dividends—paid in dollars—as an important reference point. Again, the corollary from this placebo test is that changes in dividend policy are important because of an endogenously chosen and acknowledged reference point, not because changes in this neighborhood would otherwise have been economically important.

5. Conclusion

Standard dividend signaling theories posit that executives use dividends to destroy some firm value and thereby signal that plenty remains. The money burning typically takes the form of tax-inefficient distributions, forgone profitable investment, or costly external finance. The executives who actually set dividend policy overwhelmingly reject these ideas—yet, at the same time, are equally adamant that "dividends are a signal" to shareholders and that cutting them has negative consequences.

We develop what we believe to be a more realistic signaling approach. We use loss aversion, which is widely supported in both the experimental and empirical literatures, to create a model in which past dividends are reference points against which future dividends are judged. The theory is consistent with important aspects of the data—including survey evidence, patterns of market reaction to dividend announcements, dividend smoothing, and the Lintner partial-adjustment model—and the preference for regular dividends over regular repurchases. We also find support for its broader intuition that dividends are paid in ways that make them memorable and thus serve as stronger reference points and signals.

Appendix A. Proofs

Proposition 1. Dividends are weakly monotonic in profits. The equilibrium level of dividends $d_1 = d_1(\varepsilon_1)$ is a non-decreasing function of ε_1.

Proof of Proposition 1. Consider some pure-strategy equilibrium in which a manager with period-1 earnings x pays dividends $d_1 = d(x)$. Suppose that the equilibrium function $d(x)$ is not weakly increasing. Then there exists some $x < y$ with $d(x) > d(y)$. For this to be an equilibrium, it must be that the y-type manager does not have an incentive to deviate to paying $d(x)$, while the x-type manager does not have an incentive to deviate to paying $d(y)$. Note that both of these deviations are feasible within the budget constraints of these two managers. These conditions can be written:

$$\lambda F(d(x), d_0) + (1-\lambda) V_0(d(x), d_0; x) \geqslant \lambda F(d(y), d_0) + (1-\lambda) V_0(d(y), d_0; x)$$

$$\lambda F(d(y), d_0) + (1-\lambda) V_0(d(y), d_0; y) \geqslant \lambda F(d(x), d_0) + (1-\lambda) V_0(d(x), d_0; y)$$

Isolating the λ terms to the left side of the first equation and the right side of the second equation, these become:

$$\lambda F(d(x), d_0) - \lambda F(d(y), d_0) \geqslant (1-\lambda) V_0(d(y), d_0; x) - (1-\lambda) V_0(d(x), d_0; x)$$

$$(1-\lambda) V_0(d(y), d_0; y) - (1-\lambda) V_0(d(x), d_0; y) \geqslant \lambda F(d(x), d_0) - \lambda F(d(y), d_0)$$

Recall from above that V_0 is supermodular. That is, higher-type managers have a "stronger" V_0-preference for high dividends than do lower-type managers. Because $x < y$ and $d(x) > d(y)$, this implies that:

$$V_0(d(x), d_0; y) - V_0(d(y), d_0; y) \geqslant V_0(d(x), d_0; x) - V_0(d(y), d_0; x)$$

Multiplying by -1 and $(1 - \lambda)$, this allows us to link the last two inequalities to give:

$$\lambda F(d(x), d_0) - \lambda F(d(y), d_0) \geqslant \lambda F(d(x), d_0) - \lambda F(d(y), d_0)$$

Dividing by λ and adding the negative terms to each side yields:

$$F(d(x), d_0) + F(d(y), d_0) \geqslant F(d(x), d_0) + \lambda F(d(y), d_0)$$

Of course, this inequality must in fact be an equality. Working backward, this implies that each of these inequalities must have been equalities all along. The x-type manager and y-type manager are each indifferent between paying $d(x)$ or $d(y)$. This can only be sustained if the investors (accurately) believe that the firm with the lower dividends in fact has the better earnings. This is an unstable sunspot equilibrium, which we rule out by requiring whenever a manager is indifferent between paying $d(x)$ or $d(y)$, he pays the higher of the two. This suffices to prove the proposition.

Derivation of x_1. Having established the existence of a solution with the desired functional form, we solve for x_1 by imposing the equilibrium conditions on the proposed solution. The derivation of x_1 can be broken into three cases: $x_1 < 2\alpha$, $x_1 = 2\alpha$, and $x_1 > 2\alpha$. The case where $x_1 < 2\alpha$ is infeasible, as it implies $x_1 < \frac{x_1}{2} + \alpha$, so a manager with earnings $\epsilon_1 = x_1$ is unable to pay the dividend $\frac{x_1}{2} + \alpha$. This case can then be ignored. In the equality case, x_1 is immediately pinned down by assumption.

In the final case with $x_1 > 2\alpha$, we have $\frac{x_1}{2} + \alpha < x_1$. Then a manager with earnings $\varepsilon_1 < x_1$ by a sufficiently small amount has the ability to pay dividends $d_1 = \frac{x_1}{2} + \alpha$, imitating the higher-type manager. Of course, the higher-type manager has the same option of imitating the lower-type manager. In this unconstrained case, the manager with $\varepsilon_1 = x_1$ must be indifferent between paying the pooling level of dividends or the separating level of dividends.

The indifference condition gives:

$$\lambda \frac{d_0 + x_1}{2} + (1 - \lambda) \left(1 + x_1 + (1 - \beta) d_0 - b\beta \left(d_0 - \frac{x_1}{2} \right)^2 1_{d_0 > \frac{x_1}{2}} \right)$$

$$= \lambda x_1 + (1 - \lambda)(1 + x_1 + (1 - \beta) \left(\frac{x_1}{2} + \alpha \right) - b\beta \left(\frac{x_1}{2} + \alpha - \frac{x_1}{2} \right)^2 1_{\frac{x_1}{2} + \alpha > \frac{x_1}{2}})$$

This simplifies to:

$$\frac{\lambda}{1 - \lambda} \frac{d_0}{2} + (1 - \beta) d_0 - b\beta \left(d_0 - \frac{x_1}{2} \right)^2 1_{d_0 > \frac{x_1}{2}} = \frac{\lambda}{1 - \lambda} \frac{x_1}{2} + ((1 - \beta) \left(\frac{x_1}{2} + \alpha \right) - b\beta\alpha^2)$$

In the case where $d_0 \leqslant \frac{x_1}{2}$, this becomes:

$$x_1 = \frac{\frac{\lambda}{1 - \lambda} d_0 + (1 - \beta) 2(d_0 - \alpha) + 2b\beta\alpha^2}{\frac{\lambda}{1 - \lambda} + (1 - \beta)}$$

as in the text. In the remaining case where $d_0 > \frac{x_1}{2}$, the indifference condition becomes:

$$0 = b\beta (d_0 - \frac{x_1}{2})^2 + \left(\frac{\lambda}{1 - \lambda} + 1 - \beta \right) \frac{x_1}{2} - ((1 - \beta)(d_0 - \alpha) + b\beta\alpha^2 + \frac{\lambda}{1 - \lambda} \frac{d_0}{2})$$

as in the body of the text.

Proposition 4. In the equilibrium described in Proposition 3, x_1 and the market reaction to $d_1 < d_0$ are increasing in b.

Proof of Proposition 4. x_1 is defined in Proposition 3 as satisfying one of the following two equations:

$$b\beta\left(d_0 - \frac{x_1}{2}\right)^2 + \left(1 - \beta + \frac{\lambda}{1-\lambda}\right)\frac{x_1}{2} - \left((1-\beta)(d_0 - \alpha) + b\beta\alpha^2 + \frac{\lambda}{1-\lambda}\frac{d_0}{2}\right) = 0 \text{ when } x_1 < 2d_0, \text{ or}$$

$$x_1 = \frac{\frac{\lambda}{1-\lambda}d_0 + 2(1-\beta)(d_0 - \alpha) + 2b\beta\alpha^2}{\frac{\lambda}{1-\lambda} + (1-\beta)} \text{ when } x_1 \geqslant 2d_0.$$

Consider the case where $x_1 < 2d_0$ and a comparative static in b. Note first that $\frac{\partial \alpha}{\partial b} = -\frac{\alpha}{b}$. Then:

$$\frac{\partial x_1}{\partial b}2b\beta\left(d_0 - \frac{x_1}{2}\right) + \beta\left(d_0 - \frac{x_1}{2}\right)^2 + \left(1 - \beta + \frac{\lambda}{1-\lambda}\right)\frac{1}{2}\frac{\partial x_1}{\partial b} - (1-\beta)\frac{\alpha}{b} - \beta\alpha^2 + 2b\beta\frac{\alpha}{b} = 0$$

Rearranging:

$$\frac{\partial x_1}{\partial b}\left(2b\beta\left(d_0 - \frac{x_1}{2}\right) + \frac{1}{2}\left(1 - \beta + \frac{\lambda}{1-\lambda}\right)\right) = \beta\alpha^2 - 2\beta\alpha + (1-\beta)\frac{\alpha}{b} - \beta\left(d_0 - \frac{x_1}{2}\right)^2$$

In the case where $x_1 < 2d_0$ the coefficient on $\frac{\partial x_1}{\partial b}$ on the left is positive. The right-hand side is positive if and only if:

$$\beta\left(d_0 - \frac{x_1}{2}\right)^2 \leqslant \frac{\alpha}{b}(b\beta\alpha - 2b\beta + (1-\beta))$$

Solving this out completely is difficult, but we can state necessary conditions by requiring that the right-hand side of this inequality be positive. As α, b are each positive, this condition becomes:

$$0 \leqslant b\beta(\alpha - 2) + 1 - \beta$$

By choice of λ, α may be made arbitrarily large, so this can be satisfied.

In the simpler case where $x_1 \geqslant 2d_0$, the comparative static is:

$$\frac{\partial x_1}{\partial b} = \frac{2(1-\beta)\left(\frac{\alpha}{b}\right) + 2\beta\alpha^2 + 4\beta\alpha}{\frac{\lambda}{1-\lambda} + (1-\beta)} = \frac{2\alpha}{\frac{\lambda}{1-\lambda} + (1-\beta)}\left((1-\beta)b^{-1} + \beta\alpha + 2\beta\right)$$

which is positive, as desired.

Appendix B. U.S. Dividends and Returns Data

The primary sample of U.S. dividends data in Table 1 is from the Center for Research on Security Prices (CRSP) database. We start with all records in the event database with a distribution code (CRSP: DISTCD) equal to 1232. These are ordinary taxable dividends paid at a quarterly frequency. We further limit the sample to firms with a share code (CRSP: SHRCD) of 10 or 11. This restricts our attention to ordinary common shares and eliminates most companies incorporated outside the United States, Americus Trust Components, closed-end funds, and REITs. Such firms have dividend policies that may have reference points denominated in non-dollar currencies or have regulatory or contractual restrictions on dividend policy. We further eliminate dividend payments of $0, dividend payments greater than $2 per share (these are rare, and we wish to avoid skewing computations of nominal changes), and dividends for which there is no declaration date (CRSP: DCLRDT). The CRSP data start in 1926, but restricting attention to more recent periods does not change economic or statistical conclusions.

Note that by eliminating $0 dividend payments, we eliminate dividend omissions and initiations. Initiations are somewhat removed from deviating from a pool of similar firms. Omissions can

be driven by qualitatively different considerations, and we want to avoid any asymmetry in announcement effects to be driven simply by omissions.

Our stock returns and volume data are also from CRSP. To measure the reaction to a dividend announcement, we compute a three-day abnormal stock return around the declaration date. This is the simple return (CRSP: RET) for the firm in the day before, the day of, and the day after a dividend declaration minus the return of the CRSP value-weighted index over the same window. We also measure volume from the dividend declaration through three days after. We normalize this volume by taking the log difference between the average daily declaration date volume and the average daily volume in the previous 90 calendar days.

Appendix C. ADR and Parents Dividends and Returns Data

The ADRs and matched parents described in Table 5 are from Datastream and cover the period from 1990 through 2009. We restrict the sample to firms with an ADR traded on the NYSE, the Nasdaq, and other U.S. over-the-counter exchanges. This gives us a preliminary list of 4,916 Datastream codes for ADRs and their parents. Despite this large initial number of potential firms, the coverage and quality of Datastream dividend data is much lower than CRSP. Some of the parents appear more than once, meaning that there is more than one ADR for a given parent firm. We treat these as separate observations.

We gather information on dividends paid per share (Datastream: DD) in each month for these Datastream codes. We restrict attention to the following dividend types (Datastream: DT): QTR, HYR, YR, FIN, INT. Semiannual and annual dividends are more common abroad, so we include them. We also include dividends designated as final and intermediate under the assumption that many of these are regular dividends during the course of a fiscal year. We exclude a small number of observations where an ADR pays a dividend in a foreign currency, despite apparently trading on a U.S. exchange, or the parent pays a dividend in U.S. dollars. These are likely data errors. We are able to find 19,046 dividends for ADRs and 32,177 dividends for their parents. Given the smaller quantity of data, we use split-adjusted values, so we can examine changes in more cases.

Our interest is whether or not a reference point is created through the payment of ADR dividends. When we compute changes, we require that the dividend type be constant from one period to the next. Quarterly dividends are reported to be more common in the ADRs in Datastream than in their matched parents, for reasons that are not clear, so we lose more data when we look for clean changes in the parent sample. The dividend type typically stays the same in consecutive records for ADRs, while the dividend type is the same in only 9,196 of 29,211 consecutive parent records.

The dividends per share for the parents are paid in a wide range of currencies, so the levels of dividends per share range widely. There are many small dividend payments in more valuable currencies and many large ones in less valuable currencies. There is no unambiguous way to put all of these currencies on level terms without losing the essence of a reference point analysis, so we leave them in raw terms.

We compute announcement returns for the ADR sample by merging declaration dates from Worldscope (Datastream: DECQ1–DECQ4) to Datastream return indexes (Datastream: RI) for the five-day window surrounding the declaration date to accommodate the lower quality of Worldscope's declaration dates. Dividend payments are matched to declaration dates that occur for up to three previous months to increase coverage.

References

Abel, A. B. 1990. Asset prices under habit formation and catching up with the Joneses. *American Economic Review* 80:38–42.

Aerts, W., G. Van Campenhout, and T. Van Caneghem. 2008. Clustering in dividends: Do managers rely on cognitive reference points? *Journal of Economic Psychology* 29:276–84.

Aharony, J., and I. Swary. 1980. Quarterly dividend and earnings announcements and stockholders' returns: An empirical analysis. *Journal of Finance* 35:1–12.

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

Arkes, H., D. Hirshleifer, D. Jiang, and S. S. Lim. 2008. Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decisions Processes* 105:67–81.

Atkinson, R. C., and R. M. Shiffrin. 1968. Human memory: A proposed system and its control processes. In *The Psychology of learning and motivation,* vol. 2, 89–195. Ed. K. W. Spence and J. T. Spence. New York: Academic Press.

Baker, M. P., S. Nagel, and J. Wurgler. 2007. The effect of dividends on consumption. *Brookings Papers on Economic Activity* 38(1):231–92.

Baker, M., X. Pan, and J. Wurgler. 2012. The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics* 106:49–71.

Baker, M. P., and J. Wurgler. 2012. Behavioral corporate finance: An updated survey. In *The Handbook of the Economics of Finance,* vol. 2A, 357–424. Ed. G. M. Constantinides, M. Harris, and R. M. Stulz. New York: Elsevier.

Baker, M., and Y. Xuan. 2015. Under new management: Equity issues and the attribution of past returns. *Journal of Financial Economics*, forthcoming.

Barberis, N., M. Huang, and R. Thaler. 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review* 96:1069–90.

Bebchuk, L., and L. Stole. 1993. Do short-term objectives lead to under-or overinvestment in long-term projects? *Journal of Finance* 48:719–29.

Benartzi, S., R. Michaely, and R. Thaler. 1997. Do changes in dividends signal the future or the past? *Journal of Finance* 52:1007–34.

Bernheim, B. D. 1991. Tax policy and the dividend puzzle. *Rand Journal of Economics* 22:455–76.

Bhattacharya, S. 1979. Imperfect information, dividend policy, and the bird in the hand fallacy. *Bell Journal of Economics and Management Science* 10:259–70.

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

Charest, G. 1978. Dividend information, stock returns, and market efficiency. *Journal of Financial Economics* 6:297–330.

Constantinides, G. 1990. Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy* 98:519–43.

Damodaran, A. 1989. The weekend effect in information releases: A study of earnings and dividend announcements. *Review of Financial Studies* 2:607–23.

Eades, K. M., P. J. Hess, and E. H. Kim. 1985. Market rationality and dividend announcements. *Journal of Financial Economics* 14:581–604.

Eysenck, M. W., and M. C. Eysenck. 1980. Effects of processing depth, distinctiveness, and word frequency on retention. *British Journal of Psychology* 71:263–74.

Fama, E. F., L. Fisher, M. C. Jensen, and R. Roll. 1969. The adjustment of stock prices to new information. *International Economic Review* 10:1–21.

Fama, E. F., and H. Babiak. 1968. Dividend policy: An empirical analysis. *Journal of the American Statistical Association* 63:1132–61.

Genesove, D., and C. Mayer. 2001. Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics* 116:1233–60.

Grullon, G., and R. Michaely. 2002. Dividends, share repurchases, and the substitution hypothesis. *Journal of Finance* 57:1649–84.

Grullon, G., R. Michaely, and B. Swaminathan. 2002. Are dividend changes a sign of firm maturity? *Journal of Business* 75:387–424.

Guttman, I., O. Kadan, and E. Kandel. 2010. Dividend stickiness and strategic pooling. *Review of Financial Studies* 23:4455–95.

Heath, C., S. Huddart, and M. Lang. 1999. Psychological factors and stock option exercise. *Quarterly Journal of Economics* 114: 601–27.

John, K., and J. Williams. 1985. Dividends, dilution, and taxes: A signalling equilibrium. *Journal of Finance* 40:1053–70.

Kahneman, D., J. Knetch, and R. H. Thaler. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy* 98:1325–48.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–91.

Koszegi, B., and M. Rabin. 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics* 121:1133–66.

———. 2009. Reference-dependent consumption plans. *American Economic Review* 99:909–36.

Kumar, P. 1988. Shareholder-manager conflict and the information content of dividends. *Review of Financial Studies* 1:111–36.

Lambrecht, B. M., and S. C. Myers. 2012. A Lintner model of dividends and managerial rents. *Journal of Finance* 67:1761–810.

Lang, H. P., and R. H. Litzenberger. 1989. Dividend announcements: Cash flow signalling vs. free cash flow hypothesis? *Journal of Financial Economics* 24:181–91.

Leary, M. T., and R. Michaely. 2011. Determinants of dividend smoothing: Empirical evidence. *Review of Financial Studies* 24:3197–249.

Leland, H. E., and D. H. Pyle. 1977. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance* 32:371–87.

Lintner, J. 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review* 46:97–113.

Markowitz, H. 1952. The utility of wealth. *Journal of Political Economy* 60:151–58.

Michaely, R., and M. R. Roberts. 2012. Corporate dividend policies: Lessons from private firms. *Review of Financial Studies* 25:711–46.

Miller, M. H. 1986. Behavioral rationality in finance: The case of dividends. *Journal of Business* 59:S451–68.

Miller, M. H., and K. Rock. 1985. Dividend policy under asymmetric information. *Journal of Finance* 40:1031–51.

Odean, T. 1998. Are investors reluctant to realize their losses? *Journal of Finance* 53:1775–98.

Palmere, M., S. L. Benton, J. A. Glover, and R. Ronning. 1983. Elaboration and the recall of main ideas in prose. *Journal of Educational Psychology* 75:898–907.

Samuelson, W., and R. Zeckhauser. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1:7–59.

Shefrin, H. M., and M. Statman. 1984. Explaining investor preference for cash dividends. *Journal of Financial Economics* 13:253–82.

———. 1985. The disposition to sell winners too early and ride losers too long. *Journal of Finance* 40:777–90.

Stein, J. 1989. Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics* 104:655–69.

Thaler, R. H. 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1:39–60.

———. 1999. Mental accounting matters. *Journal of Behavioral Decision Making* 12:183–206.

Tversky, A., and D. Kahneman. 1991. Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics* 106:1039–61.

Tyler, S. W., P. T. Hertel, M. C. McCallum, and H. C. Ellis. 1979. Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory* 5:607–17.

Yoon, P. S., and L. T. Starks. 1995. Signaling, investment opportunities, and dividend announcements. *Review of Financial Studies* 8:995–1018.