Signaling With Reference Points:

Behavioral Foundations for the Lintner Model of Dividends[*]

Malcolm Baker
Harvard Business School and NBER

Jeffrey Wurgler
NYU Stern School of Business and NBER

PRELIMINARY AND INCOMPLETE DRAFT

October 11, 2010

Abstract

We propose a signaling model in which agents value dividends relative to a reference point of prior dividends and exhibit loss aversion, as in a prospect theory value function. Two versions of the model are developed, one in which the manager's utility suffers if the dividend falls below an endogenous prior dividend, and another in which investors sell if this occurs. Managers of firms with strong earnings separate themselves by paying high dividends and still retaining enough earnings to be likely to pay the same dividend next period. Equilibrium dividend policies follow a Lintner partial-adjustment model. We argue that the model accounts for major patterns in dividend policy better than signaling models based on public destruction of value, and we find empirical support for some of its novel predictions.

---

[*] We thank Chris Allen for excellent research assistance. Baker gratefully acknowledges the Division of Research of the Harvard Business School for financial support.

> *I assume [the Lintner model] to be a behavioral model, not only from its form, but because no one has yet been able to derive it as the solution to a maximization problem, despite 30 years of trying.*
>
> --- Merton Miller (1986, p. S467).

I. Introduction

The first-order facts of dividend policy are agreed upon by executives, and what they say has been confirmed in the data. A recent survey of 384 executives by Brav, Graham, Harvey, and Michaely (2005) found that they try to avoid reducing dividends per share (93.8% agreed); that they try to maintain a smooth dividend stream (89.6%); that they are reluctant to make changes that might have to be reversed (77.9%) because there are negative consequences to reducing dividends (88.1%) given that they convey information to investors (80%). The responses are consistent with Lintner's (1956) own survey results and interviews, his famed partial-adjustment model, and a large empirical literature.

Yet the very strongest views expressed in the Brav et al. survey are reserved for mechanisms behind traditional theories of dividend signaling. Executives viewed these as broadly misguided. The notion 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 idea of signaling through costly taxes did not receive much more support (16.6%). These findings cast some doubt on the mechanisms driving signaling models by Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), and Allen, Bernardo, and Welch (2000), among others.

In this paper we use prospect theory of Kahneman and Tversky (1979) to build a signaling model of dividend policy with behavioral foundations. We focus on two features of the prospect theory value function: that values and perceptions are based on losses and gains relative to a reference point; and, that there is more disutility from losses than there is utility from equal-size gains. Reference-dependence and loss aversion are supported by a considerable literature in psychology and a growing body of evidence from finance and economics, as we discuss later.

The essence of the model is that agents evaluate current dividends against the reference point established by past dividends. Because agents 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. The model also yields novel predictions, which we test.

We develop two versions of the model. The first version of the model uses reference point *preferences* as the mechanism for costly signaling. The manager's utility function reflects both a preference for a high stock price and for avoiding a future dividend cut relative to the reference point set by the current dividend. Thus, a reference point effect is built directly into the manager's utility function (which may reflect his perception of the investor utility function). Investors know the shape of the manager's utility function and infer current earnings from dividends paid. The current dividend, and share price, will be greater for firms with high current earnings and thus the savings to insure against a future dividend cut. For this reason, the current dividend can credibly separate profitable and unprofitable firms. This version leads to a Lintner partial-adjustment policy.

The second version of the model, which is more natural, uses reference point *behavior* as the mechanism for signaling. Here, the reference point effect appears in a less direct manner. Some behavioral investors are prone to sell a stock, depressing its price, if its current dividend is below an endogenous reference point set by the past dividend. Again, the likelihood of having to cut the dividend is high for firms with low savings entering the second period, which are firms with low profits and high dividends in the first period. With enough behavioral investors, the model also implies a Lintner policy.

Models based on reference points and loss aversion have some appealing properties. They are conceptually simple and, more importantly, they are based on empirically supported assumptions. They do not ask profitable firms to destroy fundamental value—burn money—in order to distinguish themselves. The second version seems particularly consistent with the views of those who actually set dividend policy as in Brav et al. (2005), while the first version is at least as compatible with those views as existing signaling theories. Being inherently multiperiod and concerned with dividends relative to a reference point set by past dividends, the model naturally explains the real-world focus on changes in dividends, whereas some other signaling theories focus on levels. The model provides (two) behavioral foundations for the Lintner model (Miller (1986) was correct). We discuss other stylized facts about dividend policy and conclude that our model does not explain all of them, but it does at least as well as other signaling theories.

Our model also makes some novel predictions. They revolve around the insight that a dividend that creates a reference point forms a powerful signal. Psychological evidence suggests that memory and salience play a role in the formation of reference points. This suggests that dividends per share will be concentrated in round numbers, which are more memorable and deviations from which are thus more noticeable. In contrast, standard models are continuous in the sense that they ascribe no particular role to round number dividends. Repetition of a particular dividend level also ingrains a reference point.

The data largely support these predictions. Both dividend per share levels and changes are made in round numbers, such as multiples of five or ten cents. Managers that raise dividends strive to exceed round number thresholds. The market reacts asymmetrically when past levels are not reached versus when they are exceeded (a known result), especially when these changes cross a round number threshold (a new result). As our model predicts, this asymmetry is more pronounced when the same dividend per share has been paid for more than four consecutive quarters. As a placebo test, we examine ADRs, where we find, not surprisingly, that the reference point and round numbers are denominated in foreign currency. In this sample, there is nothing special about the market's reaction around zero dividend changes in U.S. dollars.

Other papers have made connections between dividends and reference points or prospect theory. One that uses prospect theory even more heavily than we do is Shefrin and Statman (1984). They 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. Their theory is an important contribution but quite different from ours. Dividends serve no signaling function in their theory, do not follow a Lintner policy, and can be evaluated against a reference level of no dividend, for instance. Another important contribution is Lambrecht and Myers (2010). In their model, managers maximize the present value of the utility of rents that they can extract from profits. They smooth dividends because they have habit formation preferences and rents move in lockstep with dividends given the budget constraint. Lambrecht and Myers do, importantly, derive a Lintner policy in equilibrium, but there are a variety of differences with our theory and predictions. We discuss these papers and other related research. More generally, our paper adds to the literature on behavioral corporate finance surveyed by Baker, Ruback, and Wurgler (2007).

Section II reviews the relevant literature on reference-dependent utility and the formation of reference points. Section III describes the models. Section IV discusses the model's compatibility with the major empirical facts of dividend policy. Section V describes data and presents empirical results for the novel predictions. Section VI concludes.

II. 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.

A. Reference-Dependent Utility and Loss Aversion

We will 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.) Finally, a related phenomenon is the 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.

B. Reference Points

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 prospect theory does not provide answers to these questions. The relevant reference point depends on the setting and, in static choice settings, is frequently unambiguous. For example, in the applications and experiments above, the reference point is obviously 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 prospect theory value function.

Shefrin and Statman (1984) apply these ideas to explain why investors like dividends, although their perspective is very different. Shefrin and Statman argue that investors may 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 its 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 (2009) argue that the stock price that a new CEO inherits is an important reference point for raising new equity. The general 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, which is unlikely to be memorable. 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. Some phone companies sell phone numbers that include round numbers or several repeated digits at a premium.

A stock's 52-week high provides an interesting example of a memorable number that, at least 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 matter for the pricing and deal success in mergers and acquisitions (Baker, Pan, and Wurgler (2010)). Given that an individual shareholder's purchase price also affects her trading behavior, this is also an example of how behavior and perceptions of value may depend on multiple reference points.

C. Past Dividends as Reference Points

This discussion shows that theory alone cannot identify "the" reference point. The typical research process is to consider the setting, hypothesize the nature of the reference point(s), and then see if the empirical evidence supports the hypothesis. In this paper we hypothesize that 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. 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 announced at discrete and regular intervals, and often with some degree of ceremony and fanfare, which encourages the formation of memory. The same level is often repeated for many quarters in a row, further encouraging memory. We will show that they cluster at round numbers, and that changes are commonly in round-number intervals or designed to meet or exceed a round-number threshold. The memorability of the dividend is central to our theory—it increases their power as reference points and, consequently, as signals.

III. A Model of Signaling With Dividends as Reference Points

We present two versions of a dividend-signaling model with reference dependence. The model use biases in investor and managerial behavior, not willful destruction of firm value through investment distortions or taxes, to provide the costly signaling mechanism.

A. Version 1: Reference-Dependent Utility

The first model uses reference point *preferences* as the mechanism for costly signaling. The reference point appears directly in the manager's objective function. The manager dislikes paying a dividend in the second period that is below an endogenous reference point. The likelihood of dividends falling below the reference point is high for firms with low savings entering the second period. These are firms with low profits and high dividends in the first period. For this reason, dividend policy in the first period can credibly separate profitable and unprofitable firms. The basic idea is simple and similar to committing to a deadline. If it is personally costly for the announcer to miss the deadline in utility terms once a public announcement has been made, the announcement itself has some signal value to observers who know the shape of the utility function.

1. Setup

This version of the model has two periods: $t = 1$ and 2. There are two players: a benevolent manager and a risk neutral investor. In the first period, the manager pays an initial dividend and forms a reference point, and the investor forms expectations about dividends in the next period. In the second period, the manager pays a liquidating dividend. There is no discounting. The manager cares about the current stock price P_1 gross of first period dividends as well as the total level of dividends.

The manager's utility function takes the usual form of Miller and Rock (1985) or Stein (1989), combining some weight on the dividend-adjusted stock market price and some weight on his utility of second period dividends:

$$E_m\left[\lambda \hat{P}_1 + (1-\lambda)u(d_1, d_2)\right], \text{ where } \hat{P}_1 = P_1 + d_1 \tag{1}$$

where P_1, d_1 and d_2 are the stock market price and dividends for the firm, u is the manager's utility function for second period dividends, and E_m is the manager's expectations operator. We use dividend-adjusted prices instead of raw prices to eliminate the mechanical downward effect of a dividend on stock price. The usual argument for this general sort of utility function 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.

The interesting aspect of this model is that the manager has a kink in his preferences for second period dividends d_2 around an endogenously determined reference point of d^*:

$$u(d_1, d_2) = d_1 + d^* + (d_2 - d^*)\{d_2 > d^*\} + \beta(d_2 - d^*)\{d_2 < d^*\}. \tag{2}$$

Put simply, the manager cares about fundamental value, or total dividend payments, but with a twist. The level of the reference point d^* comes from historical firm dividend policy, and β is greater than one so that dividends below the reference point are more painful than symmetric dividends above the reference point. This utility function is in the spirit of prospect theory with a kink at a reference point. We leave out the complexity of curvature. From this point forward, the reference point d^* simply equals first period dividends d_1 by assumption. In reality, the reference point d^* and the intensity of the reference point β may be determined by a long history of levels and changes in dividend policy. We test for this in the empirical section.

For simplicity, the manager has no control over fundamental value F, and so it is excluded from his utility. Note that this is a bit different from a traditional signaling model where the manager must destroy firm value to impress the capital markets. There is no fundamental agency problem as there is in Lambrecht and Myers (2010). The manager is not able to keep the cash for himself, and no real value is created or destroyed with dividend policy. This 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:

$$\widetilde{F} = \varepsilon_1 + \varepsilon_2. \tag{3}$$

Think of these as cash earnings that are not observable to outside investors. 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 in learning ε_1 over the investor. Otherwise, there is no signaling problem. Second, the payment of the observable dividend must form the manager'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, we assume that the cash earnings are i.i.d. with a uniform distribution, $\varepsilon_t \sim U[0,2]$.

There is no new equity or debt available simply to finance the payment of dividends. The most the manager can pay in the first period is ε_1, and the most he can pay in a later period t is ε_t plus any savings from previous periods. This leads to the following constraints:

$$0 \le d_1 \le \varepsilon_1 \text{ and } d_2 = \varepsilon_1 + \varepsilon_2 - d_1 . \tag{4}$$

These constraints are not additional assumptions, but implications of the absence of new financing and a benevolent manager.

2. Equilibrium

Prices are determined competitively. Note that P_2 is simply equal to d_2. By contrast, P_1 comes from investor demand, and by extension expectations of d_2. Rather than starting from utility, we skip to a demand function for the risk neutral investor. Demand takes a standard form:

$$w = \gamma\left(E[d_2] - P_1\right), \tag{5}$$

where the investor's demand w is sensitive to mispricing. Risk neutrality means that $\gamma \to \infty$. Market clearing requires that the total demand equal the supply of shares outstanding. If the investor were risk averse, this would add an extra term that depended on the quantity of shares. Risk neutrality means that price is simply:

$$P_1 = E_i[d_2 \mid d_1] = E_i[\varepsilon_1 \mid d_1] + 1 - d_1 . \tag{6}$$

The first term is investors' expectation of this period's cash earnings conditional on dividend policy, the second term is the expectation of next period's cash earnings, which neither player knows, and the third term is the observable dividend paid that reduces savings for the second period dividend payment.

We can now substitute the price from Eq. (6) and the utility of second period dividends from Eq. (1) into the manager's utility function,

$$E_m\left[\lambda\hat{P}_1 + (1-\lambda)u(d_2)\right] = \lambda\left[E_i\left[\varepsilon_1 \mid d_1\right]+1\right] + (1-\lambda)\left[1+\varepsilon_1 - \frac{\beta-1}{4}(2d_1-\varepsilon_1)^2\right]. \tag{7}$$

The second term is the weight (1 – λ) times the manager's expected second period utility, using the definition of d_2 from Eq. (4), and taking expectations over ε_2 over the range from zero to 2, under the assumption that first period dividends are high enough that the second period dividend has a chance of falling below the reference point, or $d_1 \geq \frac{1}{2}\varepsilon_1$. Note that if first period dividends are exactly $\frac{1}{2}\varepsilon_1$, so that the reference point is never triggered, then second term is simply the expectation of second period cash earnings, or 1, plus first period cash earnings ε_1.

We only consider equilibria where the manager pays out a linear function of the first period cash earnings, or $d_1 = \frac{1}{\alpha}\varepsilon_1 + c$, so that with rational investor expectations $E_i\left[\varepsilon_1 \mid d_1\right] = \alpha d_1$. To see whether this is sustainable, we need to check whether the manager will deviate. The manager knows ε_1 in Eq. (7), but can only control d_1, so dividend policy has two offsetting effects in the first order condition,

$$\lambda\alpha - (1-\lambda)\cdot(\beta-1)\cdot(2d_1-\varepsilon_1) = 0. \tag{8}$$

On one hand, increasing the dividend raises the investor's estimate of firm value through the expectation of ε_1. This increases the manager's utility. On the other hand, increasing the dividend raises the manager's reference point and in turn raises the probability that this higher reference point cannot be satisfied with second period dividends. This decreases utility.

Dividend policy can be described as follows. In equilibrium, the manager pays half of the first period cash earnings, or $\alpha = 2$, plus an extra amount c that depends on horizons λ and the intensity of the reference point preferences β,

$$d_1 = \tfrac{1}{2}\varepsilon_1 + \frac{\lambda}{1-\lambda} \cdot \frac{1}{\beta - 1}. \tag{9}$$

Intuitively, dividends are higher when the manager's horizons are short and when the intensity of reference point preferences is low. Recall that Eq. (7) applies only when $d_1 \geq \tfrac{1}{2}\varepsilon_1$. This will hold for all relevant parameter values. If the manager puts a sufficiently high weight on current stock prices or reference point effects are mild with $\beta < 1 + \tfrac{2}{\varepsilon_1} \cdot \tfrac{\lambda}{1-\lambda}$, then the second term can exceed $\tfrac{1}{2}\varepsilon_1$, and the budget constraint of Eq. (4) will be violated by dividend policy in Eq. (9). In this case, because of high incentives to signal and low costs of signaling, the equilibrium quite naturally involves a high dividend equilibrium, where $d_1 = \varepsilon_1$.

In all but the most extreme signaling equilibria, the manager follows a Lintner-type dividend policy, with dividends adjusting partway toward cash earnings. Putting the two cases together, we have the following range of equilibria.

Proposition 1. Reference-dependent utility and dividend policy. *Strong reference point effects* $\beta > 1 + \tfrac{2}{\varepsilon_1}\tfrac{\lambda}{1-\lambda}$ *lead to a low dividend equilibrium* $d_1 = \tfrac{1}{2}\varepsilon_1 + \tfrac{\lambda}{1-\lambda} \cdot \tfrac{1}{\beta-1}$ *and a strong announcement effect* $E_i\left[\varepsilon_1 \mid d_1\right] = 2d_1$. *Weak reference point effects* $\beta < 1 + \tfrac{2}{\varepsilon_1}\tfrac{\lambda}{1-\lambda}$ *lead to a high dividend equilibrium* $d_1 = \varepsilon_1$ *and a weak announcement effect* $E_i\left[\varepsilon_1 \mid d_1\right] = d_1$.

B. Version 2: Reference-Dependent Behavior

The second version of the model uses reference point *behavior* as the mechanism for costly signaling. Some investors are prone to sell stock if its dividend does not exceed an endogenous reference point. This has a negative impact on stock price. The likelihood of this negative impact on stock price is high for firms with low savings entering the second period. As in the previous version, these are firms with low profits and high dividends in the first period, so dividend policy in the first period can credibly separate profitable and unprofitable firms.

1. Setup

This version of the model needs three periods: $t = 1$, 2, and 3. This version involves three players: a benevolent firm manager, a risk-neutral behavioral investor, and a risk-averse arbitrageur. The behavioral investor is rational in all respects except one. He will sell his shares with probability p if the firm pays a dividend that is below his endogenously determined reference point. We use an exogenous p in place of β in the previous version of the model as a way to keep track of the intensity of the reference point. The level of the reference point d^* comes from historical firm dividend policy.

Nothing interesting happens in the third period other than the payment of a liquidating dividend. In the second period, the manager pays a second dividend, the behavioral investor reacts to the dividend decision, given his reference point, and both investors form expectations about the final liquidating dividend. In the first period, the manager pays an initial dividend, the behavioral investor forms a reference point, and both investors form expectations about dividends over the next two periods. There is no discounting in this version, either.

The manager's utility function is as before but with an extra period. He places some weight on the dividend-adjusted stock market prices and some weight on fundamental value:

$$E_m\left[\tfrac{1}{2}\left(\hat{P}_1 + \hat{P}_2\right) + \widetilde{F}\right], \text{ where } \hat{P}_t = P_t + \sum\nolimits_1^t d_t \tag{10}$$

and P_1, P_2, d_1, and d_2 are the stock market prices and dividends for the firm at $t = 1$ and 2, F is the uncertain fundamental value of the firm, and E_m is the manager's expectations operator. We earlier discussed the logic behind this sort of utility function and why we use dividend-adjusted prices instead of raw prices. The fundamental value of the firm appears in three installments:

$$\widetilde{F} = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 . \tag{11}$$

To isolate the unique mechanisms of this version of the model, it shares many features with the first version. To briefly restate, the manager has no control over fundamental value F. There is no fundamental agency problem. The manager is not able to keep the cash for himself, and no real value is created or destroyed with dividend policy. Cash earnings are not observable to the investors, otherwise there is no signaling problem, and the payment of the observable dividend forms the behavioral investor's reference point, not the firm's reported financials. Otherwise, the manager has no lever to signal his information about ε_1. Cash earnings are i.i.d. with a uniform distribution, $\varepsilon_t \sim U[0,2]$.

Once again, there is no new equity or debt available to finance the payment of dividends. The most the manager can pay in the first period is ε_1, and the most he can pay in a later period t is ε_t plus any savings from previous periods. This leads to the following constraints:

$$0 \le d_1 \le \varepsilon_1 ,\ 0 \le d_2 \le \varepsilon_1 + \varepsilon_2 - d_1 ,\ d_3 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 - d_1 - d_2 . \tag{12}$$

2. Equilibrium

Prices are determined competitively. Note that P_3 is simply equal to d_3. By contrast, P_1 and P_2 come from investor demand, and by extension expectations of d_2 and d_3. We again skip to investor demand functions of a standard form:

$$w_{i,t} = \gamma_{i,t}\left(E\left[P_{t+1} + d_{t+1}\right] - P_t\right), \tag{13}$$

where the risk-neutral behavioral investor's demand $w_{b,t}$ is very sensitive to mispricing, with $\gamma_{b,1}$ taking a large value $\gamma \rightarrow \infty$ at $t = 1$. At $t = 2$, $\gamma_{b,2}$ is equal to 0 with probability p, if the current dividend payment falls below the behavioral investor's reference point d^*, and $\gamma_{b,2}$ is equal to γ if the dividend exceeds his reference point or with probability (1-p) if it does not. From this point forward, the behavioral investor's reference point d^* is d_1 by assumption. Again, in reality, the reference point d^* and p may be determined by a long history of levels and changes in dividend policy. The risk-averse arbitrageur's demand $w_{a,t}$ is less sensitive to mispricing, $\gamma_{a,t} = \gamma_a << \gamma$.

The behavioral investors' demand in Eq. (13) is, to be precise, best described as reference point *behavior*. It could derive from strong reference point preferences, of course, but the essential feature is that the behavioral investor sells his shares if dividends come in below his reference point. The previous version of the model involved prospect theory preferences more explicitly, where the manager internalized his own or a behavioral investor's utility loss directly in the second period.

Market clearing requires that the total demand of the behavioral investor and the arbitrageur equal the supply of shares outstanding Q. Using backward induction, we compute P_2 first. There is no reason to hold back dividends in the second period. To maximize P_2, the manager will maximize demand by paying the largest dividend possible, so that

$$d_1 + d_2 = \varepsilon_1 + \varepsilon_2 \,. \tag{14}$$

There are two possible aggregate demand functions at $t = 2$. The first is when the reference point is met, $d_2 \geq d^* = d_1$. Then price is set by the behavioral investor, and $P_2^{d \geq d^*} \rightarrow E_b[\varepsilon_3] = 1$. The second case is when the reference point is not met, $d_2 < d^* = d_1$. Then there is a probability p that prices are set by the arbitrageur. If so, $Q = w_{a,t} = \gamma_a \left(E_a[\varepsilon_3] - P_2\right)$.

Combining this with the probability (1-p) that the price is still set by the behavioral investor nonetheless, gives a lower expected price when the second dividend falls below d^*,

$$P_2^{d<d^*} = 1 - \frac{pQ}{\gamma_a}. \tag{15}$$

The basic idea here is that the behavioral investor is the ideal holder of the shares, but his preferences lead him to sell the shares to an arbitrageur, who perhaps ultimately sells them to a new behavioral investor. This setup is meant to capture a possible dislocation in the firm's shares from a dividend cut that leads to a fire sale price. In this model, the repricing occurs instantly. More realistically, behavioral investors might sell with a lag, possibly creating momentum effects rather than instant overreaction.

Next, we compute prices in the first period. The behavioral investor sets these, because there is no historical reference point for dividend policy. This means that the first period price is a probability-weighted average of the two scenarios at $t = 2$. The behavioral investor is quasi-rational in the sense that he recognizes that the first period dividend will create reference point behavior that he cannot control in the second period:

$$P_1 = 2 + \left(E_b[\varepsilon_1 \mid d_1] - d_1\right) - \frac{pQ\left(d_1 - \frac{1}{2}E_b[\varepsilon_1 \mid d_1]\right)}{\gamma_a}\left\{d_1 > \tfrac{1}{2}E_b[\varepsilon_1 \mid d_1]\right\} \tag{16}$$

The first term is the expectation of unknown future dividends in periods 2 and 3. The second term is the level of savings from the first period to be paid out with a delay, and the third term is the expected reduction in the second period price arising from the possibility of a sale of shares from the behavioral investor to the arbitrageur. This sale will not occur if dividend policy is sufficiently conservative.

Equilibrium dividend policy depends on the parameter values: the manager pays out a fraction of cash earnings $d_1 = \frac{4p}{p+2\gamma_a/Q}\varepsilon_1$. Provided the intensity of reference point behavior p and the limits to arbitrage are sufficiently large, the manager will not pay out the full amount ε_1. Instead, he will follow a Lintner-type dividend policy, setting dividends d_1 to move partially toward cash earnings ε_1. We build up this logic by analyzing two special cases first.

As a first case, consider a possible equilibrium where the manager pays out the maximum $d_1 = \varepsilon_1$, so that with rational investor expectations $E_b[\varepsilon_1 \mid d_1] = d_1$. To see whether this is sustainable, we need to check whether the manager will deviate. His objective is to maximize the average stock price gross of dividends paid in the first two periods, or

$$\max\ E_m[P_1 + d_1 + P_2 + d_1 + d_2] = E_b[\varepsilon_1 \mid d_1]\left(1 + \frac{pQ}{2\gamma_a}\right) - d_1\left(\frac{2pQ}{\gamma_a}\right) + 4 + \varepsilon_1\left(1 + \frac{pQ}{2\gamma_a}\right). \quad (17)$$

The manager knows ε_1, but can only control d_1, so only the first two terms are relevant. The objective function shows that dividend policy has two offsetting effects. On one hand, in the first term, increasing the dividend raises the behavioral investor's estimate of firm value through the expectation of ε_1. On the other hand, in the second term, increasing the dividend raises the behavioral investor's reference point and in turn raises the probability that he will sell his shares at $t = 2$ when this higher reference point is not satisfied with second period dividends. This results in a lower second period share price, which the manager does not like.

From the equilibrium level of dividends of ε_1, a 0.001 reduction in the dividend would increase the manager's utility by $0.001 \cdot \left(\frac{3pQ}{2\gamma_1} - 1\right)$. So this is an equilibrium only for a sufficiently small p or a sufficiently large $\frac{\gamma_1}{Q}$. The intuition is as follows. When the reference point has a small effect on the behavioral investor or if the arbitrageur's risk bearing capacity is large relative to the size of the firm, the manager will pay out as much as possible to increase the current stock price. If the reference point has a large effect and there are considerable limits to arbitrage, the manager will hold back some cash earnings and engage in at least some dividend smoothing, making this equilibrium unsustainable.

As a second case, consider another possible equilibrium, at the other end of the spectrum, which has the manager paying out $d_1 = \frac{1}{2}\varepsilon_1$, thus eliminating any possibility of a price distortions at $t = 2$. By saving a half of the current cash earnings, the manager ensures that the behavioral investor's reference point will be met, even with the worst case second period results. For this equilibrium, investors believe the reference point will be satisfied, so Eq. (17) is modified as follows:

$$\max\ E_m\left[P_1 + d_1 + P_2 + d_1 + d_2\right] = E_b\left[\varepsilon_1 \mid d_1\right] - \left(d_1 - \tfrac{1}{2}\varepsilon_1\right)\left(\frac{pQ}{\gamma_a}\right)\left\{d_1 > \tfrac{1}{2}\varepsilon_1\right\} + 4 + \varepsilon_1. \tag{18}$$

Again, the manager can only control the first two terms. This time, because investors believe the reference point will not be tested, each 0.001 increase in the dividend raises expectations of ε_1 by 0.002.

From the equilibrium level of dividends of $\frac{1}{2}\varepsilon_1$, a 0.001 increase in the dividend would increase the manager's utility by $0.001\cdot\left(2 - \frac{pQ}{\gamma_1}\right)$. Therefore, this is an equilibrium only for a sufficiently large p or a sufficiently small $\frac{\gamma_1}{Q}$. The intuition is the same as before.

To solve for a more general case, assume that the manager uses the payout rule $d_1 = \frac{1}{\alpha}\varepsilon_1$, where α is strictly between 1 and 2. In this case, because α is less than 2 and there is some expected price distortion in equilibrium, the manager's objective function follows Eq. (17). From the equilibrium level of dividends of $\frac{1}{\alpha}\varepsilon_1$, a small change in the dividend in either direction would change the manager's utility by $\pm 0.001\cdot\left(\alpha - \frac{(4-\alpha)pQ}{2\gamma_1}\right)$. As a result, this equilibrium is sustainable only when

$$\alpha = \frac{4p}{p + 2\frac{\gamma_a}{Q}}. \qquad (19)$$

This subsumes the other two cases. Moreover, the intuition for *level* of dividends is the same as the logic for the *existence* of an equilibrium in the two special cases. The level of dividends is lower when the reference point has a large effect on the behavioral investor or if the arbitrageur's risk bearing capacity is small relative to the size of the firm. In these situations, the desire to smooth dividends over time outweighs the desire to signal current cash earnings. This is somewhat more knife-edged, because the manager is no worse off from deviating from the equilibrium level of dividends.

Putting the three cases together, we have the following range of equilibria.

Proposition 2. Reference-dependent behavior and dividend policy. *Strong reference point effects* $p > 2\cdot\frac{\gamma_a}{Q}$ *lead to a low dividend equilibrium* $d_1 = \frac{1}{2}\varepsilon_1$ *and a strong announcement effect* $E_b\left[\varepsilon_1 \mid d_1\right] = 2d_1$. *Intermediate reference point effects* $\frac{2}{3}\cdot\frac{\gamma_a}{Q} < p < 2\cdot\frac{\gamma_a}{Q}$ *lead to intermediate outcomes* $d_1 = \frac{1}{4p}\left(p + 2\frac{\gamma_a}{Q}\right)\varepsilon_1$ *and* $E_b\left[\varepsilon_1 \mid d_1\right] = 4p\left(1/p + 2\frac{\gamma_a}{Q}\right)d_1$. *Weak reference point effects* $p < \frac{2}{3}\cdot\frac{\gamma_a}{Q}$ *lead to a high dividend equilibrium* $d_1 = \varepsilon_1$ *and a weak announcement effect* $E_i\left[\varepsilon_1 \mid d_1\right] = d_1$.

IV. The Model and Prior Evidence

Dividend policy is an area so awash with empirical facts that any new model must be judged as much on its ability to fit those facts as on the success of its novel predictions. We consider several stylized facts here in light of the model. While it certainly cannot explain all of the known facts, a model of signaling with reference points appears to capture many of them at least as well as existing approaches, of which the best known are Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), and Allen, Bernardo, and Welch (2000). Our model may also perform as well in some respects as theories based on agency problems, catering motives, or clientele effects, although to keep the discussion manageable we will not make such comparisons.

A. Surveys

Dividend policy is an explicit choice of executives (more precisely, the board). The proposition that their behavior may be guided not by their own hands but by an unseen higher market force, and therefore survey evidence should be disregarded, is clearly inappropriate here. We therefore view the fact that our model is consistent with what managers say about dividend policy as an important success.

The strongest results fit nicely with the reference point 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.

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. summarize taxes as of "second-order importance" (p. 521).

Brav et al. followed up on 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" (pp. 522-523). Note that our model doesn't rely on voluntary destruction of value or real economic cost to create the opportunity for a credible signal. Although the mechanisms and assumptions behind our own model were not explicitly assessed in the Brav et al. survey, it is hard to imagine they would receive less support than this.

Finally, it is notable that standard signaling theories do not naturally focus on dividends per share. Dividend policy tends to be defined in more "economic" terms such as dividend yield or payout ratio, which are less salient to the average investor. Dividend policy measured in these units would not make natural reference points, however, perhaps explaining why stability of the level of dividends per share is the most common target. We will return to the salience of dividends per share in our own empirical tests.

B. Dividend Policy and the Lintner Model

As Miller (1986) hypothesized, the Lintner (1956) model can be given behavioral foundations. Given reasonable parameter values, both versions of our model predict that dividends follow a partial-adjustment policy and are, more generally, smoothed relative to earnings (Fama and Babiak (1968)).

In the first and simplest version of the model, dividends are simply half of current earnings when the manager is sufficiently concerned with meeting the reference point of last period's dividends. In the second version, there is some heterogeneity in the speed of adjustment coefficient (or more precisely, payout ratio). Namely, it is lower when arbitrage forces are weak relative to the number of behavioral investors. We are not aware of any direct tests of this point, but it is intuitively consistent with the fact that mature, liquid, stable firms have higher payout ratios than small firms that are harder to trade and arbitrage. Brav et al. (2005) find that average target payout ratios have fallen over time, consistent with the secular increase in the fraction of younger, smaller firms in the public markets (Fama and French (2001)). (Of course, the smoothest policy of all is to pay no dividends and thereby take no risk of having to cut them. Our model would predict this if the behavioral investors are sufficiently trigger-happy.)

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. Lambrecht and Myers (2010)'s model is highly compatible with the Lintner model and smoothing. It is not a signaling model so we do not include it in our horse race. It has many appealing features; on the other hand, it seems unrealistic that thousands of large, established, public U.S. firms smooth their payouts because a coalition of habit-formation managers prefer to smooth their stealing. The required magnitude of rent extraction alone may often be too great even for firms with the poorest of corporate governance.

C. Announcement Effects

Even if executives disavow standard signaling models, it is clear that shareholders care about dividends. 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. See also Charest (1978) and subsequent papers.

That dividend cuts would be received especially poorly is incorporated in the setup of the second version of our model, where behavioral investors are assumed to sell if the dividend does not meet the previous dividend. In the first model, the mechanism is slightly more substantive, and built on investors' knowledge that the manager will cut dividends only if required by poor earnings realizations. To explicitly derive the asymmetry in announcement effects, however, we would need a model that is at least one period longer. In any case, none of the standard signaling models offers a direct explanation for the asymmetry in announcement effects. We conclude that signaling with dividends as reference points explains the empirical facts reviewed here as well as or better than standard signaling models.

V. New Predictions and Tests

The power of dividends as a signal is directly proportional to their use as reference points. Firms that wish to signal in the manner of our model will not hide their dividend, especially when it is not a decrease. Firms in our model will design dividends to be as salient and easy to recall as possible. The novel predictions of our model are based on these principles.

By contrast, standard signaling theories feature highly sophisticated investors who can solve complicated signal extraction problems; the least of their troubles would be remembering that last period's dividend was $0.1323811 per share. Indeed, for static one-period models, there is no need for investors to remember anything.

A. Salience, Ease of Recall, and Repetition

What makes a number like dividends per share memorable? As noted earlier, round numbers are easier to remember. Quarterly repetition of the same dividend helps as well. In our main sample of U.S. dividend payers, we examine whether dividends and dividend changes concentrate on round numbers, whether the use of salient numbers affects how the market reacts to dividend changes, and whether repetition increases the strength of a particular dividend level as a reference point as measured by the market reaction to changes. After this, we consider a sample of ADRs and how dividend reference points fail to translate across currencies.

B. Main Sample

Our primary sample of U.S. dividend payments is summarized in Table 1. We obtain dividend data 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 U.S., 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 also eliminate dividend payments of 0, dividend payments greater than \$2.00 per share, and dividends for which there is no declaration date (CRSP:DCLRDT). Otherwise, the entire CRSP database ending in 2009 appears in Table 1. The data here start with the beginning of the CRSP file in 1926, but restricting attention to more recent periods does not change the economic or statistical conclusions below.

Our main variable of interest is the raw dividend payment per share. It is easier to think about dividend policy as clearing a threshold or creating a new reference point in raw, rather than split-adjusted terms. The median dividend payment is \$0.195. For changes in dividends, we have a choice. We can examine only changes where no split has occurred since the last dividend or we can look at split adjusted differences. For the CRSP sample, where quantity of data is not a problem, we look only at pure changes, where no split has occurred. Even with the more restrictive definition of pure changes, we have almost 250,000 observations. The majority of quarterly dividend changes are zero, and only a small fraction are negative. We also look at whether a dividend change clears a threshold of \$0.10, \$0.05, or \$0.025. For the median dividend payment, reaching the next \$0.10, for example, would require an increase of \$0.05.

We also examine the length of a dividend streak, under the assumption that a longer dividend streak ingrains a reference point and makes the market reaction to missing it stronger. The median streak in our sample is 4, meaning that the typical dividend decision follows four quarters of constant dividends. There is considerable heterogeneity in streaks, with streaks of more than 12 quarters being not uncommon.

To measure the market 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. On average, a dividend declaration is met with approximately a 20 basis point abnormal return. The median abnormal return is also zero. This is a sample of firms that did not omit a dividend, so a slightly positive average is not surprising. We also measure volume over the period 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.

C. Round Numbers and Reaching for Thresholds

Dividends are paid in round numbers. This is apparent in a simple histogram of dividends per share in Panel A of Figure 1. A nickel per share is the modal dividend, a dime the second most common amount, and a quarter is very nearly the third most common amount. There are notable spikes at other round multiples of $0.05. Panel B shows that the most common values for the second and third digits are 0.050, followed by 0.000, and to a much lesser extent 0.025 and 0.075. Other round multiples of 0.01 are somewhat less common, and non-round values are rare.

Dividend changes are also made in round numbers. Figure 2 shows dividend changes when no split occurs between dividend payments. The most obvious patterns in dividend changes are the very large mass at exactly zero, a discontinuity between negative and positive changes even when the mass point at zero is removed, and considerably more clustering just above zero than just below zero. Panel A shows the discontinuity at zero. Little else is even perceptible because the density at zero renders the rest of the distribution inconsequential. When we remove zero changes in Panel B, clustering just to the right of zero is much more apparent. Again, the density above zero is many times the density at equivalent changes below zero. When we further split the sample into increases and decreases, it is also clear that the left tail of the distribution of dividend cuts in Panel D is longer than the right tail of dividend increases in Panel C.

Another pattern is the tendency for increases to reach a threshold, presumably contributing to the salience of the new level. By threshold, we are referring to the next round number in dividends per share. 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.

Figure 3 shows this pattern. When we center the change in dividends—for dividend increases only—on one of these thresholds, it is apparent that the modal increase is exactly to the next threshold. In each case, we look in the neighborhood of the threshold, within but not including $0.025 above and below the threshold. For example, for the firm that is currently paying $0.11, Panel A shows that paying $0.20 is much more likely than paying $0.19 or $0.21.

Panel B shows the same result around the next \$0.05 threshold, and Panel C shows the next \$0.025 threshold.

One question is whether this is simply restating the fact from Figure 1 that firms tend to pay dividends in round numbers. We also check situations where a firm is not starting at a \$0.025 threshold in Panels D, E, and F. The same pattern appears. Clearly, boards think of communicating dividend policy in an easy to recall dollar per share quantity, rather the alternative of deciding on a very specific dollar amount and dividing this equally among shareholders.

D. Market Reaction

Figure 4 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 3-day abnormal return for each group. The pattern in both Panels is similar. Dividend cuts are greeted with a larger negative return than dividend increases of the same magnitude. The difference is roughly a factor of two—remarkably similar to estimates of the degree of loss aversion from other settings. In fact, the whole response curve is strikingly similar to the prospect theory value function, including its features of concavity in gains and convexity in losses. While this is not a direct prediction of the simple managerial utility function in Model 1, a more elaborate model of investor preferences could in principle deliver this sort of pattern.

We examine this pattern somewhat more formally in Table 2, where we estimate piecewise linear regressions of the market reaction on the change in dividends per share to trace out the patterns in Figure 4 in a regression framework. We are particularly interested in the shift in slope below and above zero. The first regression is a simple linear regression. Each \$0.01 change in dividends leads to a 9 basis point market reaction.

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 71 basis points for each \$0.01 change. Small increases in dividends up to \$0.025 are greeted with a market reaction of 35 basis points, or approximately half the slope that we observe in dividend 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 coefficients between -\$0.10 and zero to those between zero and +\$0.10. The slope for dividend cuts is larger both economically and statistically, as one might have guessed from Figure 4. In the final specification, 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. The conclusions are the same.

The analysis thus far is consistent with both versions of the model. Table 3 provides evidence that is particularly suggestive of the second version, where behavioral investors sell if they are disappointed by the dividend. The table shows that the stronger market reaction comes with stronger volume. We repeat the analysis in Table 2 but replace the market reaction with abnormal value as the dependent variable. Both dividend 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 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 667 basis points, or six percent more than normal volume. Similar dividend increases are also associated with higher volume but the rate is somewhat smaller at 570 basis points. Once again, the joint test of the differential sensitivity of volume to dividend changes above and below zero is statistically strong: The market reacts with greater volume following a dividend cut.

We also look for patterns around threshold dividend changes. Instead of sorting the sample around zero dividend change, we sort it around the next \$0.10, \$0.05, or \$0.025 threshold. This means implicitly that we are capturing both the threshold above and the threshold below the current level of dividends in our sorts in Figure 5 and our piecewise linear regressions in Table 4, which themselves use threshold breakpoints. Figure 5 shows a similar pattern. Changes that do not cross a round number threshold elicit a neutral market reaction, changes that cross a threshold from below have a positive reaction similar to that in Figure 4, and changes that cross a threshold from above have a slightly stronger negative reaction than in Figure 4. The numerical results in Table 4 show this somewhat more clearly. The change in market reaction per \$0.01 change in dividends below the lower threshold is 100 basis points, versus 71 basis points in Table 2. The same comparison for thresholds of \$0.05 is 66 basis points versus 57 basis points. For dividend increases the differences are smaller at 39 versus 35 basis points, and for a \$0.05 increase, the effect is actually somewhat smaller at 22 versus 27 basis points. In short, threshold effects are important on the downside, suggesting that round number thresholds are important reference points.

Another, and perhaps cleaner, test of reference points is to examine dividend streaks. If memory is an important part of reference point formation, then repeated dividends of the same amount per share may be stronger reference points. Hence, cutting or raising a dividend after a long streak may have stronger market reactions. We partition the sample into three categories: Decisions following a change in the prior quarter; decisions following no change for up to four quarters, the periodicity of the typical annual board review of dividend policy; and dividend decisions following no change for more than four quarters. The basic idea is that long streaks constitute stronger reference points, so the patterns in Table 2 should be more pronounced as the streak lengthens. Again, we do this analysis two ways: Once with sorts in Figure 6 and once with a piecewise linear regression in Table 5. The results are as expected. The no streak sample has essentially no difference between the effect of a $0.01 decrease and a $0.01 increase in the neighborhood of zero. Meanwhile, the patterns are successively stronger for the short and long streak samples.

Take 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 93 basis points per $0.01 change in dividends just below zero, versus 71 basis points in Table 2. The market reaction to a dividend increase is also larger at 52 versus 35 basis points in Table 2, despite being half of the reaction to a dividend cut in Table 4. Taken together, these results suggest that repetition increases the strength of a reference point.

E. BP-Amoco: A Case Study

Our last set of tests 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 an ADR trading in the U.S. is by definition unable to create a reference point in two different currencies simultaneously.

The case of BP-Amoco 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 retained. Ultimately, dividend policy during the transition was managed carefully so as not to upset dollar-dividend reference points that had been created for Amoco shareholders over many years, as they now owned a large fraction of BP shares.

F. ADR Sample

For a broader analysis, we start with a list of ADRs and matched parents from Datastream over the period from 1990 through 2009 described in Table 6. We restrict the sample to firms with an ADR traded on the New York Stock Exchange, the NASDAQ, and other U.S. OTC 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, so our tendency in forming a sample is one of inclusion rather than restricting attention to the cleanest situations. 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.

For these Datastream codes, we gather information on dividends paid per share (Datastream:DD) in each month. We restrict attention to the following dividend types (Datastream:DT): QTR, HYR, YR, FIN, INT. While we would like to limit attention only to quarterly dividends, semi-annual and annual dividends are more common abroad. 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 starting quantity of data, we use split adjusted values, so that we can examine changes in more cases.

Our primary 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, from Yen to Euro, so the levels of dividends per share are sometimes an order of magnitude higher. The median for ADRs is \$0.194. The median dividend payment for the parent sample is 1.2, which includes 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 essential nature of a reference point analysis, so we leave them in raw terms.

Our specific tests involve the market reaction around zero dividend changes for ADRs. 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. Dividend payments are matched to declaration dates that occur for up to three previous months in an attempt to increase coverage. We also use a slightly larger window to capture the lower quality of Worldscope's declaration dates. The average dividend announcement return is approximately 110 basis points, though the median is again zero.

G. A Placebo Test

We conduct a final test that confirms what is clear in the BP-Amoco case. Namely, that zero change in dividends has no special significance for ADRs. In other words, investors do not care about dividend cuts per se, rather they care about a cut from a mutually agreed upon reference point. Because reference points cannot hold simultaneously in two currencies, ADR dividends in most cases freely cross the zero change boundary and the market reaction is similarly unremarkable in this range.

Figure 8 shows the dividend policy of ADRs measured in both dollars and local currency. Dividend changes in US 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 2. Part of this is because we broadened the sample as much as possible, perhaps at the cost of including some special or liquidating dividends, and part of this may be because the Datastream data is lower quality.

We examine the market reaction to these changes in Table 7. The first observation is that the relationship between dividend changes and the market reaction is everywhere less economically and statistically significant. The R-squared drops from 0.0127 to 0.0004. Moreover, there is no clear pattern in the neighborhood of zero. The reaction to cuts is on average almost the same as the reaction to increases. Together, these results suggest that neither ADR boards nor investors view past dividends – paid in dollars – as an important reference point. The corollary from this placebo test is that changes in dividend policy are important because of an endogenously reference point, not because changes in this neighborhood would otherwise have been economically important.

VI. Conclusion

Standard dividend signaling theories posit that executives use dividends to destroy some firm value and thereby signal that plenty of value remains. The money burning takes the form of tax-inefficient distributions, foregone 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 a more realistic signaling theory. We use core features of the prospect theory value function to create two versions of a model in which past dividends are reference points against which future dividends are judged. The theory is consistent with several important aspects of the data, including survey evidence, patterns of market reaction to dividend announcements, and dividend smoothing and the Lintner partial-adjustment model. We also find support for its novel prediction that dividends are formed in ways that make them memorable and thus stronger reference points.

References

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

Allen, Franklin, Antonio E. Bernardo, and Ivo Welch, 2000, A theory of dividends based on tax clienteles, *Journal of Finance* 55, 2499-2536.

Arkes, Hal, David Hirshleifer, Danling Jiang, and Sonya 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 (Volume 2)*, K.W. Spence and J.T. Spence, eds., New York: Academic Press, 89-195.

Baker, Malcolm P., Stefan Nagel, and Jeffrey Wurgler, 2007, The effect of dividends on consumption, *Brookings Papers on Economic Activity*, 277-291.

Baker, Malcolm, Xin Pan, and Jeffrey Wurgler, 2009, A reference point theory of mergers and acquisitions, Harvard University working paper.

Baker, Malcolm P., Richard R. Ruback, and Jeffrey Wurgler, 2007, Behavioral corporate finance: A survey, In *The Handbook of Corporate Finance: Empirical Corporate Finance*, Espen Eckbo ed, New York: Elsevier/North Holland.

Baker, Malcolm, and Yuhai Xuan, 2009, Under new management: equity issues and the attribution of past returns, Harvard University working paper.

Bernheim, B. Douglas, 1991, Tax policy and the dividend puzzle, *Rand Journal of Economics* 22, 455-476.

Bhattacharya, Sudipto, 1979, Imperfect information, dividend policy, and the bird in the hand fallacy, *Bell Journal of Economics and Management Science* 10, 259-270.

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni 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, George, 1990, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519-543.

Fama, Eugene F., and Henry Babiak, 1968, Dividend policy: An empirical analysis, *Journal of the American Statistical Association* 63, 1132-1161.

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60, 3-43.

Heath, Chip, Stephen Huddart, and Mark Lang, 1999, Psychological factors and stock option exercise, *Quarterly Journal of Economics* 114, 601-627.

Genesove, David, and Christopher Mayer, 2001, Loss aversion and seller behavior: Evidence from the housing market, *Quarterly Journal of Economics* 116, 1233-1260.

John, Kose, and Joseph Williams, 1985, Dividends, dilution, and taxes: A signalling equilibrium, *Journal of Finance* 40, 1053-1070.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263-291.

Kahneman, Daniel, J. Knetch, and Richard H. Thaler, 1990, Experimental tests of the endowment effect and the Coase theorem, *Journal of Political Economy* 98, 1325-1348.

Koszegi, Botond, and Matthew Rabin, 2006, A model of reference-dependent preferences, *Quarterly Journal of Economics* 121, 1133-1166.

Koszegi, Botond, and Matthew Rabin, 2009, Reference-dependent consumption plans, *American Economic Review* 99, 909-936.

Kumar, P., 1988, Shareholder-manager conflict and the information content of dividends, *Review of Financial Studies* 1, 111-136.

Lambrecht, Bart M., and Stewart C. Myers, 2010, A Lintner model of dividends and managerial rents, MIT Sloan School of Management working paper.

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

Markowitz, Harry, 1952, The utility of wealth, *Journal of Political Economy* 60, 151-158.

Miller, Merton H., 1986, Behavioral rationality in finance: The case of dividends, *Journal of Business* 59, S451-S468.

Miller, Merton H., and Kevin Rock, 1985, Dividend policy under asymmetric information, *Journal of Finance* 40, 1031-1051.

Odean, Terrence, 1998, Are investors reluctant to realize their losses? *Journal of Finance* 53, 1775-1798.

Samuelson, William, and Richard Zeckhauser, 1988, Status quo bias in decision making, *Journal of Risk and Uncertainty* 1, 7-59.

Shefrin, Hersh M., and Meir Statman, 1984, Explaining investor preference for cash dividends, *Journal of Financial Economics* 13, 253-282.

Shefrin, Hersh M., and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long, *Journal of Finance* 40, 777-790.

Stein, Jeremy, 1989, Efficient capital markets, inefficient firms: A model of myopic corporate behavior, *Quarterly Journal of Economics* 104, 655-669.

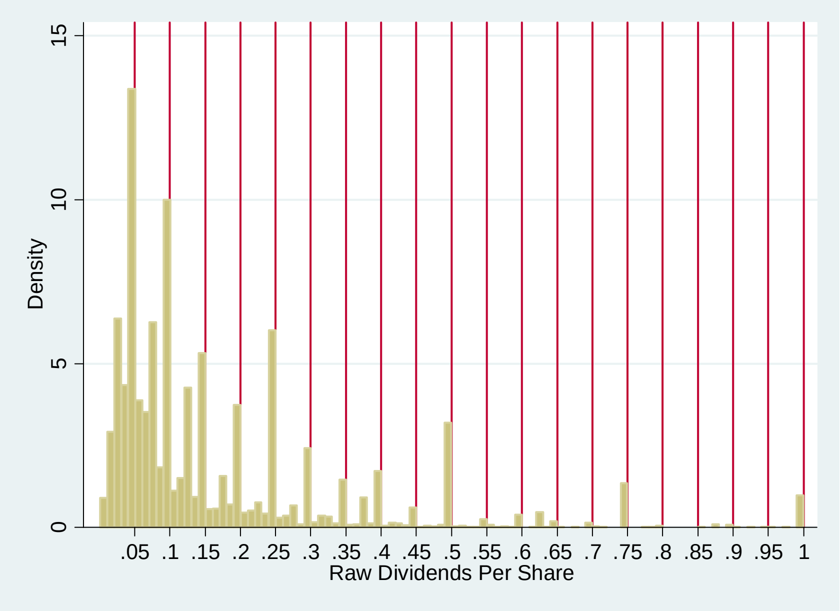
Thaler, Richard H., 1980, Toward a positive theory of consumer choice, *Journal of Economic Behavior and Organization* 1, 39-60.

Thaler, Richard H., 1999, Mental accounting matters, *Journal of Behavioral Decision Making* 12, 183-206.

Tversky, Amos, and Daniel Kahneman, 1991, Loss aversion in riskless choice: A reference-dependent model, *Quarterly Journal of Economics* 106, 1039-1061.

Figure 1. Dividends per share. Histogram of dividends per share and the second and third digits of dividends per share. Panel A shows the distribution of announced dividends per share, while Panel B shows the distribution of the second and third decimal in the announced dividends per share. The sample includes all records from the CRSP event file with a Distribution Codes (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 US, Americus Trust Components, closed-end funds, and REITs) and nonmissing data on the amount of the dividend and the declaration date.

Panel A. Dividends per share.



Panel B. The second and third decimal in dividends per share.

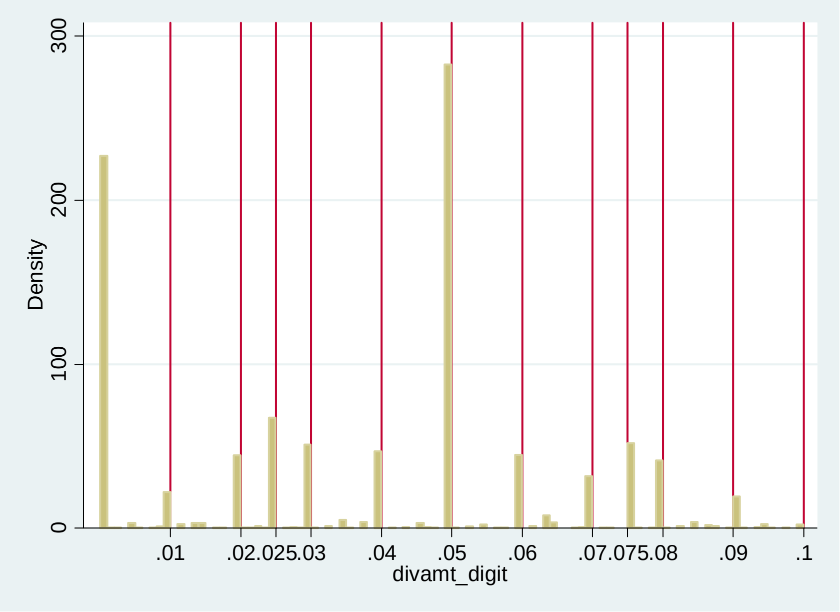
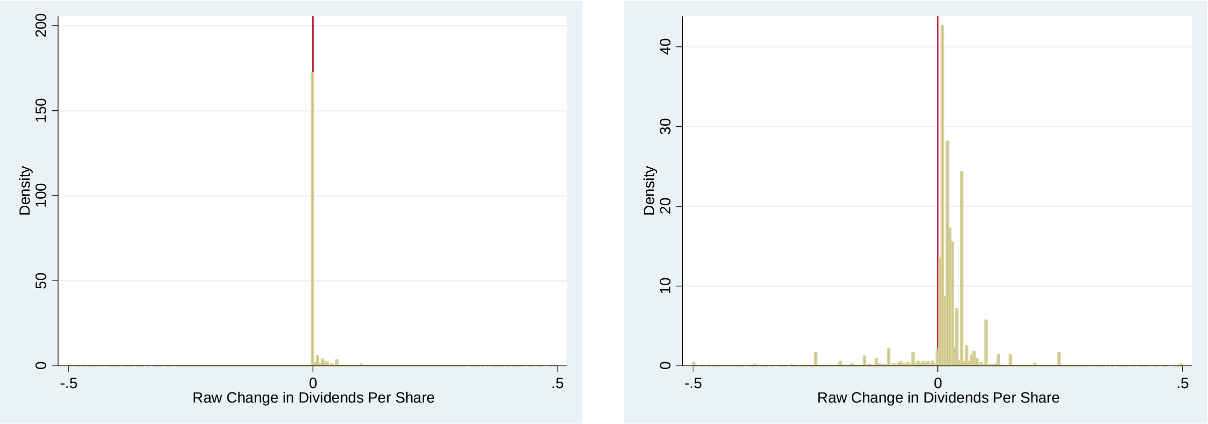
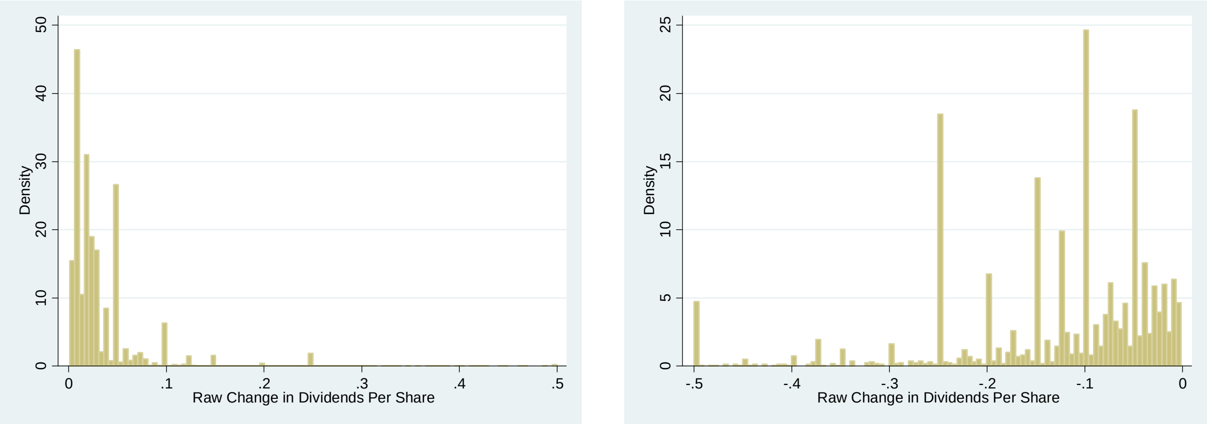


Figure 2. Changes in dividends per share. Histogram of changes in dividends per share. Panel A shows the distribution of changes in dividends per share, Panel B shows the distribution of changes in dividends per share, excluding zero, Panel C shows the distribution of dividend per share increases, and Panel D shows the distribution of dividend per share decreases. The sample from Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.



Panel A. Changes in dividends

Panel B. Excluding zero

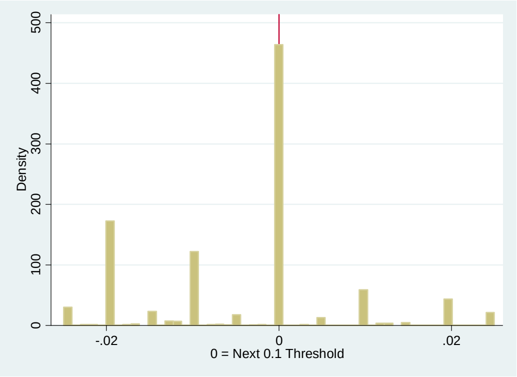


Panel C. Increases in dividends

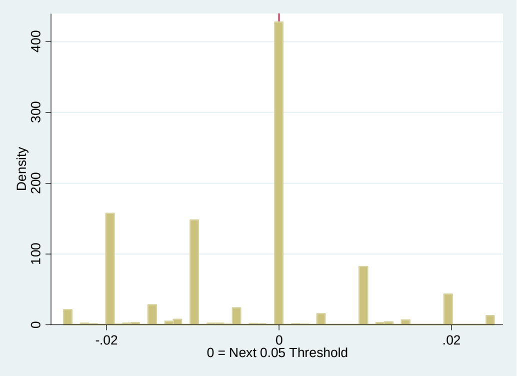
Panel D. Decreases in dividends

Figure 3. Reaching thresholds in dividends per share. Histogram of changes in dividends per share with changes centered on the next threshold. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Panel A, B, and C show the distribution of changes centered on the next round number. Panels D, E, and F show the distribution changes centered on the next round number, when the current dividend is not a round number. The sample from Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

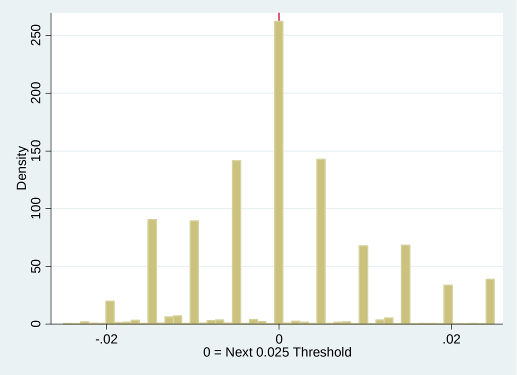
Panel A. Threshold = 0.1



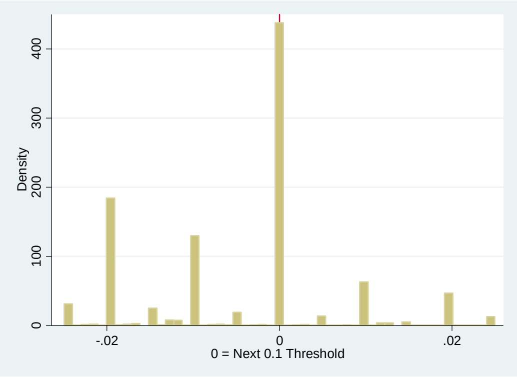
Panel B. Threshold = 0.05



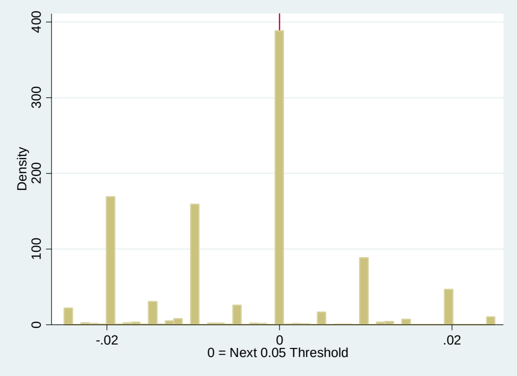
Panel C. Threshold = 0.025



Panel D. Threshold = 0.1



Panel E. Threshold = 0.05



Panel F. Threshold = 0.025

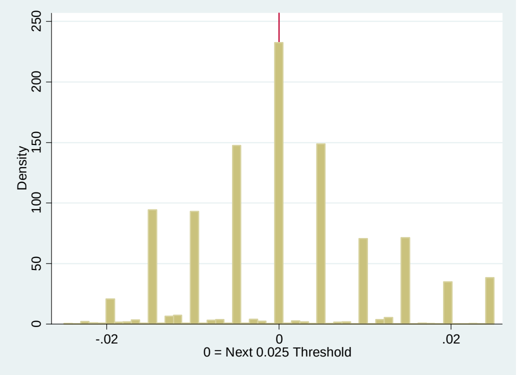
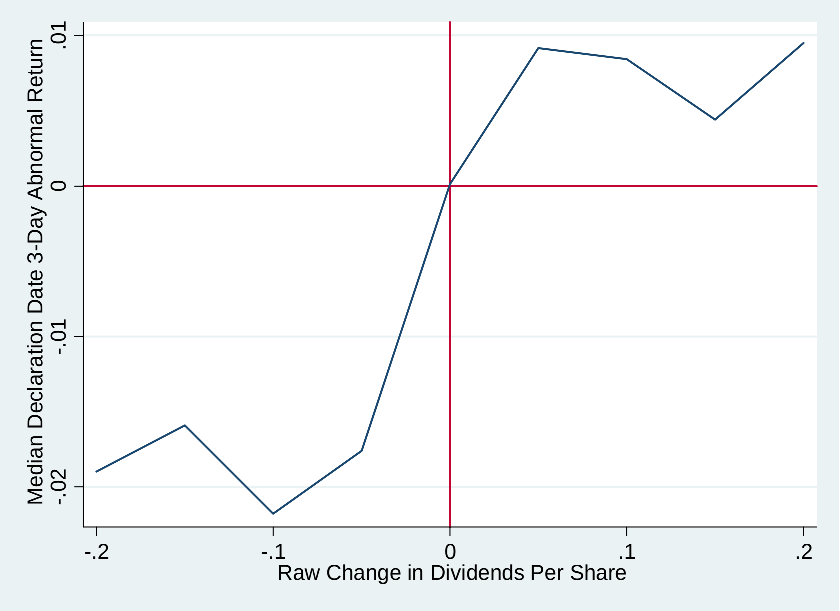


Figure 4. Market reaction to changes in dividends per share. Average 3-day abnormal return by change in dividends per share. Panel A groups changes in dividends per share into groups of 0.05, while Panel B groups changes in dividends per share into groups of 0.025. The groups are formed by rounding the changes in dividends per share down to the nearest threshold. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

Panel A. Changes in dividends per share are grouped to the nearest 0.05



Panel B. Changes in dividends per share are grouped to the nearest 0.025

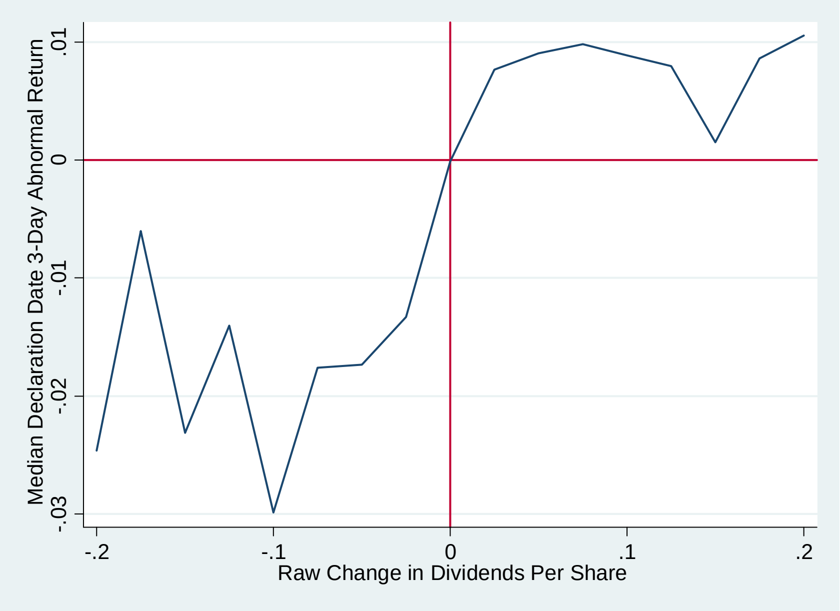
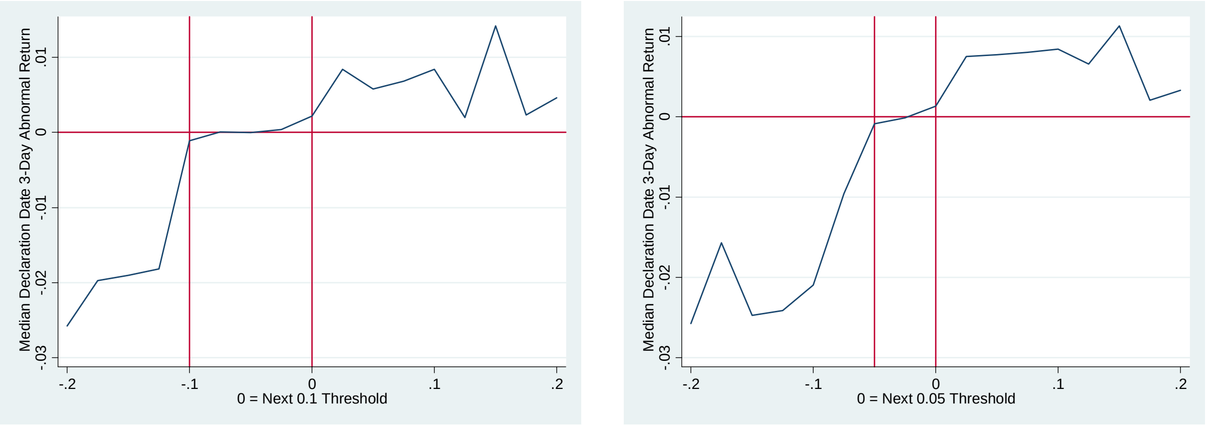


Figure 5. Market reaction to threshold changes in dividends per share. Average 3-day abnormal return by change in dividends per share relative to the next threshold. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Panel A shows changes in dividends per share centered around the next 0.1 threshold, Panel B shows changes in dividends per share centered around the next 0.05 threshold, and Panel C shows changes in dividends per share centered around the next 0.025 threshold. Announcement returns between the two lines correspond to changes in dividends that do not cross a threshold from above or below. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.


Panel A. Threshold = 0.1

Panel B. Threshold = 0.05

Panel C. Threshold = 0.025

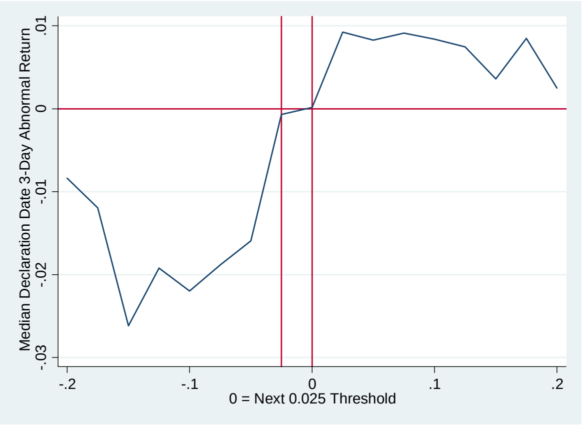


Figure 6. Market reaction following dividend per share streaks. Average 3-day abnormal return by change in dividends per share. Changes in dividends per share are sorted into groups of 0.05, by rounding the raw dividend per share down to the nearest threshold. We partition the sample into situations 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 of 4 or Less), and where the dividend was not changed in the previous four periods (Streak of More Than 4). 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

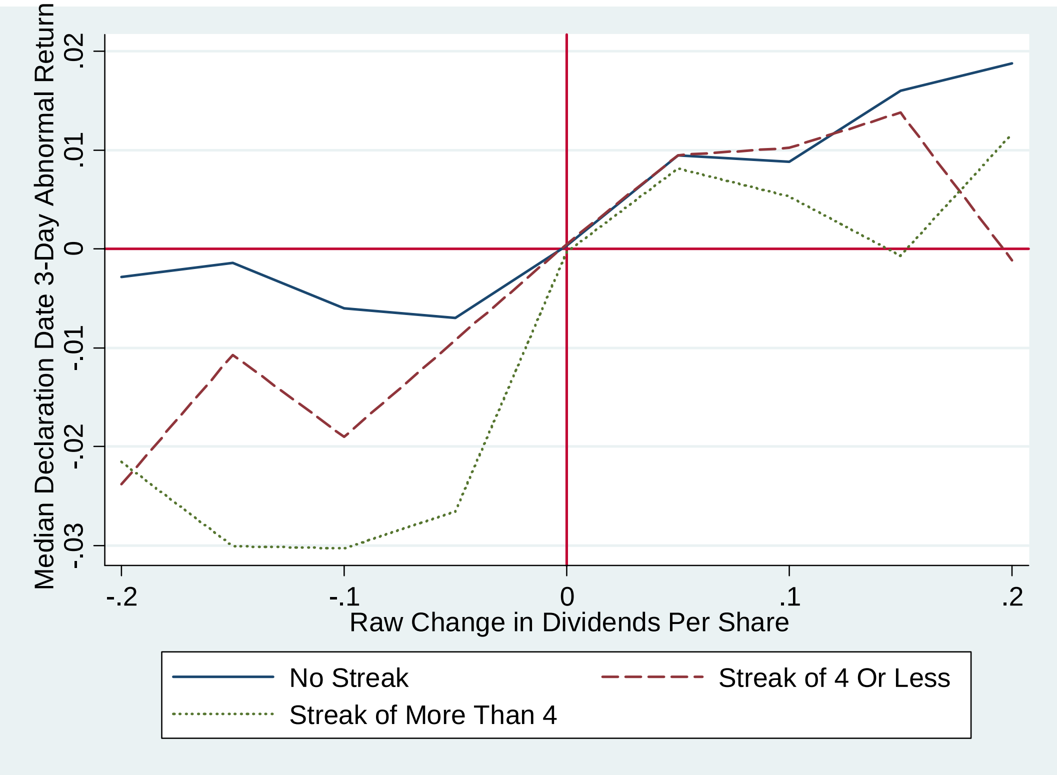


Figure 7. BP-Amoco dividend policy. Split-adjusted dividends per share for BP, Amoco, and the merged company. BP and Amoco merged in December of 1998 forming BP-Amoco. Panels A and C show dividend levels and changes prior to the merger. Panels B and D show dividends after the merger.

Panel A. Pre-Merger Dividend Levels

Panel B. Post-Merger Dividend Levels

Panel A. Pre-Merger Dividend Changes

Panel B. Post-Merger Dividend Changes

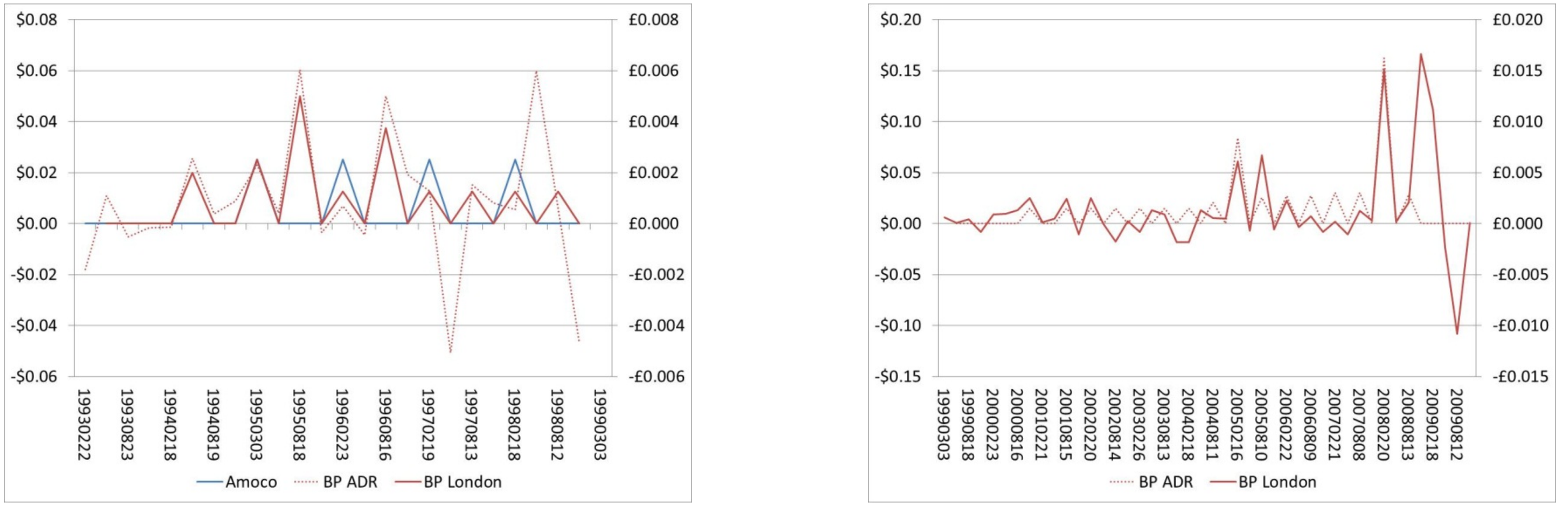
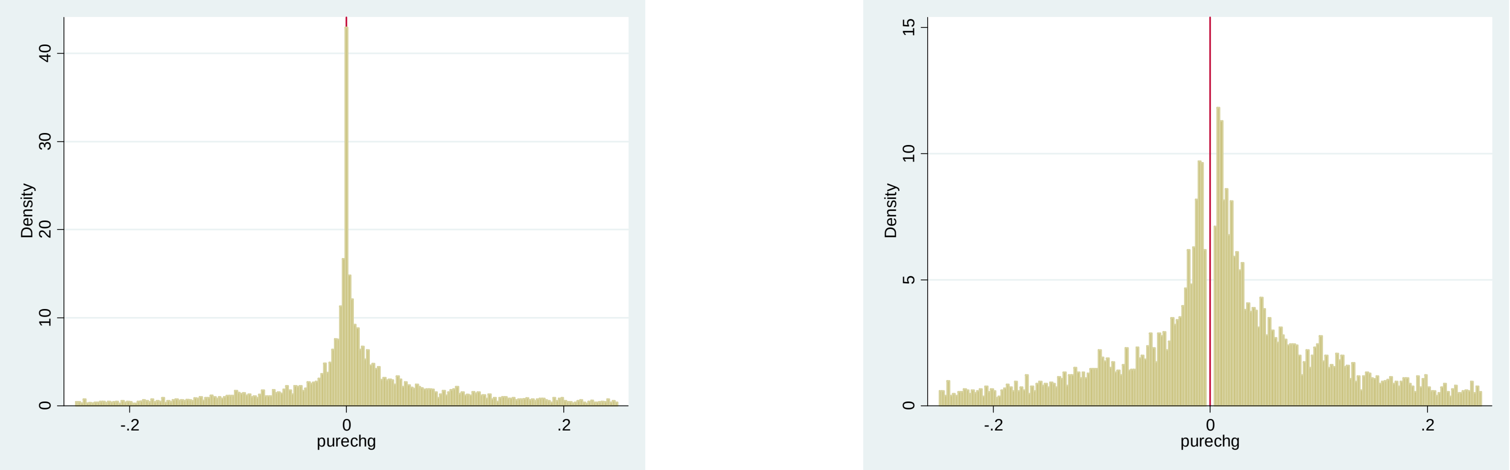
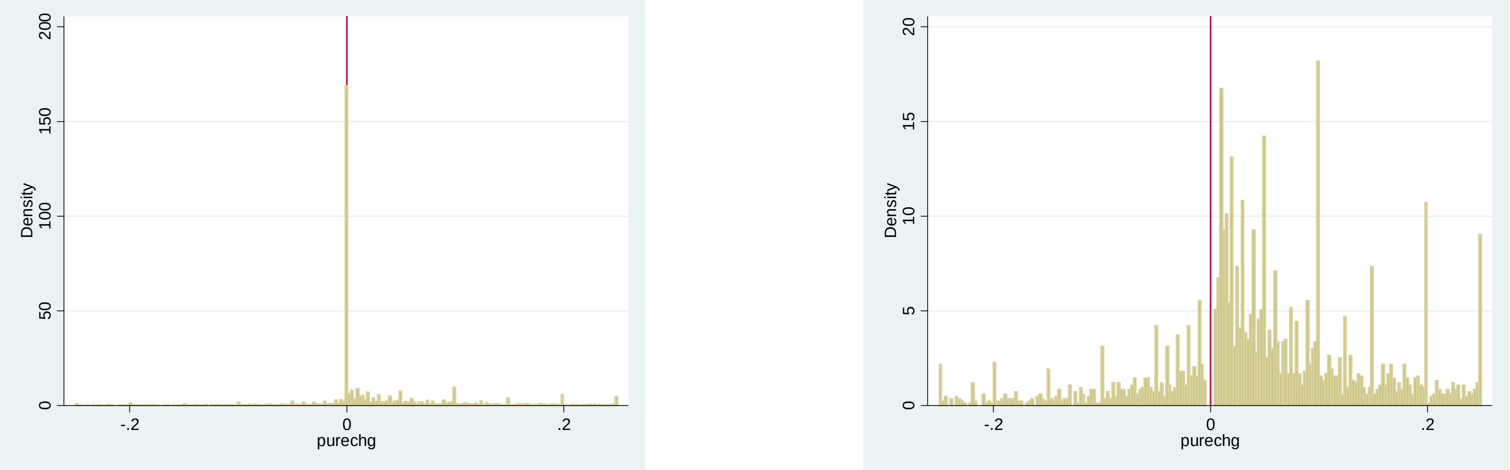


Figure 8. Changes in dividends per share. Histogram of changes in dividends per share. Simple changes in consecutive dividends per share (DD) from Datastream for the following Datastream types: QTR, HYR, YR, FIN, INT. Panels A and C shows the distribution of changes in dividends per share, and Panels B and D show the distribution of changes in dividends per share, excluding zero. Panels A and B show ADRs, and Panels C and D show parent companies.


Panel A. ADRs, Changes in Dividends Per Share

Panel B. ADRs, Changes in Dividends Per Share, Excluding Zero


Panel A. Parents, Changes in Dividends Per Share

Panel B. Parents, Changes in Dividends Per Share, Excluding Zero

Table 1. Summary statistics. The sample includes all records from the CRSP event file with a Distribution Codes (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 US, Americus Trust Components, closed-end funds, and REITs) and nonmissing data on Dividends Per Share and the declaration date. Second and Third Digit are the second and third digits after the decimal place in Dividends Per Share. Change in Dividends Per Share is only computed over three-month windows with no stock splits. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Dividends Per Share are measured relative to a threshold rather than lagged Dividends Per Share. Constant Dividend Streak is the number of past periods where Dividends Per Share remained unchanged. 3-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.

	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
Second and Third Digit	391,865	0.038	0.045	0.028	0.000	0.010	0.060	0.085
Change in Dividends Per Share	250,123	0.003	0.000	0.031	0.000	0.000	0.000	0.030
Change to Next 0.1 Threshold	250,123	-0.059	-0.050	0.042	-0.100	-0.090	-0.040	-0.010
Change to Next 0.05 Threshold	250,123	-0.032	-0.032	0.035	-0.050	-0.050	-0.020	0.000
Change to Next 0.025 Threshold	250,123	-0.015	-0.020	0.032	-0.025	-0.025	-0.005	0.015
Constant Dividend Streak	383,809	6.670	4.000	8.264	1.000	2.000	8.000	23.000
3-Day Announcement Return	249,498	0.002	0.000	0.039	-0.052	-0.017	0.019	0.062

Table 2. Market reaction to changes in dividends per share. Piecewise linear regressions of average 3-day abnormal return on change in dividends per share. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

	1		*2*		*3*	
	Coef	**[T-stat]**	**Coef**	**[T-stat] (p-val)**	**Coef**	**[T-stat] (p-val)**
b	9.89	[19.52]				
$b_{-\infty,-0.2}$			-3.49	[-2.30]	-3.50	[-2.31]
$b_{-0.2,-0.1}$			-7.21	[-1.58]	-7.04	[-1.54]
$b_{-0.1,-0.05}$			19.80	[2.16]	17.10	[1.98]
$b_{-0.05,-0.025}$			38.51	[1.55]	57.33	[9.43]
$b_{-0.025,0}$			71.05	[3.69]		
$b_{0,\ 0.025}$			35.28	[22.82]	27.44	[34.35]
$b_{0.025,\ 0.05}$			12.41	[4.81]		
$b_{0.05,\ 0.1}$			5.84	[2.43]	1.40	[0.61]
$b_{0.1,\ 0.2}$			-7.62	[-3.23]	-7.16	[-3.04]
$b_{0.2,\ \infty}$			0.61	[0.51]	0.59	[0.49]
N		249,498		249,498		249,498
R^2		0.0058		0.0127		0.0125
$b_{-0.1,0}-b_{0,\ 0.1}$			75.84	(0.000)	45.60	(0.000)

Table 3. Market reaction to changes in dividends per share: Abnormal Volume. Piecewise linear regressions of average 3-day abnormal return on change in dividends per share. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

	1		*2*		*3*	
	Coef	**[T-stat]**	**Coef**	**[T-stat] (p-val)**	**Coef**	**[T-stat] (p-val)**
b	31.81	[5.11]				
$b_{-\infty,-0.2}$			77.45	[3.70]	77.48	[3.70]
$b_{-0.2,-0.1}$			138.95	[2.43]	138.40	[2.42]
$b_{-0.1,-0.05}$			-363.84	[-3.17]	-355.33	[-3.29]
$b_{-0.05,-0.025}$			-561.95	[-1.66]	-620.92	[-8.32]
$b_{-0.025,0}$			-666.94	[-2.47]		
$b_{0, 0.025}$			577.04	[16.19]	513.11	[28.40]
$b_{0.025, 0.05}$			390.53	[6.67]		
$b_{0.05, 0.1}$			78.59	[1.66]	42.31	[0.95]
$b_{0.1, 0.2}$			-71.60	[-1.55]	-67.88	[-1.47]
$b_{0.2, \infty}$			-11.75	[-0.60]	-11.97	[-0.61]
N		219,835		219,835		219,835
R^2		0.0001		0.0064		0.0064
$b_{-0.1,0}-b_{0, 0.1}$			546.58	(0.000)	420.82	(0.000)

Table 4. Market reaction to changes in dividends per share: Clearing a threshold. Piecewise linear regressions of average 3-day abnormal return on change in dividends per share relative to the next threshold. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Column 2 shows changes in dividends per share centered on the next 0.025 threshold, and Column 3 shows changes in dividends per share centered around the next 0.05 threshold. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

	1		*2*		*3*	
	Coef	**[T-stat]**	**Coef**	**[T-stat] (p-val)**	**Coef**	**[T-stat] (p-val)**
b	9.28	[20.21]				
$b_{-\infty,-0.2}$			-4.52	[-3.06]	-4.93	[-3.58]
$b_{-0.2,-0.1}$			-1.21	[-0.24]	1.36	[0.31]
$b_{-0.1,-0.05}$			18.60	[1.65]	66.30	[11.79]
$b_{-0.05,-0.025}$			99.87	[6.70]	5.93	[13.59]
$b_{-0.025,0}$			7.49	[8.95]		
$b_{0,\ 0.025}$			39.57	[21.10]	22.08	[15.04]
$b_{0.025,\ 0.05}$			-0.68	[-0.18]		
$b_{0.05,\ 0.1}$			7.10	[2.19]	-0.15	[-0.05]
$b_{0.1,\ 0.2}$			-8.46	[-3.59]	-7.62	[-3.25]
$b_{0.2,\ \infty}$			1.02	[0.85]	1.33	[1.13]
N		249,498		249,498		249,498
R^2		0.0056		0.0116		0.0095
$b_{-0.05,-0.0025}-b_{0,\ 0.0025}$			60.30	(0.000)		
$b_{-0.1,-0.05}-b_{0,\ 0.05}$					44.22	(0.000)

Table 5. Market reaction following dividend per share streaks. Average 3-day abnormal return by change in 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 of 4 or Less), and where the dividend was not changed in the previous four periods (Streak of More Than 4). T-stats and p-values are on differences from the No Streak sample. 3-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 Figure 1 and Table 1 is further limited to changes over a three-month window with no stock splits.

	No Streak		*Streak<=4*		*Streak>4*	
	Coef	**[T-stat]**	**Coef**	**[Diff T-stat]**	**Coef**	**[Diff T-stat]**
$b_{-\infty,-0.2}$	-6.13	[-3.03]	-3.32	[1.09]	-1.41	[2.30]
$b_{-0.2,-0.1}$	14.66	[1.68]	-4.79	[-1.83]	-14.43	[-3.33]
$b_{-0.1,-0.05}$	-1.26	[-0.06]	38.22	[1.39]	9.13	[0.41]
$b_{-0.05,-0.025}$	9.72	[0.20]	-35.71	[-0.72]	72.79	[0.94]
$b_{-0.025,0}$	25.03	[0.77]	89.89	[1.51]	93.01	[1.36]
$b_{0,\ 0.025}$	46.22	[8.00]	28.67	[-2.90]	52.43	[0.92]
$b_{0.025,\ 0.05}$	8.36	[0.79]	14.58	[0.57]	3.71	[-0.40]
$b_{0.05,\ 0.1}$	9.98	[0.90]	12.01	[0.18]	-6.73	[-1.42]
$b_{0.1,\ 0.2}$	-8.79	[-0.97]	-9.15	[-0.04]	-4.48	[0.44]
$b_{0.2,\ \infty}$	1.59	[0.58]	-1.60	[-1.10]	2.48	[0.25]
N						249,498
R^2						0.0139
$b_{-0.1,0}-b_{0,\ 0.1}$	-31.06	(0.156)				
$b_{-0.1,0}-b_{0,\ 0.1}$ - No Streak $b_{-0.1,0}-b_{0,\ 0.1}$			68.21	(0.015)	156.58	(0.000)

Table 6. Summary statistics. The sample includes all ADRs from Datastream and their matched parents with nonmissing data on dividends per share. The sample includes all ADRs from Datastream with a trading history between 1990 and 2009 on a US exchange in US$. 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 5-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 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.

	N	**Mean**	**Median**	**SD**	**5**	**25**	**75**	**95**
Panel A. US ADRs								
Dividends Per Share, Split Adjusted	19,046	0.386	0.194	1.897	0.0	0.1	0.4	1.2
Change in Dividends Per Share	16,203	0.015	0.002	0.836	-0.3	0.0	0.1	0.4
5-Day Announcement Return	8,472	0.011	0.000	0.247	-0.1	0.0	0.0	0.1
Panel B. Parents								
Dividends Per Share, Split Adjusted	32,177	41.925	1.200	873.329	0.0	0.2	5.5	33.7
Change in Dividends Per Share	9,196	-4.914	0.005	1948.531	-2.3	0.0	0.2	5.0
Change in Dividends Per Share, All	29,211	0.095	0.000	1137.445	-4.4	0.0	0.2	5.7

Table 7. Market reaction to changes in dividends per share for ADRs. Piecewise linear regressions of average 5-day return on clean changes in dividends per share. The sample is described in Table 6. We compute 5-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 use the Datastream return index (RI) up to five weekdays surrounding the announcement date, with no adjustment for market movements.

	ADR			
	Coef	**[T-stat] (p-val)**	**Coef**	**[T-stat] (p-val)**
$b_{-\infty,-0.2}$	-0.05	[-0.12]	-0.05	[-0.13]
$b_{-0.2,-0.1}$	18.63	[1.17]	19.78	[1.25]
$b_{-0.1,-0.05}$	-7.69	[-0.17]	-20.08	[-0.55]
$b_{-0.05,-0.025}$	-45.81	[-0.58]	5.42	[0.19]
$b_{-0.025,0}$	47.05	[1.04]		
$b_{0, 0.025}$	-40.71	[-0.68]	-7.01	[-0.31]
$b_{0.025, 0.05}$	30.99	[0.83]		
$b_{0.05, 0.1}$	4.82	[0.18]	14.47	[0.50]
$b_{0.1, 0.2}$	-5.32	[-0.35]	-6.42	[-0.43]
$b_{0.2, \infty}$	0.09	[0.50]	0.10	[0.51]
N		8,472		8,472
R^2		0.0004		0.0003
$b_{-0.1,0}-b_{0, 0.1}$	-1.55	(0.985)	-22.12	(0.638)