Dividends as Reference Points: A Behavioral Signaling Approach*

Malcolm Baker
Harvard Business School and NBER

Jeffrey Wurgler
NYU Stern School of Business and NBER

July 15, 2012

Abstract

We outline a dividend signaling approach in which rational managers signal firm strength to investors who are loss averse to reductions in dividends relative to the reference point set by prior dividends. Managers with strong but unobservable cash earnings separate themselves by paying high dividends but retain enough earnings to be likely not to fall short of the same level next period. The model is consistent with several features of the data, including equilibrium dividend policies similar to a Lintner partial-adjustment model; modal dividend changes of zero; stronger market reactions to dividend cuts than increases; relative infrequency and irregularity of repurchases versus dividends; and a core mechanism that does not center on public destruction of value, a notion that managers reject in surveys. Supportive new tests involve nominal levels and changes of dividends per share, announcement effects, and reference point currencies of ADR dividends.

* For helpful comments we thank Christa Bouwman, Harry DeAngelo, Wayne Ferson, Xavier Gabaix, Cam Harvey, David Hirshleifer, Steven Huddart, Anthony Lynch, Roni Michaely, Holger Mueller, Erin Smith, and seminar participants at the American Finance Association annual meeting, Dartmouth, the International Monetary Fund, the Jackson Hole Finance Conference, NYU Stern, the SFS Cavalcade at the University of Virginia, Tilburg University, the University of Amsterdam, the University of Miami, the University of Michigan, the University of Southern California, the University of Wisconsin, and Wharton. We thank Chris Allen for excellent research assistance. Baker gratefully acknowledges the Division of Research of the Harvard Business School for financial support.

I. Introduction

Managers share a number of common views about their dividend policies, as shown in the survey by Brav, Graham, Harvey, and Michaely (2005). They strive to avoid reducing dividends per share (of the 384 managers surveyed, 93.8% agreed); they try to maintain a smooth dividend stream (89.6%); and, they are reluctant to make changes that might have to be reversed (77.9%). They follow such policies because they believe that there are negative consequences to reducing dividends (88.1%), which they believe convey information to investors (80%). While caution is merited in interpreting any survey responses, the Brav et al. results are further consistent with Lintner's (1956) own survey and interviews, his partial-adjustment model, and a large empirical literature demonstrating a significant response to dividend announcements.

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

In this paper we use prospect theory of Kahneman and Tversky (1979) to motivate a signaling model of dividend policy with behavioral foundations. We focus on two features of the prospect theory value function. We use the concept of reference-dependence: values and perceptions are based on losses and gains relative to a reference point. We also assume loss aversion: there is a kink at the reference point whereby marginal utility is discontinuously lower in the domain of losses. Reference-dependence and loss aversion are supported by a considerable literature in psychology, finance and economics, as we briefly review later.

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

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

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

The approach suggests some new tests as well. They revolve around the idea that a dividend that creates a reference point forms a powerful signal. Psychological evidence, as well as casual introspection, shows that memory and salience play a role in the formation of reference points. One potential empirical manifestation is that dividends per share will be concentrated in round numbers, which are more memorable and deviations from which are thus more noticeable. (Static models, by contrast, are continuous ascribe no particular role to round number dividends

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

and cannot treat changes at all.) Another possible empirical manifestation is that repetition of a particular dividend level also ingrains a reference point. These are not direct predictions of the model, because we do not model memory and salience, but rather are predictions of the general intuitions of signaling around a moveable reference point.

In empirical tests we confirm that dividend levels and changes are typically made in round numbers, e.g. 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 (a new result). This asymmetry is more pronounced when the same dividend per share has been paid for several consecutive quarters. As a placebo test, we examine ADRs, where we find that the reference point and round numbers are denominated in foreign currency. In this sample, there is no clustering and nothing special about the market's reaction around zero dividend changes in U.S. dollars. Investors have settled on a reference point currency even though the dividend's economic value to some investors, and perhaps the firm's ability to pay that value, fluctuates with the exchange rate.

The approach complements other papers that connect dividends and reference points or prospect theory. In particular, Shefrin and Statman (1984) argue that dividends improve the utility of investors with prospect theory value functions if they also mentally account (Thaler (1999)) for dividends and capital gains and losses separately. This approach is however very different in that dividends serve no signaling purpose, rely on curvature in the value function, and have no particular relationship to a Lintner policy. Another recent contribution is Lambrecht and Myers (2012). In their model, managers maximize the present value of the utility of rents that they can extract from profits. They smooth dividends and prefer a Lintner policy because they have habit formation preferences and rents move in lockstep with dividends given the budget constraint. We discuss these papers and other related research. More generally, our paper adds to the literature on behavioral corporate finance surveyed by Baker and Wurgler (2012).

Section II reviews the relevant literature on reference-dependent utility. Section III describes the model. Section IV discusses its compatibility with known facts of dividend policy. Section V tests 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. We first review this literature more generally, then apply it to dividends.

A. Reference-Dependent Utility and Loss Aversion

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

Tversky and Kahneman (1991) review the classic literature on loss aversion. Kahneman and Tversky (1979) introduced loss aversion to reflect then-known patterns in choice behavior. The subsequent literature suggests its relevance in a wide range of applications. One implication of loss aversion is what Thaler (1980) termed the endowment effect. Kahneman, Knetch, and

Thaler (1990) found that the value of an item increases when it is considered already in one's endowment. A literature has developed on differences between the willingness to pay for a small improvement versus willingness to accept a small loss, another reflection of loss aversion. (These literatures suggest the ballpark figure that losses matter slightly more than twice as much as gains, a figure that we employ in our numerical example.) Finally, a related phenomenon is status quo bias. Samuelson and Zeckhauser (1988) documented a preference for the status quo even when costs of change are small relative to potential benefits, such as in choices about medical plans.

B. Reference Points

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

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

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

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

Past circumstances can also supply powerful reference points. Genesove and Mayer (2001) find that people resist selling their homes below 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 idea of one's prior consumption as a reference point for the utility of current consumption is represented through internal habit formation preferences as in Constantinides (1990).

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

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

C. Past Dividends as Reference Points

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

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

III. A Model of Signaling With Dividends as Reference Points

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

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

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

A. Setup

The model has two periods: $t = 1$ and 2. There are two players: a benevolent manager and an investor with reference dependent preferences. In the first period, the investor arrives with an exogenous reference point for dividends d_0. In some ways, this is a single snapshot in a multiperiod model. While the inherited reference point could in principle be endogenized, we believe the technical costs would be large compared to the benefits in extra realism or intuition. The manager also receives private information about cash earnings ε_1 and pays a dividend d_1 in the first period. This dividend forms a new reference point for the second, liquidating dividend d_2. Today's dividend d_1 relative to d_0 tells the investor something about the manager's private information and hence the value of the firm. In the second period, the manager simply pays d_2. There is no discounting.

Manager utility. The manager cares about what the investor thinks about ε_1 today. He also cares about the investor's long run utility. This objective function is in the spirit of standard signaling models like Leland and Pyle (1977), Miller and Rock (1985), or Stein (1989), which use weighted averages of the dividend-adjusted stock market price and the investor's long-run utility. Our simplified objective function is:

$$E_m\left[\lambda E_i\left[\varepsilon_1\right]+\left(1-\lambda\right)u\left(d_1,d_2 \mid d_0\right)\right], \tag{1}$$

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

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

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

Put simply, the investor cares about fundamental value, or total dividend payments, but with a twist. The level of the reference point comes from historical firm dividend policy, and b is greater than zero so that the marginal utility of dividends falls discontinuously at the reference

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

point; the pain of coming up short is high. This utility function is in the spirit of prospect theory with a kink at a reference point. We leave out the complexity of curvature. The second-period reference point equals first-period dividends d_1 by assumption. In reality, the reference point and the intensity of the reference point b may be determined by a long history of levels and changes in dividend policy. The fact that each dividend payment forms a separate reference point also requires a practice of narrow framing. This is not a reference point applied to total ending wealth, but rather a reference point applied much more narrowly both across stocks and time, in the spirit of Barberis, Huang, and Thaler (2006).

Information. In the absence of a signaling motive, the first-best would have a benevolent manager saving any resources above the first-period reference point, both to lower the reference point created for the second period and to save resources to meet this lower reference point in the event of low second-period earnings. We now introduce the signaling problem. For simplicity, the manager has no control over the cash earnings of the firm. Note that this is a bit different from a traditional signaling model where the manager must destroy firm value to impress the capital markets. There is also no fundamental agency problem as there is in Lambrecht and Myers (2012). The manager is not able to keep the cash for himself, and no firm value is intentionally created or destroyed with dividend policy.[3] 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,

$$\varepsilon_1 + \varepsilon_2 . \tag{3}$$

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

Think of these as cash earnings that are not observable to the investor. This is obviously an extreme assumption of asymmetric information. It is worth noting the key elements of the assumption, which might each seem more reasonable. First, the manager must have some informational advantage in learning ε_1 over the investor. Otherwise, there is no signaling problem. Second, the payment of the observable dividend must form the investor's reference point, not the firm's reported financials, such as earnings per share or cash balances. Otherwise, the manager has no lever to signal his information about ε_1. For simplicity, we assume that the second-period cash earnings are have a uniform distribution, $\varepsilon_2 \sim U[0,2]$.

We have considered extensions of the model where the source of the information asymmetry is over ε_2, a quantity that would not appear explicitly in any financial statements. This assumption produces similar results, although the effects of the budget constraint described in the next paragraph can change. The simpler model that we examine here has a mechanical link between type in terms of firm quality and current resources.

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

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

These follow from the assumptions of no new financing and a benevolent manager.

B. Equilibrium

There are three effects that appear in the manager's objective function in Eq. (1). First, there is sometimes an advantage to paying out dividends immediately. Consider a first-period dividend below the reference point d_0. Setting aside the effect on the second-period reference point, these dividends will be valued on the margin at b+1 times the payout, instead of simply the payout. This is a net benefit to investor utility in Eq. (2) of bd_1. Above d_0, there is no marginal benefit from merely shifting payout from the second period forward. Second, by increasing the dividend today, the investor's estimate $E_i[\varepsilon_1]$ of the unobservable cash earnings rises through an equilibrium set of beliefs that map dividend policy to cash earnings. This enters into the manager's utility function directly in Eq. (1). Third, increasing the dividend in the first period, for either of these rationales, produces an expected future cost to investor utility that comes from the possibility of falling short of the reference point set for the second period.

Substituting in the budget constraint from Eq. (4), and taking expectations over the $\varepsilon_2 \sim U[0,2]$ distribution, leads to an expected cost conditional on today's dividend of

$$b\left(d_1 - \frac{\varepsilon_1}{2}\right)^2, \tag{5}$$

provided d_1 is more than half of cash earnings. Intuitively, there is no cost if the manager adopts a conservative dividend policy of paying half of cash earnings. The expected cost is quadratic as dividends rise from this point and increasing in the intensity of the reference point b>0.

Combining the three motivations, the manager's utility function from Eq. (1) simplifies to

$$(1-\lambda)b(d_1 - d_0)\mathbf{1}_{d_1<d_0} + \lambda E_i[\varepsilon_1 \mid d_1] - (1-\lambda)b\left(d_1 - \frac{\varepsilon_1}{2}\right)^2 \mathbf{1}_{d_1>\frac{\varepsilon_1}{2}}. \tag{6}$$

The cost of falling short of the initial reference point is relevant only at low levels of first-period dividends, the signaling motive is present for all levels, and the expected cost of falling short of the new reference point is relevant only when first-period savings alone cannot cover it.

Given these considerations, there are three ranges of dividend policies in equilibrium. There is a high payout ratio for firms with the extra motivation to clear the initial reference point of d_0. Next, managers cluster at d_0 once this motivation drops out of Eq. (6). Finally, there is a lower payout ratio for firms well above the initial reference point, who nonetheless pay higher dividends to separate themselves from each other and from the pool at d_0. The form of the partial pooling equilibrium is not dissimilar to that in Guttman et al. (2010), though the nature of the signaling mechanism is quite different and perhaps more importantly, in our approach it is clear why dividends pool at the prior period's level as opposed to some arbitrary level.

Proposition 1. *There exists an equilibrium dividend policy in which*

$$d_1 = \varepsilon_1 \text{ for } \varepsilon_1 < d_0$$

$$d_1 = d_0 \text{ for } d_0 < \varepsilon_1 < \varepsilon^*$$

$$d_1 = \tfrac{1}{2}\varepsilon_1 + \tfrac{\lambda}{1-\lambda}\cdot\tfrac{1}{b} \text{ for } \varepsilon_1 > \varepsilon^*$$

with associated equilibrium beliefs of:

$$E_i[\varepsilon_1 \mid d_1] = d_1 \text{ for } d_1 < d_0$$

$$E_i[\varepsilon_1 \mid d_1] = \tfrac{1}{2}\left(d_0 + \varepsilon^*\right) \text{ for } d_1 = d_0$$

$$E_i[\varepsilon_1 \mid d_1] = 2\left(d_1 - \tfrac{\lambda}{1-\lambda}\cdot\tfrac{1}{b}\right) \text{ for } d_1 > \tfrac{1}{2}\varepsilon^* + \tfrac{\lambda}{1-\lambda}\cdot\tfrac{1}{b}.$$

where ε^* solves:

$$\lambda\tfrac{1}{2}\left(\varepsilon^* - d_0\right) - (1-\lambda)b\left(\left(\tfrac{\lambda}{1-\lambda}\cdot\tfrac{1}{b}\right)^2 - (d_0 - \tfrac{1}{2}\varepsilon^*)^2\right) = 0.$$

Incentive compatibility here requires a manager with $\varepsilon_1 < d_0$ to be willing to pay $d_1 = \varepsilon_1$. This is not essential. There are other slightly more complicated and perhaps more realistic equilibria, where there is a discontinuous drop in dividends just to the left of d_0 as well as to the right. These equilibria also allow for lower levels of b. More importantly, a manager with $\varepsilon_1 = \varepsilon^*$

must be indifferent between paying d_0 and paying $\frac{1}{2}\varepsilon^* + \frac{\lambda}{1-\lambda}\cdot\frac{1}{b}$. For this to hold, the signaling benefit of shifting the investor's expectations from $\frac{1}{2}\left(d_0 + \varepsilon^*\right)$ to ε^* must equal the cost differential of evaluating Eq. (5) at d_0 and at $\frac{1}{2}\varepsilon^* + \frac{\lambda}{1-\lambda}\cdot\frac{1}{b}$. This leads to the simple quadratic equation in the proposition. There is no claim of uniqueness, so a sufficient proof and simplest illustration of Proposition 1 is a numerical example.

Example. Suppose the manager cares equally about stock price and utility, i.e. $\lambda = 0.5$. If d_0 is 1 and b is 2.5, then Proposition 1 indicates that the equilibrium cutoff ε^* is 1.6. For ε_1 above 1.6, the first-period dividend is $\frac{1}{2}\varepsilon_1 + 0.4$. This exactly trades off the marginal signaling benefit per unit of dividends of 2.0, using investor beliefs implicit in Proposition 1, against the second-period marginal cost, i.e. the first derivative of Eq. (5), of $5\cdot\left(d_1 - \frac{1}{2}\varepsilon_1\right) = 5\cdot\left(0.4\right) = 2.0$. For ε_1 between 1.0 and 1.6, the first-period dividend is simply d_0, or 1.0. At ε_1 equal to 1.6, where there should be indifference, the signaling benefit of separating from this pool is 1.6 minus the average of $\varepsilon^* = 1.6$ and $d_0 = 1.0$, which is 0.3. The cost from Eq. (5) is $2.5 \cdot (0.4^2 - 0.2^2) = 0.3$. This cost is decreasing in ε_1, so there is no incentive for any of the managers clustered at $d_1 = 1.0$ to raise the first-period dividend. For ε_1 below 1.0, the first-period dividend is ε_1. Here, the manager is limited by the budget constraint. The marginal first-period benefit per unit of dividends of 2.5 plus the marginal signaling benefit per unit of dividends of 1.0, using investor beliefs implicit in Proposition 1, totals 3.5. This exceeds the second-period marginal cost just below d_0 of $5\cdot\left(1 - 0.5\right) = 2.5$. So, dividends are at a corner solution of $d_1 = \varepsilon_1$ from Eq. (4).

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

Another way to see this is by plotting the histogram of dividend changes in Panels B and C of Figure 1. We assume that first-period cash earnings ε_1 are normally distributed,[4] with a mean of $\frac{1}{2}\left(d_0 + \varepsilon^*\right)$. There is a spike in the distribution at the reference point in Panel B. Moreover, even when we remove this spike in Panel C, there is still a jump in the distribution moving from the range just to the left of the reference point to the range just to the right. The distribution of dividend changes otherwise has a lower and longer tail of larger dividend cuts to the left of the reference point.

Finally, we plot the market reaction to dividend announcements in Panel D of Figure 1. This is measured as the percentage change in expected utility in Eq. (1) from before the announcement. The interesting behavior is in the narrow range around the reference point. The drop in utility per unit of dividends is steeper to the left of the reference point than to the right. Missing the reference point by just a tiny amount leads to a drop of $\frac{1}{2}\left(\varepsilon^* - d_0\right)$ in the investor's expectation. But it takes a discontinuous increase in dividends of $\frac{1}{2}\varepsilon^* + \frac{\lambda}{1-\lambda} \cdot \frac{1}{b} - d_0$ to achieve the equivalent increase in expectations. As a result, there is a kink in the reaction at exactly d_0.

[4] This is for purposes of illustration, but a sum of uniform variables converges quickly to a normal distribution, so if one conceives ε_1 as such a sum then the relationship to the uniform distribution of ε_2 in the model is not entirely unnatural, particularly in the dividend initiation case.

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

Proposition 2. *In the equilibrium described in Proposition 1, ε* and the market reaction to $d_1 < d_0$ are increasing in b.*

Put another way, Proposition 2 says that there is more clustering of dividends at the reference point d_0 as the intensity of reference point preferences increases. As a result, the market reacts more negatively to a narrow miss. More information is revealed.

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

C. Reference dependent behavior

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

IV. The Model and Prior Evidence

Dividend policy is so awash with empirical facts that any new model could be assessed almost as much on its ability to fit those facts as on the success of novel predictions. We consider several stylized facts here in light of the model. While it certainly cannot explain all of the facts, a model of signaling with reference points captures some 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). To keep the discussion finite we will not make broad comparisons to models based on agency problems, catering motives, clientele effects, or other non-signaling issues.

A. Surveys

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

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

While standard signaling theories also predict that lower dividends are associated with lower market values, executives reject them as based on unrealistic foundations. As noted in the Introduction, executives say that they do not use dividends to show that their firm can withstand the costs and scrutiny associated with raising external capital, or to show that their firm can pass up good projects and still perform well. Only a small minority of executives endorsed signaling through taxes; Brav et al. 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).[5] Note that our model doesn't explicitly rely on public destruction of value to create a credible signal, so in that sense it would not conflict with this aspect of the evidence.

Finally, it is notable that standard signaling theories do not naturally focus on dividends per share. But it is natural to do so in the reference point context. Moreover, standard signaling

[5] In a signaling model we would interpret "competitors" as firms of different "types."

theories predict a continuous market reaction. There is nothing special about stability or the historical level of payouts, such that falling short of this level would produce a discontinuous reaction. Dividend policy is defined in more "economic" terms in standard models, such as dividend yield or payout ratio, which are less salient to 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

Miller (1986) suggested that the Lintner (1956) model was a behavioral model, as opposed to originating from a traditional maximization problem. Given reasonable parameter values, our model does have dividends following a partial-adjustment policy; more generally, they are smoothed relative to earnings (Fama and Babiak (1968)).

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

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

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

In a recent paper, Lambrecht and Myers (2012) provide an alternative motivation for the Lintner model and smoothing. It is not a signaling model so we do not include it in our contrasts with traditional signaling models. We would just argue that there remains room for additional explanations of the Lintner model, in that it is questionable that thousands of large, established, public U.S. firms smooth their payouts because a coalition of habit-formation managers prefer to smooth out their stealing. In any event, Lambrecht and Myers's model makes predictions involving habit persistence, adjustment costs, target payout ratios, investor protection, and managerial impatience. These concepts are basically orthogonal to those of our own model and do not relate to our own tests.

C. Announcement Effects

Even if executives disavow standard signaling models, shareholders clearly 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 Yoon and Starks (1995) and subsequent papers documenting dividend announcement effects.

That dividend cuts would be received especially poorly is a prediction of the model. The main effect is that cutting a dividend, even slightly, is fully revealing and betrays the firm as one that cannot afford even that reference level dividend. In any case, none of the standard signaling models offers a direct explanation for the asymmetry in announcement effects.

D. Repurchases vs. Dividends

As Grullon and Michaely (2002) point out, some of the most popular dividend signaling models, including Bhattacharya (1979), Miller and Rock (1985), and the further investigation of Miller and Rock by Guttman et al. (2010), imply that dividends and repurchases are perfect substitutes. This is not consistent with the fact that repurchases are considerably less frequent and less regular.

The reference point approach suggests a simple reason why dividends are superior signals: There is only one number to remember. But there are at least two parameters of a repurchase program: the purchase price (or price range for a Dutch auction) and the number of shares sought. Of these, only the price is salient to investors, and obviously a firm cannot commit to repurchasing shares at the same constant price quarter after quarter. Moreover, each investor receives a dividend, but it is not required that each investor engage in a repurchase; in fact, only those selling out need to any attention at all. Put simply, repurchase programs are hard to specify in salient and memorable terms, even if they are financially equivalent.

V. New Tests

In addition to explaining known stylized facts at least as well as standard signaling model, our approach suggests some new tests. The core idea is that the power of dividends as a signal is 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. We do not model b, but the intuition of in our model is that firms will design dividends to be salient and easy to recall. Some of our tests below are based on these intuitions. 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 $0.1325 per share. Indeed, for one-period models, there is no need 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. Again, we do not model salience or memorability, so ours are tests of general intuitions surrounding the model that do not apply very clearly to other signaling models. 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 (these are rare, and we wish to avoid skewing computations of nominal changes), 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 hundreds of thousands of observations. The majority of quarterly dividend changes are zero, and only a small fraction is 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 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 3. 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. Aerts, Van Campenhout, and Van Canegham (2008) use Benford's law to show that "0" is an unusually common second digit in dividends, a similar but more limited result. A plot makes the widespread use of round numbers of various types, and specific numbers such as five and ten cents, quite obvious.

Changes are also made in round numbers. Figure 4 shows dividend changes when no split occurs between dividend payments. The most obvious patterns in dividend changes match our model in Panels B and C of Figure 1 exactly: the very large mass at exactly zero, also noted by Guttman et al. (2010) and others; 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, as we observed in Panel C of Figure 1.

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. We view this as somewhat akin to the gap in the distribution in Panel B of Figure 1. To be appreciated for raising the dividend, firms must do so in more than a trivial way.[6]

Figure 5 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 3 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. Apparently, many boards think of communicating dividend policy in an easily recalled dollar per share quantity, rather the alternative of a specific total dollar payout to be divided equally among shareholders.

[6] It is also worth keeping in mind that in practice, even a relatively small increase in the quarterly dividend that does not cross a threshold could amount to a substantial increase, and thus a meaningful signal, when considered in present value terms. In any event, the gap in the distribution in Figure 1 Panel B can be made arbitrarily small with appropriate parameter values, and thus be made to look just like the increases plotted in Figure 4.

D. Market Reaction

Figure 6 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. In fact, the whole response curve is quite similar to the pattern predicted by the model in Panel D of Figure 1. While the apparent concavity is not a direct prediction of the simple model, a more elaborate model of investor preferences could in principle deliver this sort of pattern.

We examine this pattern econometrically 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 6 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 an 11 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 76 basis points for each \$0.01 change. Small increases in dividends up to \$0.025 are greeted with a market reaction of 36 basis points, or approximately half the slope that we observe in 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 6. 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.

Table 3 provides additional evidence that dividend cuts, in particular, get investors' attention. 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 686 basis points, or almost seven percent more than normal volume. Similar dividend increases are also associated with higher volume but the rate is somewhat smaller at 606 basis points. The difference is not huge in economic terms, but the joint test of the differential sensitivity of volume to dividend changes above and below zero is statistically strong. Also, dividend decreases can be fairly large, bringing the total increase in volume to a more economically significant level.

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 7 and our piecewise linear regressions in Table 4, which themselves use threshold breakpoints. Figure 7 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 6, and changes that cross a threshold from above have a slightly stronger negative reaction than in Figure 6. 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 113 basis points, versus 76 basis points in Table 2. The same comparison for thresholds of \$0.05 is 72 basis points versus 66 basis points. For dividend increases the differences are smaller at 41 versus 36 basis points, and for a \$0.05 increase, the effect is actually somewhat smaller at 22 versus 28 basis points. In short, threshold effects are important on the downside, suggesting that round number thresholds are important reference points.

Another 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 would presumably be stronger reference points. Hence, cutting or raising a dividend after a long streak may have stronger market reactions. The basic idea is that long streaks constitute stronger reference points, so there should be more clustering and the patterns in Table 2 should be more pronounced as the streak lengthens. We plot clustering around successive streaks in Figure 8, examining streaks of length one through 16 separately. We emphasize the general rising trend clustering as the streak length increases. The exceptions are intuitive. A large number of firms reevaluate their dividend policy annually, so there is a drop in clustering after a streak of 4, 8, 12, and 16. Removing the impact of annual periodicity strengthens this conclusion.

Next, we consider the market reaction. We partition the sample into three categories: Decisions following a change in the prior quarter; decisions following no change for up to four quarters, the periodicity of the typical annual board review of dividend policy; and dividend decisions following no change for more than four quarters. Again, we do this analysis once with sorts in Figure 9 and once with a piecewise linear regression in Table 5. The results are as predicted. The no streak sample has essentially no difference between the effect of a decrease and increase in the neighborhood of zero, but the patterns are stronger as the streaks get longer.

For example, in the long streak sample, the market reaction to dividend cuts is stronger than gains, and also stronger than the unconditional coefficients in Table 2 at 106 basis points per $0.01 change in dividends just below zero, versus 76 basis points in Table 2. The market reaction to a dividend increase is also larger at 54 versus 36 basis points in Table 2, despite being half of the reaction to a dividend cut in Table 4. 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 shown in Figure 10 provides a fascinating demonstration of how the reference point is set to appeal to the relevant investor base. In December of 1998, British Petroleum acquired Amoco to form BP-Amoco. BP was listed on the London Stock Exchange but also traded through an ADR. Panel A shows that prior to the merger, Amoco had increased dividends by $0.025 each year for the prior four years. BP had increased dividends by £0.0125 semiannually for the previous two years. Not surprisingly, the dollar dividend on the ADR was hardly so regular.

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

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 quantity of data, we use split-adjusted values, so 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 announcement return is approximately 60 basis points and the median is zero.

G. A Placebo Test

We conduct a final and perhaps preferred test that reinforces the suggestion from the BP-Amoco case. Namely, that zero change in dollar 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 11 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 4. 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. In any event, it is plain that dividend policy is more often decided in local currency than dollar terms.

We examine the market reaction to these changes in Table 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 0.0136 to 0.0034. Moreover, there is no asymmetry in the neighborhood of zero. The marginal reaction to small cuts is on average about the same as the reaction to small increases. Together, these results suggest that neither ADR boards nor investors view past dividends – paid in dollars – as an important reference point. Again, the corollary from this placebo test is that changes in dividend policy are important because of an endogenously chosen and acknowledged reference point, not because changes in this neighborhood would otherwise have been economically important.

VI. Conclusion

Standard dividend signaling theories posit that executives use dividends to destroy some firm value and thereby signal that plenty remains. The money burning typically takes the form of tax-inefficient distributions, 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 what we believe to be a more realistic signaling approach. We use core features of the well-documented prospect theory value function to create 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, dividend smoothing and the Lintner partial-adjustment model, and the preference for regular dividends over regular repurchases. We also find support for its broader intuition that dividends are paid in ways that make them memorable and thus serve as stronger reference points and signals.

References

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

Aerts, Walter, Geert Van Campenhout, and Tom Van Caneghem, 2008, Clustering in dividends: Do managers rely on cognitive reference points? *Journal of Economic Psychology* 29, 276-284.

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* 38:1, 231-292.

Baker, Malcolm, Xin Pan, and Jeffrey Wurgler, 2012, The effect of reference point prices on mergers and acquisitions, *Journal of Financial Economics*, forthcoming.

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.

Barberis, Nicholas, Ming Huang, and Richard Thaler, 2006, Individual preferences, monetary gambles, and stock market participation: A Case for Narrow Framing, *American Economic Review* 96, 1069-1090.

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.

Damodaran, Aswath, 1989, The weekend effect in information releases: A study of earnings and dividend announcements, *Review of Financial Studies* 2, 607-623.

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

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.

Grullon, Gustavo, and Roni Michaely, 2002, Dividends, share repurchases, and the substitution hypothesis, *Journal of Finance* 57, 1649-1684.

Guttman, Ilan, Ohad Kadan, and Eugene Kandel, 2010, Dividend stickiness and strategic pooling, *Review of Financial Studies* 23, 4455-4495.

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, 2012, A Lintner model of dividends and managerial rents, *Journal of Finance*, forthcoming.

Leary, Mark T., and Roni Michaely, 2011, Determinants of dividend smoothing: Empirical evidence, *Review of Financial Studies* 24, 3197-3249.

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.

Michaely, Roni, and Michael R. Roberts, 2012, Corporate dividend policies: Lessons from private firms, *Review of Financial Studies* 25, 711-746.

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.

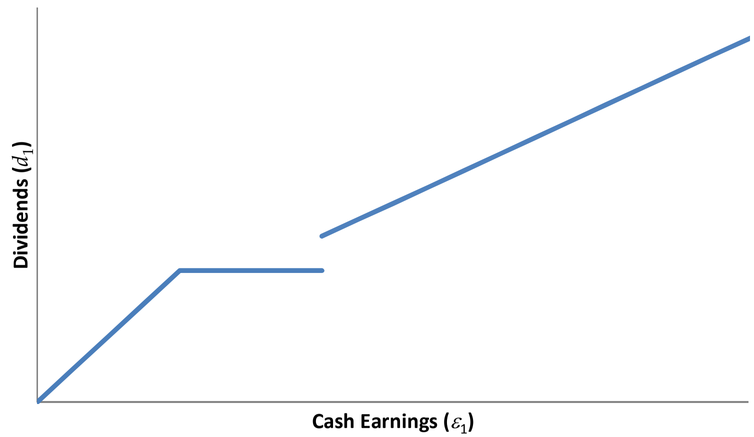
Thaler, Richard H., 1999, Mental accounting matters, *Journal of Behavioral Decision Making* 12, 183-206.

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

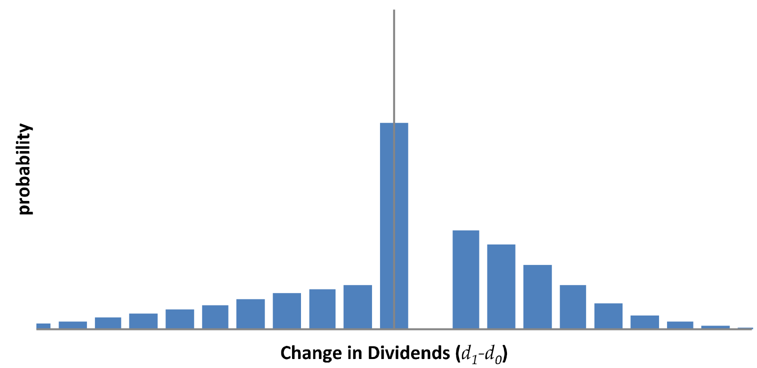
Yoon, Pyung Sig, and Laura T. Starks, 1995, Signaling, investment opportunities, and dividend announcements, *Review of Financial Studies* 8, 995-1018.

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

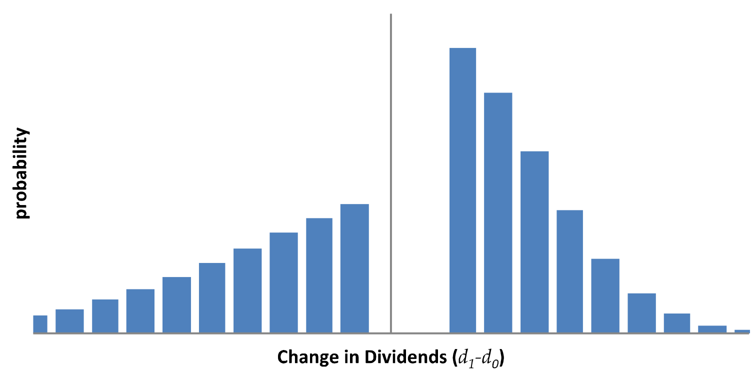
Panel A. Equilibrium dividends as a function of cash earnings



Panel B. Histogram of dividend changes



Panel C. Histogram of dividend changes, excluding zero



Panel D. Market reaction to dividend changes

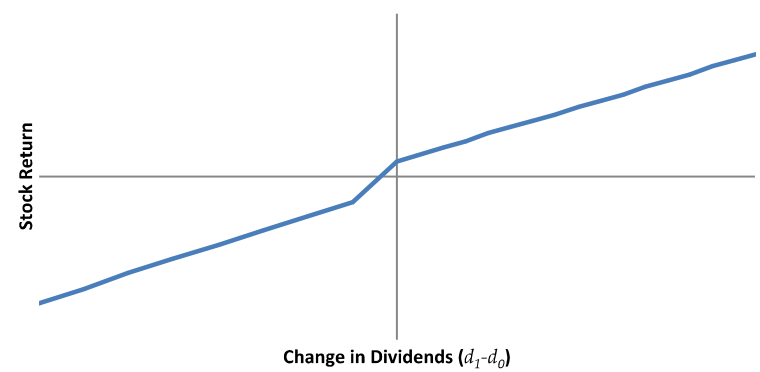
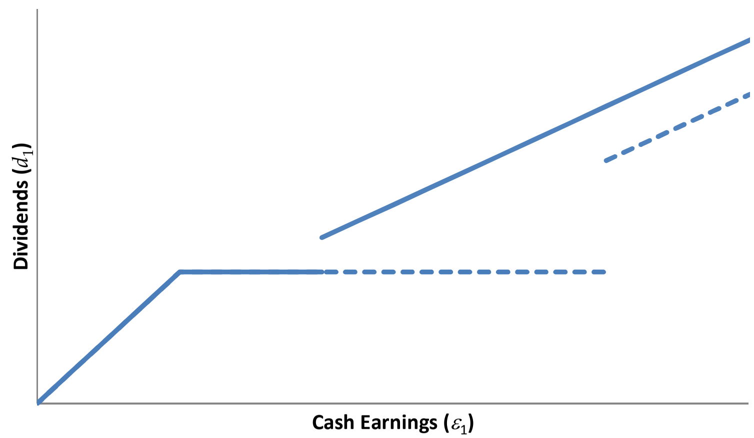
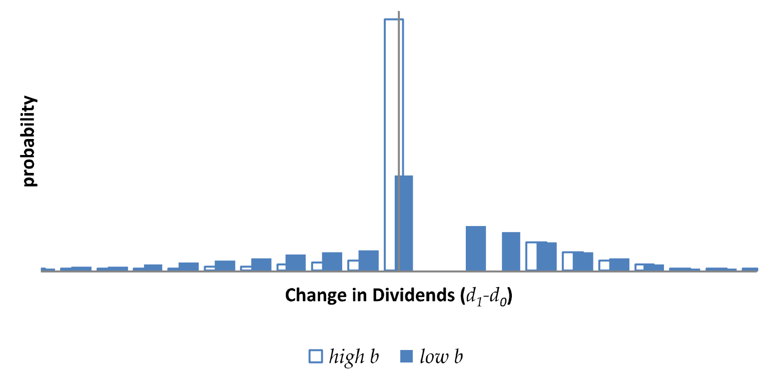


Figure 2. Equilibrium dividend policy: Weak and strong reference points. We plot the equilibrium in Proposition 1 for parameter values $d_0 = 1$ and $\lambda = 0.5$. We repeat Figure 1, comparing the results for $b = 2.1$ and $b = 2.3$.

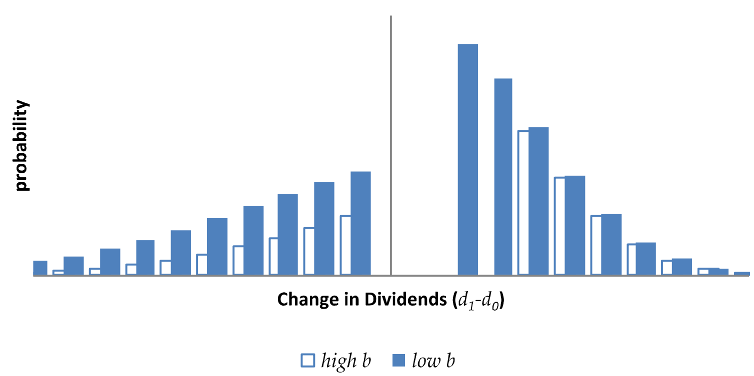
Panel A. Equilibrium dividends as a function of cash earnings



Panel B. Histogram of dividend changes



Panel C. Histogram of dividend changes, excluding zero



Panel D. Market reaction to dividend changes

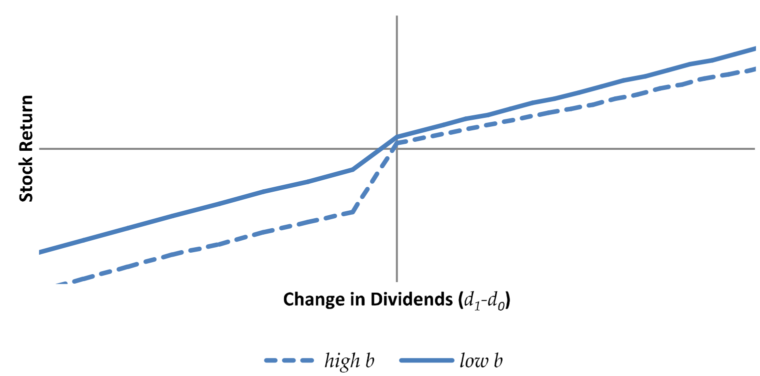
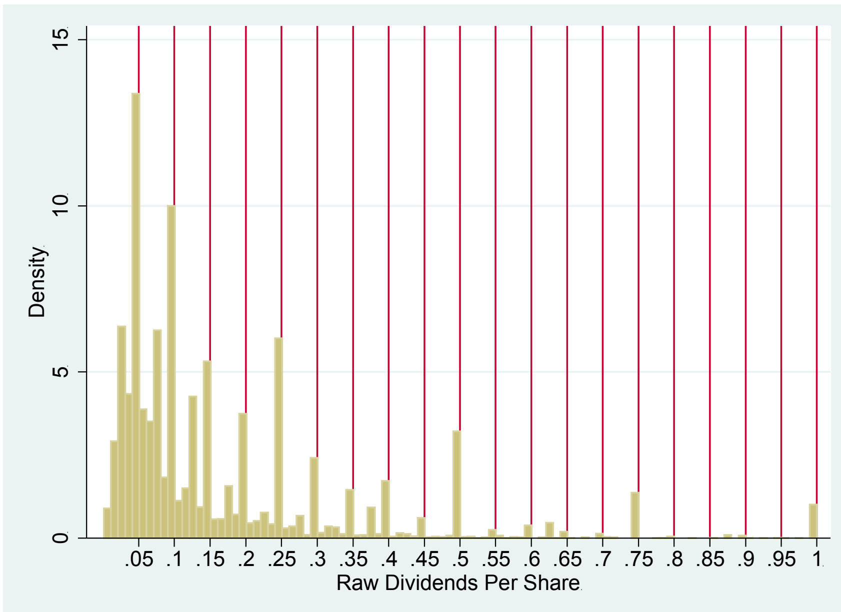


Figure 3. Dividends per share. Histogram of dividends per share and the second and third digits of dividends per share. Panel A shows the distribution of announced dividends per share, while Panel B shows the distribution of the second and third decimal in the announced dividends per share. The sample 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

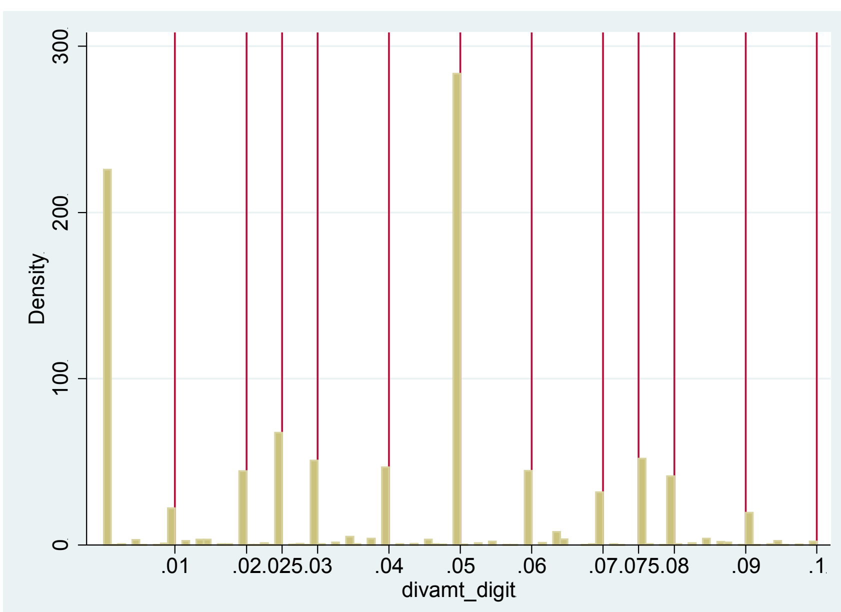
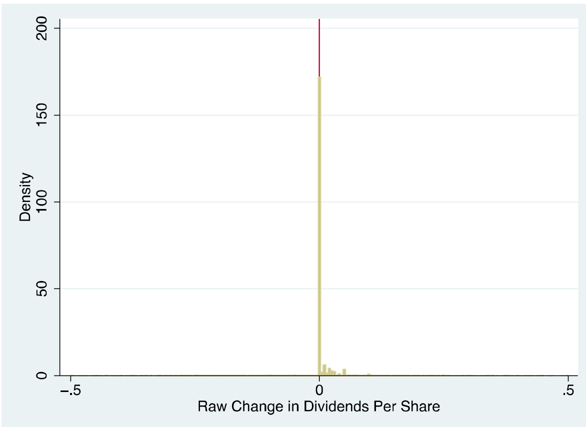
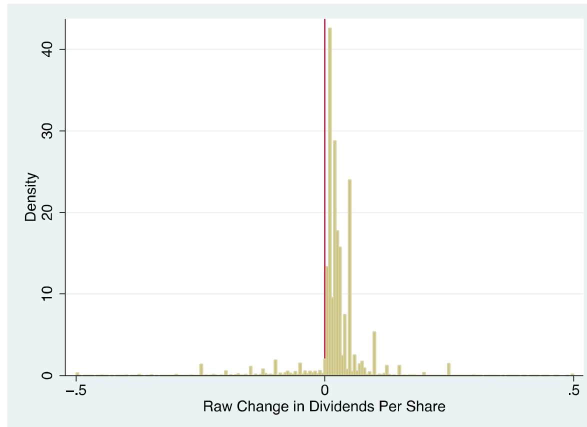


Figure 4. Changes in dividends per share. Histogram of changes in dividends per share. Panel A shows the distribution of changes in dividends per share, Panel B shows the distribution of changes in dividends per share, excluding zero, Panel C shows the distribution of dividend per share increases, and 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.

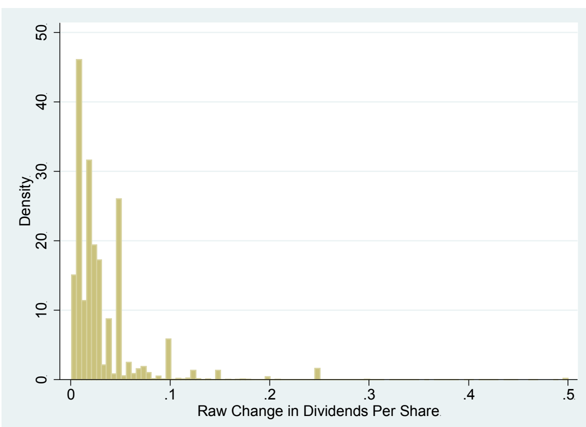
Panel A. Changes in dividends



Panel B. Excluding zero



Panel C. Increases in dividends



Panel D. Decreases in dividends

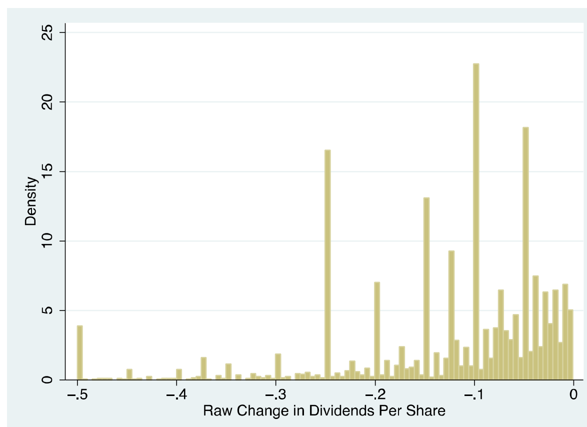
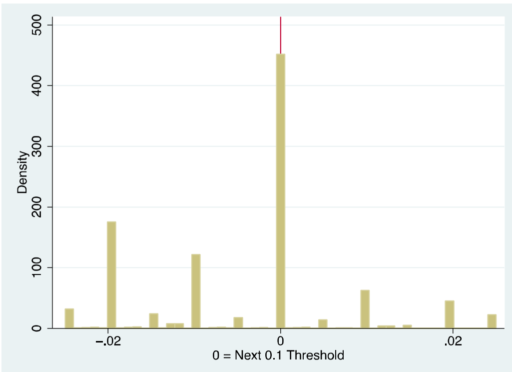
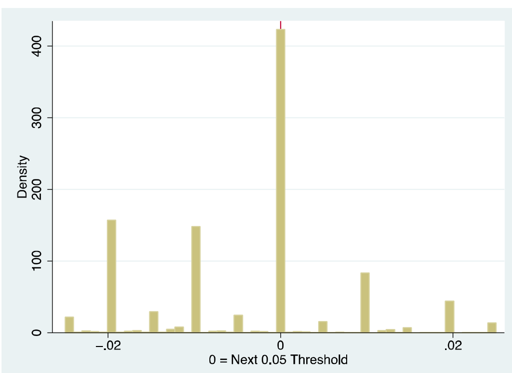


Figure 5. Reaching thresholds in dividends per share. Histogram of changes in dividends per share with changes centered on the next threshold. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Panel A, B, and C show the distribution of changes centered 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.

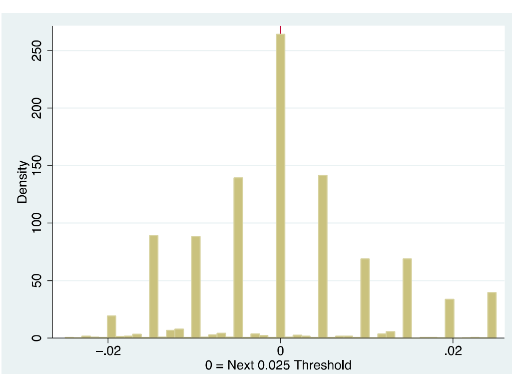
Panel A. Threshold = 0.1



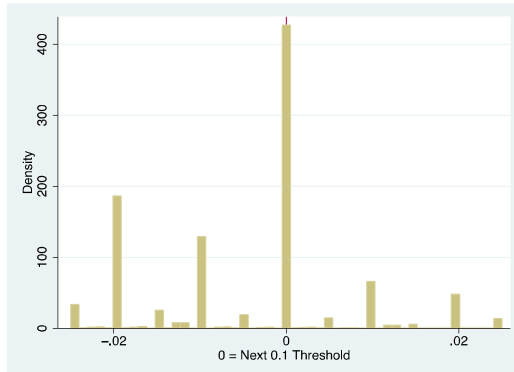
Panel B. Threshold = 0.05



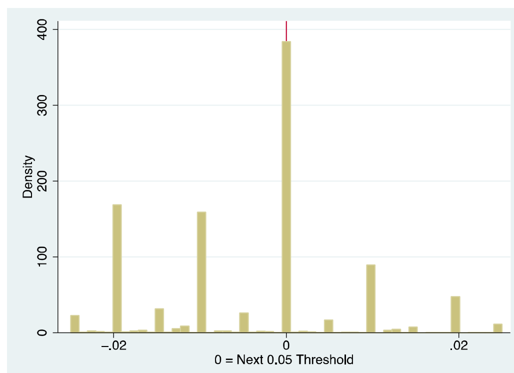
Panel C. Threshold = 0.025



Panel D. Threshold = 0.1



Panel E. Threshold = 0.05



Panel F. Threshold = 0.025

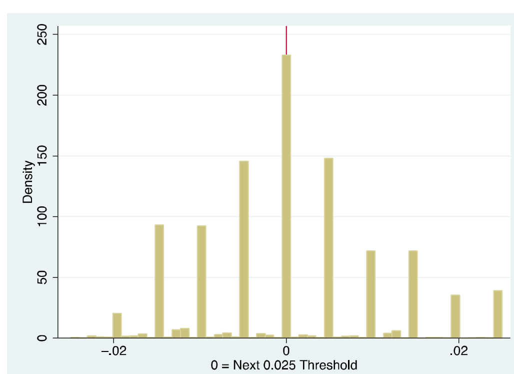
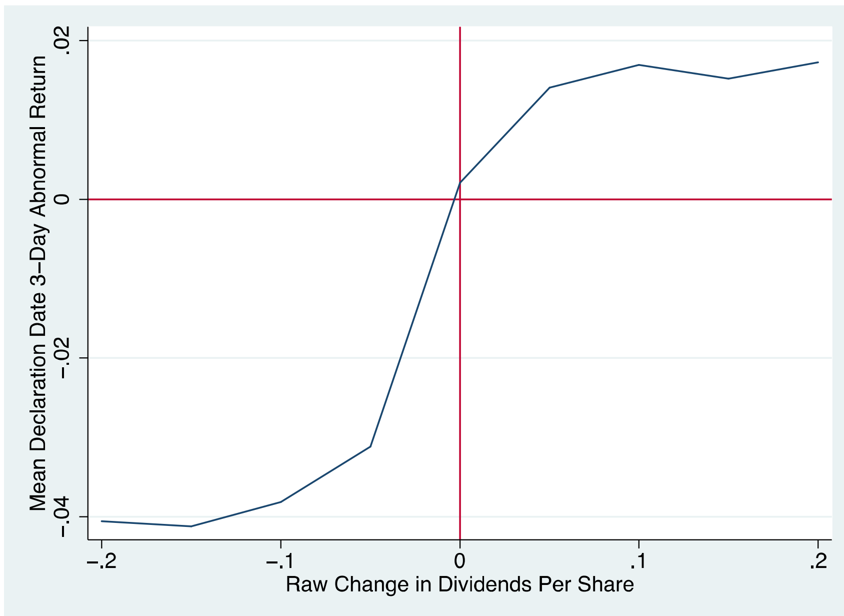


Figure 6. Market reaction to changes in dividends per share. Average 3-day abnormal return by change in dividends per share. Panel A groups changes in dividends per share into groups of 0.05, while Panel B groups changes in dividends per share into groups of 0.025. The groups are formed by rounding the 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

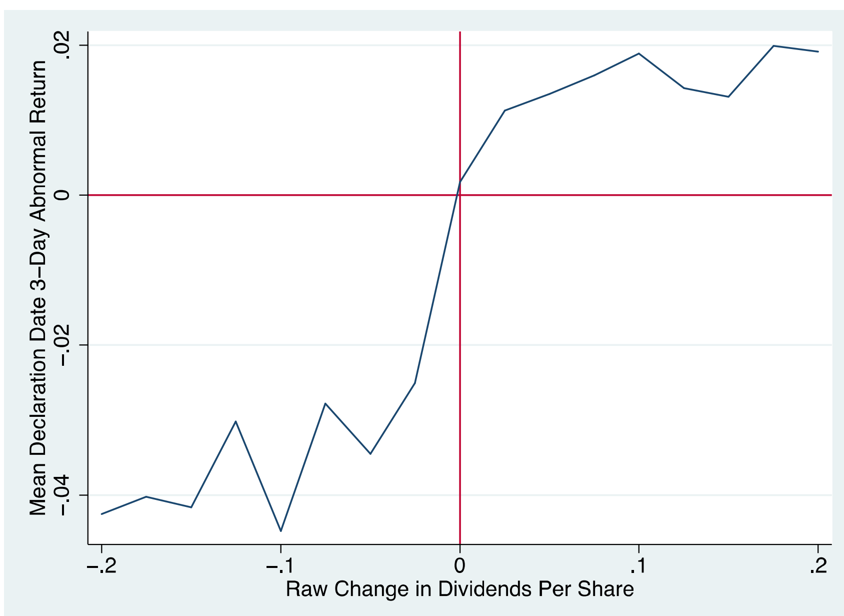
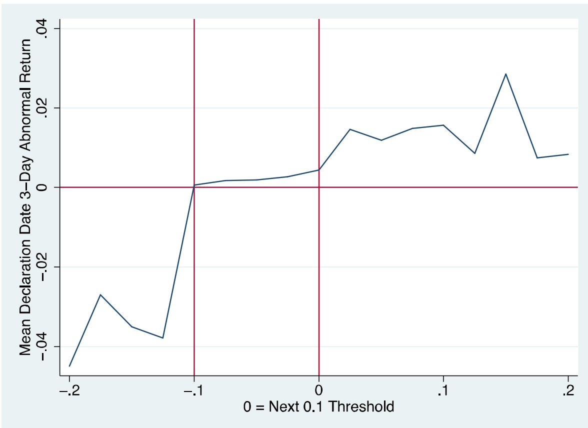
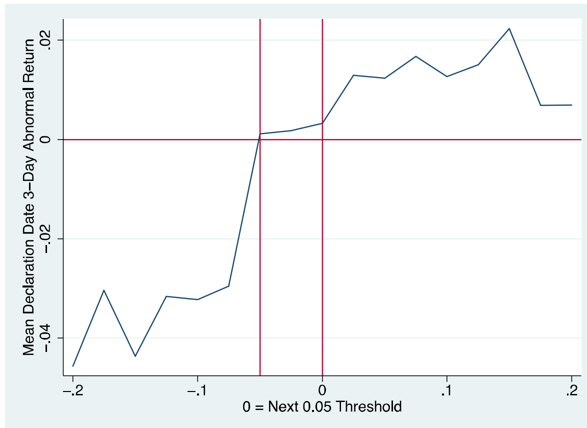


Figure 7. Market reaction to threshold changes in dividends per share. Average 3-day abnormal return by change in dividends per share relative to the next threshold. A threshold is defined as the next round number (0.1, 0.05, or 0.025) above the past level of dividends. Panel A shows changes in dividends 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

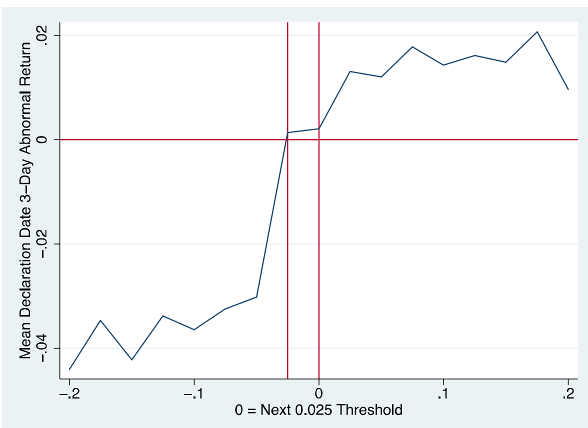


Figure 8. Clustering following dividend per share streaks. The number of observations and the clustering percentage for streaks of up to four years. We divide the sample of dividend announcements into those where the dividend changed from the previous period, a streak of 0, where the dividend remained the same after being changed the previous period, a streak of 1, and so on. The clustering percentage at a streak of 1 is the percentage of firms that do not change their dividend after having changed their dividend in the previous period. The clustering percentage at a streak of 2 is the percentage of firms that do not change their dividend after having maintained their dividend at the same level for the previous two periods. And so on.

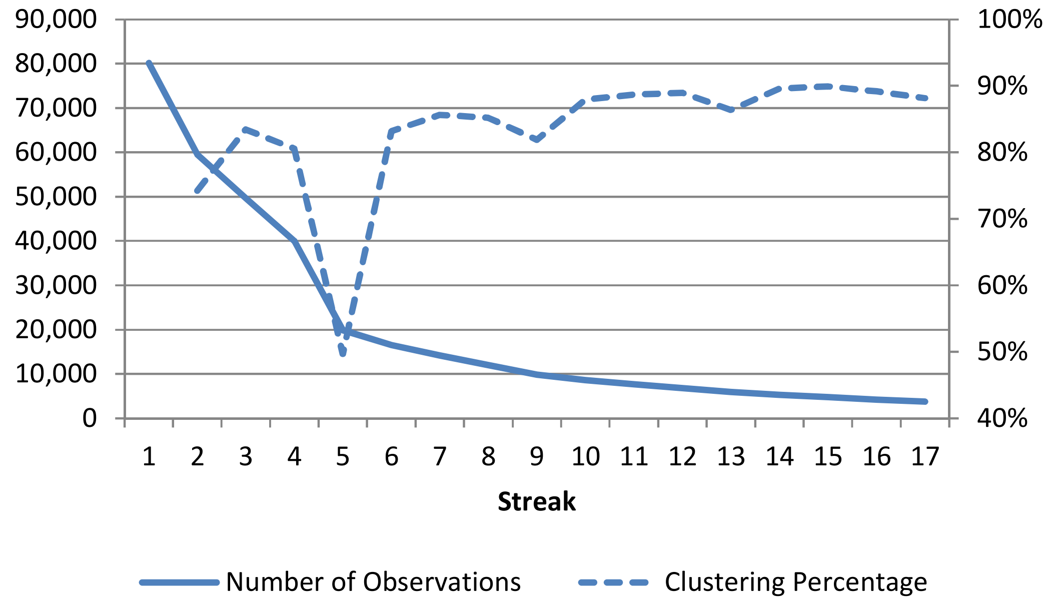


Figure 9. Market reaction following dividend per share streaks. Average 3-day abnormal return by change in dividends per share. Changes in dividends per share are sorted into groups of 0.05, by rounding the raw dividend per share down to the nearest threshold. We partition the sample into situations 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.

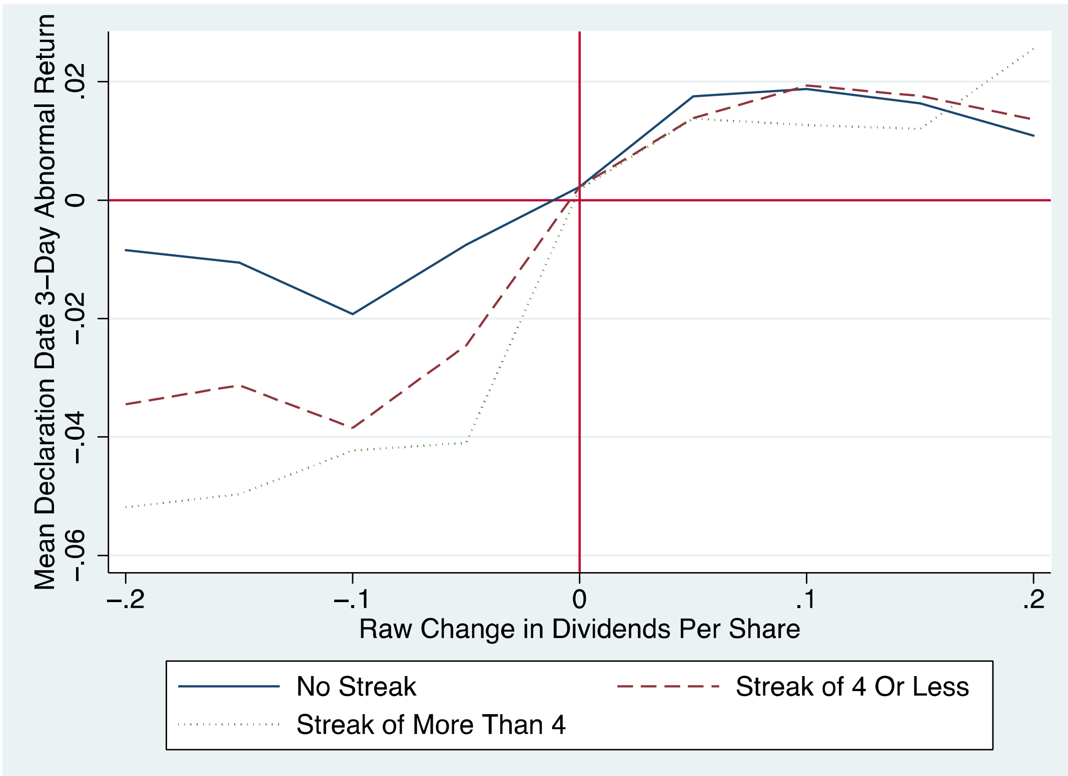
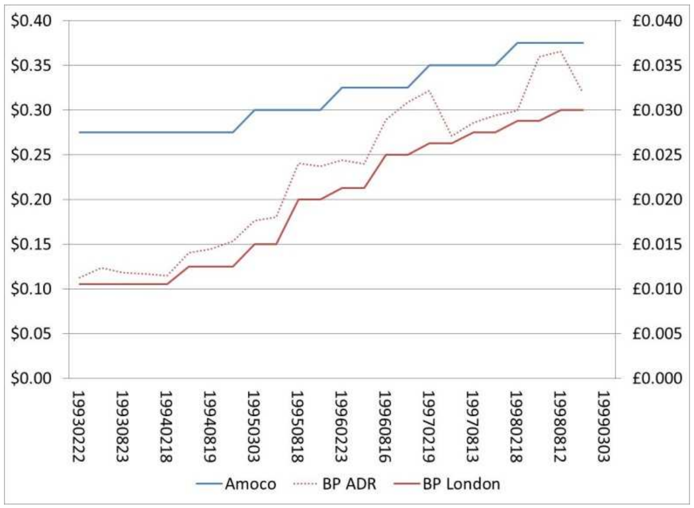
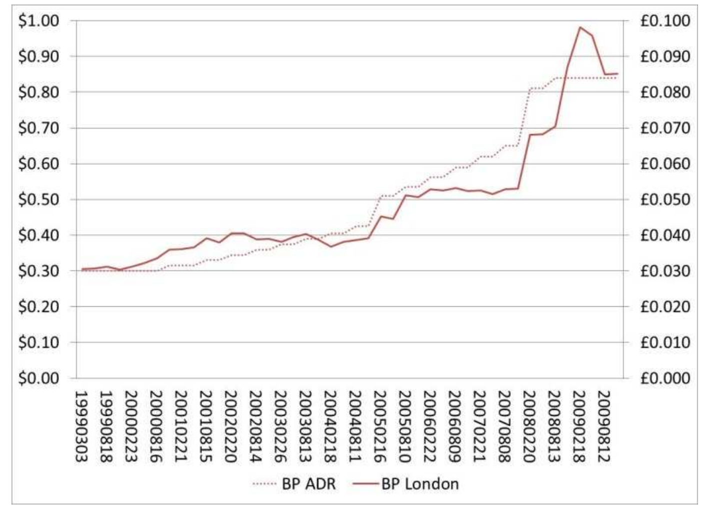


Figure 10. BP-Amoco dividend policy. Split-adjusted dividends per share for BP, Amoco, and the merged company. BP and Amoco merged in December of 1998 forming BP-Amoco. Panels A and C show dividend levels and changes prior to the merger. Panels B and D show dividends after the merger.

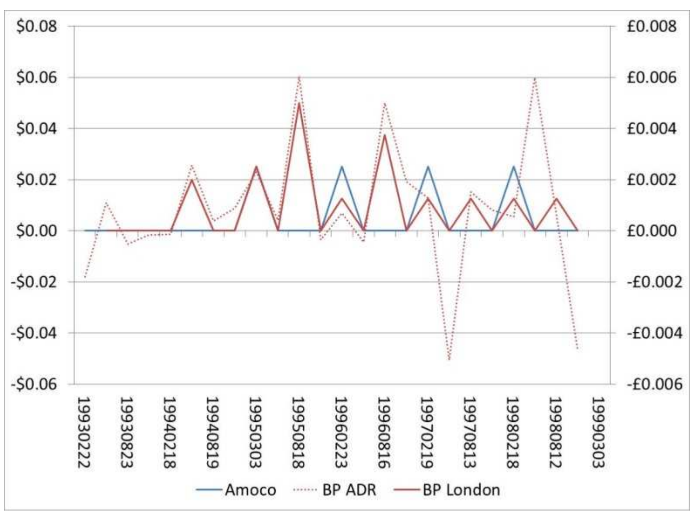
Panel A. Pre-Merger Dividend Levels



Panel B. Post-Merger Dividend Levels



Panel C. Pre-Merger Dividend Changes



Panel D. Post-Merger Dividend Changes

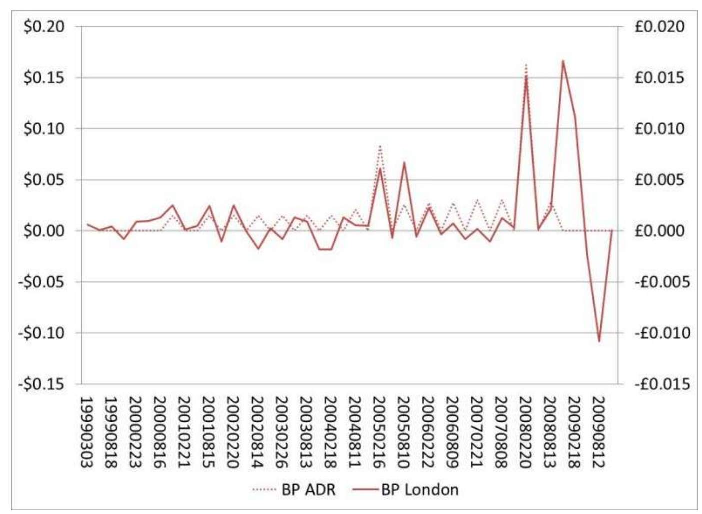
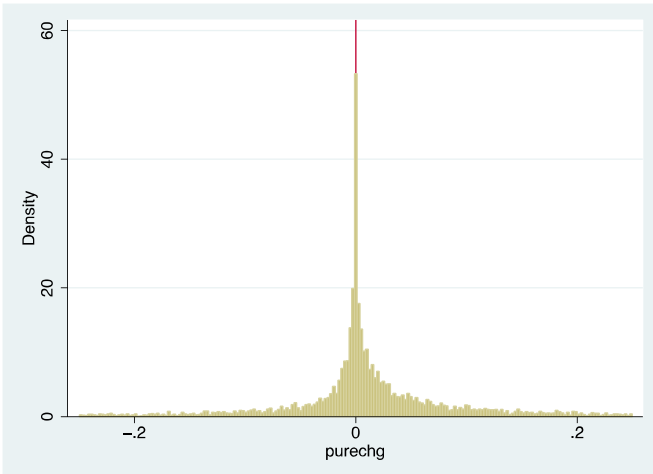
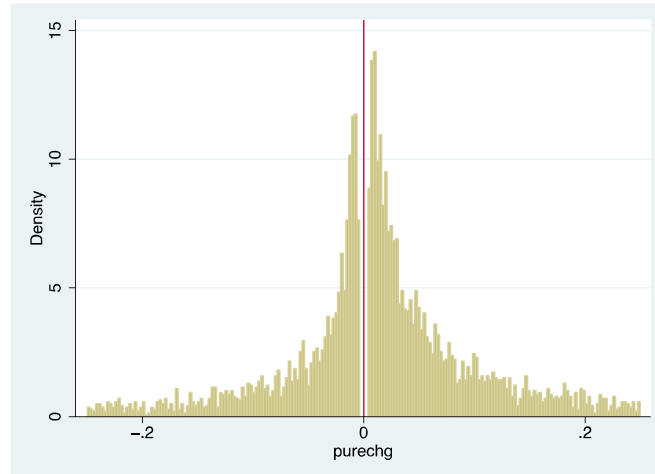


Figure 11. Changes in dividends per share. Histogram of changes in dividends per share. Simple changes in consecutive dividends per share (DD) from Datastream for the following Datastream types: QTR, HYR, YR, FIN, INT. Panels A and C shows the distribution of changes in dividends per share, and Panels 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.

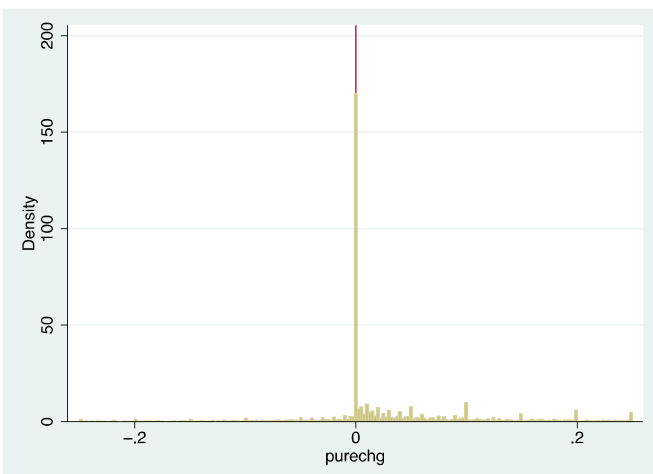
Panel A. ADRs, Changes in Dividends Per Share



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



Panel C. Parents, Changes in Dividends Per Share



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

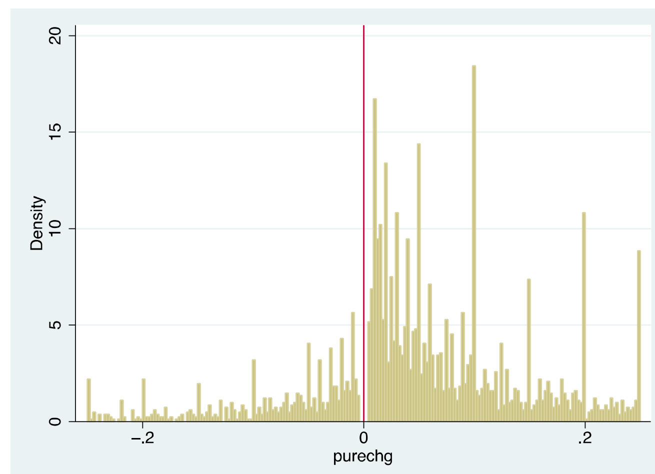


Table 1. Summary statistics. The sample includes all records from the CRSP event file with a Distribution Codes (DISTCD) of 1232 (ordinary taxable cash dividends, paid quarterly) with a Share Code (SHRCD) of 10 or 11 (ordinary common shares, excluding companies incorporated outside the US, Americus Trust 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 computed only over three-month windows (with an allowance for the fact that gaps between dividend payments are frequently two or four calendar months) with no stock splits. 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	391865	0.245	0.195	0.214	0.038	0.100	0.326	0.625
Second and Third Digit	391865	0.038	0.045	0.028	0.000	0.010	0.060	0.085
Change in Dividends Per Share	327816	0.003	0.000	0.030	0.000	0.000	0.000	0.030
Change to Next 0.1 Threshold	327816	-0.059	-0.050	0.041	-0.100	-0.087	-0.040	-0.010
Change to Next 0.05 Threshold	327816	-0.031	-0.030	0.034	-0.050	-0.050	-0.020	0.000
Change to Next 0.025 Threshold	327816	-0.015	-0.020	0.031	-0.025	-0.025	-0.005	0.015
Constant Dividend Streak	383809	6.670	4	8.264	1	2	8	23
3-Day Announcement Return	327026	0.002	0.000	0.040	-0.053	-0.017	0.019	0.063

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

Table 3. Market reaction to changes in dividends per share: Abnormal Volume. Piecewise linear regressions of average 3-day abnormal volume 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	29.73	[5.45]				
$b_{-\infty,-0.2}$			95.84	[4.93]	95.88	[4.93]
$b_{-0.2,-0.1}$			77.50	[1.52]	76.86	[1.51]
$b_{-0.1,-0.05}$			-374.41	[-3.58]	-365.25	[-3.74]
$b_{-0.05,-0.025}$			-576.20	[-1.94]		
$b_{-0.025,0}$			-686.12	[-2.94]		
$b_{-0.05,0}$					-637.73	[-9.46]
$b_{0,\ 0.05}$					522.41	[33.72]
$b_{0,\ 0.025}$			605.63	[20.10]		
$b_{0.025,\ 0.05}$			360.93	[7.18]		
$b_{0.05,\ 0.1}$			79.66	[1.89]	30.74	[0.77]
$b_{0.1,\ 0.2}$			-63.96	[-1.55]	-58.70	[-1.43]
$b_{0.2,\ \infty}$			-7.90	[-0.50]	-8.19	[-0.52]
N		288,740		288,740		288,740
R^2		0.0001		0.0069		0.0069
$b_{-0.1,0}-b_{0,\ 0.1}$			590.51	(0.000)	449.84	(0.000)

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	10.17	[23.12]				
$b_{-\infty,-0.2}$			-5.33	[-3.76]	-5.11	[-3.96]
$b_{-0.2,-0.1}$			2.56	[0.53]	2.11	[0.52]
$b_{-0.1,-0.05}$			14.60	[1.44]	72.18	[14.03]
$b_{-0.05,-0.025}$			113.31	[8.56]		
$b_{-0.025,0}$			7.13	[9.56]		
$b_{-0.05,0}$					5.93	[13.59]
$b_{0,\ 0.05}$					22.08	[15.04]
$b_{0,\ 0.025}$			40.60	[24.72]		
$b_{0.025,\ 0.05}$			-1.53	[-0.46]		
$b_{0.05,\ 0.1}$			8.31	[2.80]	0.90	[0.29]
$b_{0.1,\ 0.2}$			-5.98	[-2.69]	-5.49	[-2.47]
$b_{0.2,\ \infty}$			0.26	[0.25]	0.57	[0.55]
N		249,498		327,026		327,026
R^2		0.0056		0.0124		0.0102
$b_{-0.05,-0.025}-b_{0,\ 0.025}$			72.71	(0.000)		
$b_{-0.1,-0.05}-b_{0,\ 0.05}$					49.35	(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]**
$b_{-\infty,-0.2}$	-7.81	[-4.40]	-3.82	[1.79]	-1.68	[3.40]
$b_{-0.2,-0.1}$	18.32	[2.26]	-0.32	[-1.90]	-7.22	[-3.14]
$b_{-0.1,-0.05}$	-1.62	[-0.08]	24.03	[0.97]	1.27	[0.12]
$b_{-0.05,-0.025}$	4.03	[0.09]	9.45	[0.10]	73.21	[1.24]
$b_{-0.025,0}$	31.96	[1.22]	79.37	[1.29]	106.15	[1.86]
$b_{0, 0.025}$	44.15	[8.51]	29.48	[-2.70]	54.03	[1.65]
$b_{0.025, 0.05}$	12.43	[1.32]	15.25	[0.29]	2.30	[-0.97]
$b_{0.05, 0.1}$	3.33	[0.31]	11.29	[0.71]	-5.73	[-0.79]
$b_{0.1, 0.2}$	-4.09	[-0.51]	-5.61	[-0.18]	-1.94	[0.25]
$b_{0.2, \infty}$	0.72	[0.31]	-1.86	[-1.00]	1.41	[0.23]
N						327,026
R^2						0.0147
$b_{-0.1,0}-b_{0, 0.1}$	-25.54	(0.206)				
$b_{-0.1,0}-b_{0, 0.1}$ - No Streak $b_{-0.1,0}-b_{0, 0.1}$			87.90	(0.002)	161.09	(0.000)

Table 6. Summary statistics. The sample includes all ADRs from Datastream and their matched parents with nonmissing data on dividends per share and without an apparent practice of paying different levels of annual versus interim dividends. The sample includes all ADRs from Datastream with a trading history between 1990 and 2009 on a 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. Returns are Winsorized at the 1^{st} and 99^{th} percentiles. We look back up to three months to find a matching announcement date, and we use announcement dates from ADR-Parent pairs to enlarge the sample. The returns use the Datastream return index (RI) up to five weekdays surrounding the announcement date, with no adjustment for market movements.

	N	**Mean**	**Median**	**SD**	**5**	**25**	**75**	**95**
Panel A. US ADRs								
Dividends Per Share, Split Adjusted	10,634	0.438	0.210	2.504	0.018	0.080	0.471	1.354
Change in Dividends Per Share	8,478	0.022	0.002	1.108	-0.228	-0.012	0.044	0.301
5-Day Announcement Return	3,574	0.006	0.000	0.060	-0.088	-0.023	0.031	0.116
Panel B. Parents								
Dividends Per Share, Split Adjusted	31,828	42.307	1.199	878.1	0.0	0.2	5.5	34.0
Change in Dividends Per Share	8,968	-5.042	0.005	1973.1	-2.1	0.0	0.2	5.0
Change in Dividends Per Share, All	28,877	0.094	0.000	1144.0	-4.4	0.0	0.2	5.7

Table 7. Market reaction to changes in dividends per share for ADRs. Piecewise linear regressions of average 5-day announcement return on clean changes in dividends per share. The sample is described in Panel A of 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 are based on the Datastream return index (RI) up to five weekdays surrounding the announcement date.

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