Dividends as Reference Points: A Behavioral Signaling Model[*]

Malcolm Baker
Harvard Business School and NBER

Jeffrey Wurgler
NYU Stern School of Business and NBER

April 20, 2011

Abstract

We propose a signaling model in which investors are loss averse to reductions in dividends relative to the reference point set by prior dividends. Managers with strong but unobservable earnings separate themselves by paying high dividends and still retaining enough earnings to be likely to at least match the same dividend next period. The model matches several important features of the data, including equilibrium dividend policies that can follow a Lintner partial-adjustment model; a modal dividend change of zero; a stronger market reaction to dividend cuts than dividend increases; and a signaling mechanism that does not involve public destruction of value, a notion that managers reject in surveys. We also find empirical support for some novel predictions.

I assume [the Lintner model] to be a behavioral model, not only

from its form, but because no one has yet been able to derive it as

the solution to a maximization problem, despite 30 years of trying.

-- Merton Miller (1986, p. S467)

I.   Introduction

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

Yet the very strongest views expressed in the Brav et al. survey are reserved for mechanisms behind traditional theories of dividend signaling. Executives viewed these as broadly misguided. The notion that dividends are used to show that their firm can bear costs such as borrowing external funds or passing up investment was summarily rejected (4.4% agreement). The idea of signaling through costly taxes did not receive much more support (16.6%). These findings cast some doubt on the mechanisms driving signaling models by Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), and Allen, Bernardo, and Welch (2000), among others.

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

The essence of the model is that investors evaluate current dividends against the 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. The model also yields novel predictions, which we test.

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 must balance 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 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 exactly the reference level dividend; and managers with very strong first-period earnings pay out only a fraction,

2

raising the reference level for the future but, given their savings and average second-period earnings, to a level they are likely to be able to maintain.

The model matches important facts about dividend policy. The modal dividend change is zero. For reasonable parameter values, firms with high or stable earnings engage in a Lintner partial-adjustment policy. 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. And, notably, the model does not ask profitable firms to destroy fundamental value—burn money—in order to distinguish themselves. Its mechanism is not inconsistent with the survey evidence.

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

Our results broadly support these novel predictions. Both dividend levels and changes are 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). As our model predicts, this asymmetry is more pronounced when the same dividend per share has been paid for more than four consecutive quarters. As a placebo test, we examine ADRs, where we find, not surprisingly, that the reference point and round

numbers are denominated in foreign currency. In this sample, there is no clustering and nothing special about the market's reaction around zero dividend changes in U.S. dollars.

Other papers have made connections between dividends and reference points or prospect theory. One that uses prospect theory even more heavily than we do is Shefrin and Statman (1984). They argue that dividends improve the utility of investors with prospect theory value functions if they also mentally account (Thaler (1999)) for dividends and capital gains and losses separately. Their theory is an important contribution but quite different from ours. Dividends serve no signaling function in their theory, do not follow a Lintner policy, and can be evaluated against a reference level of no dividend, for instance. Another important contribution is Lambrecht and Myers (2010). In their model, managers maximize the present value of the utility of rents that they can extract from profits. They smooth dividends because they have habit formation preferences and rents move in lockstep with dividends given the budget constraint. Lambrecht and Myers do, importantly, derive a Lintner policy in equilibrium, but there are a variety of differences with our theory and predictions. We discuss these papers and other related research. More generally, our paper adds to the literature on behavioral corporate finance surveyed by Baker, Ruback, and Wurgler (2007).

Section II reviews the relevant literature on reference-dependent utility. 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.

A.   Reference-Dependent Utility and Loss Aversion

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

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

B.   Reference Points

If gains and losses matter, how are they defined? In other words, what is the reference point and how is it formed? Can it change? What determines its strength? Can there be multiple reference points?

The literature on prospect theory does not provide answers to these questions. The relevant reference point depends on the setting and, in static choice settings, is frequently unambiguous. For example, in the applications and experiments above, the reference point is obviously the decision maker's current position. But in many circumstances, "current position" is not always so well defined. In Abel (1990), for example, the reference point for utility includes others' current consumption levels.

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

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

Reference points can also differ in their temporal character. In dynamic situations with uncertainty, the reference point is even harder to generalize about. It may involve the future, not

just the present. In Koszegi and Rabin (2006, 2009), agents are loss averse over changes in beliefs about future outcomes such as consumption. Here, expectations about the future make up the reference point. For example, utility might depend in part on the prospect of a raise.

Past circumstances can also supply powerful reference points. Genesove and Mayer (2001) find that people resist selling their homes below its purchase price. Shefrin and Statman (1985) find that the purchase price of a security serves as a reference point. Odean (1998) confirms this, and also suggests, like Arkes, Hirshleifer, Jiang, and Lim (2008), that such reference points can change over time, albeit sluggishly. Baker and Xuan (2009) argue that the stock price that a new CEO inherits is an important reference point for raising new equity. The idea of one's prior consumption as a reference point for the utility of current consumption is represented through internal habit formation preferences as in Constantinides (1990).

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

A stock's 52-week high provides an interesting example of a memorable number that, at least for some purposes, forms a reference point. The shareholder may have a positive

association with that level. It is a specific and salient number. It can be constant (repeated and rehearsed) for up to 52 weeks, but also varies over time. Heath, Huddart, and Lang (1999) find that employee exercise of stock options doubles when the stock price tops its 52-week high. Recent peak prices matter for the pricing and deal success in mergers and acquisitions (Baker, Pan, and Wurgler (2010)). Given that an individual shareholder's purchase price also affects her trading behavior, this is also an example of how behavior and perceptions of value may depend on multiple reference points.

C.   *Past Dividends as Reference Points*

This discussion shows that theory alone cannot identify "the" reference point. The typical research process is to consider the setting, hypothesize the nature of the reference point(s), and then see if the empirical evidence supports the hypothesis. In this paper we hypothesize that past dividends are reference points against which current dividends are judged.

Our hypothesis touches on many of the concepts discussed above. The reference point we hypothesize is based on past experience, as in the disposition effect of Shefrin and Statman. It is also dynamic, as in internal habit formation. Fluctuations in the dividend upset expectations about future dividends. Baker, Nagel, and Wurgler (2007) find that many investors consume the full amount of their dividends, drawing attention further to their level.

Dividends are also packaged to be memorable. They are announced at discrete and regular intervals, and often with some degree of ceremony and fanfare, which encourages the formation of memory. The same level is often repeated for many quarters in a row, further encouraging memory. We will show that they cluster at round numbers, and that changes are commonly in round-number intervals or designed to meet or exceed a round-number threshold.

The memorability of the dividend is central to our theory—it increases their power as reference points and, consequently, as signals.

III. A Model of Signaling With Dividends as Reference Points

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

There are two key ingredients in the model. First, a reference point appears in a representative investor's objective function. In particular, there is a kink in utility, so that a $0.01 drop in dividends just below the reference point is more painful than a $0.01 increase in dividends just above. Second, the manager cares about the current estimate of firm value as well as the long term welfare of shareholders. There are other details in the model, but these are the essential elements.

Reference points shape dividend policy in several ways. On the 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 only works because firms with limited resources are unwilling to incur the expected future costs of missing an endogenous reference point.

Similar results obtain if we replace reference dependent preferences with reference dependent behavior. For example, suppose 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 managers: 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, but otherwise adds complexity and is somewhat further removed from the psychology of reference dependence.

A.   Setup

The model focuses on 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^*. 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^* 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's 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 classic 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:

$$E_m\left[\lambda E_i[\varepsilon_1] + (1-\lambda)u(d_1, d_2 \mid d^*)\right]$$ (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^*, and E_m and E_i are the expectations operators for

the manager and the investor, respectively.

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. 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.

Investor's utility. The manager's objective is standard. 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^* and the second kink is

around an endogenous reference point for second-period dividends:

$$u(d_1, d_2 \mid d^*) = d_1 + b(d_1 - d^*)\{d_1 < d^*\} + d_2 + b(d_2 - d_1)\{d_2 < d_1\}$$ (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 dividends below the reference point are more painful than symmetric

dividends above the reference point. This utility function is in the spirit of prospect theory with a

kink at a reference point. We leave out the complexity of curvature. The second-period reference

point equals first-period dividends d_1 by assumption. In reality, the reference point and the

11

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 narrow framing. This is not a reference point applied to total ending wealth, but much more narrowly both across stocks and time, in the spirit of Barberis, Huang, and Thaler (2006).

Figure 1 shows this utility function for d_1 ranging from zero to four, d_2 ranging from zero to two, a first-period reference point of d^* equal to one, and a reference point intensity b of two. This means that decreases in the dividend relative to the reference point of d^* in the first period or of d_1 in the second period are three times as costly in utility as increases. The purpose here is simply to gain intuition about the utility function—there is no optimization here or savings from period to period. Think of these exogenously specified dividend payments as simply equal to cash earnings for the moment. Panel A shows utility as a function of dividend payments d_1 and d_2 and Panels B and C show the range of utility as functions of d_1 and d_2, respectively.

In Panel B, the top of the range occurs when the second-period dividend of d_2 is at its maximum of two. When the first-period dividend ranges from zero to two, the usual prospect theory kink is observable around the first-period reference point of d^*, assumed to be equal to one. Also apparent is that higher first-period dividends can actually lower utility by increasing the second-period reference point. When the first-period dividend of exceeds the maximum second-period dividend of two, any further increase in the first-period dividend is more than offset by the cost of missing this reference point in the second period.

The bottom of the range in Panel B occurs when the second-period dividend of d_2 is at its minimum of zero. In this case, the losses of setting a higher reference point in the first period start right away, damping the slope between zero and one, and reversing it above one, after the first-period reference point of d^* is met. Without any signaling motive, a benevolent manager

would save 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.

In Panel C, the first-period reference point is still visible in overall utility, but there are a variety of ways that different second-period reference points can lead to equivalent utility levels. For example, paying no dividends in the first and second periods is equivalent to paying a first-period dividend of two and no second-period dividend. Both lead to utility of -2. Hence, "folds" in the picture are apparent from this angle, even though, for a given first-period dividend, there is the usual utility kink at the reference point d_1.

Panel D shows the benefit of holding back any resources in the first period and the key ingredient for signaling firm quality, once the first-period reference point is met. Consider a firm with first period resources of 1.2. By paying only 1.0 and saving 0.2 for the second period, there are two effects. First the second-period reference point is lower. It occurs at 1.0 instead of 1.2. Second, the reference point can be met even if second-period cash earnings only come in at 0.8. The shortfall is made up with cash savings from the first period. The difference in utility from this dividend policy is shaded in the last Panel. Integrating over possible realizations in the second period, assuming they are uniformly distributed, leads to a quadratic cost in second-period utility of increasing dividends in the first period, which we derive below in Eq. (6).

Information. 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 (2010). The manager is not able to keep the cash for himself,

13

and no real value is created or destroyed with dividend policy. This is, at least in spirit, more consistent with what managers say in surveys about their dividend policy.

The fundamental value of the firm appears in two installments,

$$\varepsilon_1 + \varepsilon_2.$$ (3)

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 manager's reference point, not the firm's reported financials, such as earnings per share or cash balances. Otherwise, the manager has no lever to signal his information about ε_1. For simplicity, we assume that the 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:

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

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

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^*. 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^*, 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 the dividend d_1 is more than half of cash earnings. Intuitively, there is no cost if the manager adopts a very 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. As discussed above, this comes from integrating over the shaded section in the Panel D of Figure 1.

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

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

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 of dividends, and the expected cost of falling short of the new reference point is relevant only when first-period savings alone cannot cover it. Intuitively, given these considerations, there are potentially 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 $d*$. Next, managers cluster at $d*$ 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 $d*$.

Proposition 1. *There exists an equilibrium where: $d_1 = \varepsilon_1$, for $\varepsilon_1 < d*$; $d_1 = d*$, for $d* < \varepsilon_1 < \varepsilon*$; and $d_1 = \frac{1}{2}\varepsilon_1 + \frac{\lambda}{1-\lambda} \cdot \frac{1}{b}$, for $\varepsilon_1 > \varepsilon*$.*

Rational expectations means that the investor believes: $E_i[\varepsilon_1 \mid d_1] = d_1$ when $d_1 < d*$; $E_i[\varepsilon_1 \mid d_1] = \frac{1}{2}(\varepsilon* + d*)$ when $d_1 = d*$; and $E_i[\varepsilon_1 \mid d_1] = 2(d_1 - \frac{\lambda}{1-\lambda} \cdot \frac{1}{b})$ when $d_1 > \frac{1}{2}\varepsilon* + \frac{\lambda}{1-\lambda} \cdot \frac{1}{b}$. Incentive compatibility requires a manager with $\varepsilon_1 < d*$ 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 $d*$ 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 $d*$ 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 $\frac{1}{2}(\varepsilon* + d*)$ to $\varepsilon*$ must equal the cost differential of evaluating Eq. (5) at $d*$ and at $\frac{1}{2}\varepsilon* + \frac{\lambda}{1-\lambda} \cdot \frac{1}{b}$. Possible equilibria $\varepsilon*$ are solutions to a quadratic equation:

$$\lambda \tfrac{1}{2}\left(\varepsilon * - d *\right) - \left(1 - \lambda\right) b\left(\left(\tfrac{\lambda}{1-\lambda} \cdot \tfrac{1}{b}\right)^2 - \left(d * - \tfrac{1}{2}\varepsilon *\right)^2\right) = 0 . \tag{7}$$

There is no claim of uniqueness, so a sufficient proof of Proposition 1 is a numerical example.

Example. Suppose the manager cares equally about stock price and utility, i.e. $\lambda = 0.5$. If $d*$ is 1 and b is 2.5, then Proposition 1 and Eq. (7) indicate that the equilibrium cutoff $\varepsilon*$ is 1.6. For ε_1 above 1.6, the first-period dividend is $\tfrac{1}{2}\varepsilon_1 + 0.4$. This exactly trades off the marginal signaling benefit per unit of dividends of 2.0, using investor beliefs implicit in Proposition 1, against the second-period marginal cost, i.e. the first derivative of Eq. (5), of $5 \cdot \left(d_1 - \tfrac{1}{2}\varepsilon_1\right) = 5 \cdot \left(0.4\right) = 2.0$. For ε_1 between 1.0 and 1.6, the first-period dividend is simply $d*$, 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 $d* = 1.0$, 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 $d*$ 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 $d*$ both to satisfy the kinked investor utility and to raise investor beliefs discontinuously from $d*$ to $\tfrac{1}{2}\left(\varepsilon * + d *\right)$. 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 $d*$. Even firms that could pay somewhat more choose not to, because of the costs of setting a high

17

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 relatively 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. These patterns are shown in Panel A of Figure 2.

Another way to see this is by plotting the histogram of dividend changes in Panels B and C of Figure 2. For this purpose, we assume that first-period cash earnings ε_1 are normally distributed, with a mean of $\frac{1}{2}(\varepsilon^* + d^*)$. 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 2. 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 $\frac{1}{2}(\varepsilon^* - d^*)$ in the investor's expectation. By contrast, it takes a discontinuous increase in dividends of $\frac{1}{2}\varepsilon^* + \frac{\lambda}{1-\lambda} \cdot \frac{1}{b} - d^*$ to achieve the same size increase in investor expectations. As a result, there is a kink in the stock return chart at exactly the reference point of d^*.

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^* is increasing in b.*

Put another way, Proposition 2 says that there is more clustering of dividends at the reference point d^* 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 3. The comparison is between b equal to 2.1 and b equal to 2.3. The exact magnitude of b is not important. If 2.1 seems uncomfortably high, it is worth noting that a similar equilibrium can be sustained at higher d^* and lower b; we have been using parameters that make for clear pictures. Because 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 $\frac{1}{2}(\varepsilon^* + d^*)$ and repeat the plots from Figure 2. Note the implications of a higher b and hence ε^*. There is more clustering and a larger spike in the distribution of dividend changes at the reference point d^*. The market reaction is more negative when dividends fall just short of the reference point, the market reaction is flatter above the reference point, so that the kink at zero in the market reaction is more pronounced.

IV.   The Model and Prior Evidence

Dividend policy is an area 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 its 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 appears to capture many of them at least

as well as existing approaches, of which the best known are Bhattacharya (1979), Miller and

Rock (1985), John and Williams (1985), Kumar (1988), Bernheim (1991), and Allen, Bernardo,

and Welch (2000). Our model may also perform as well in some respects as theories based on

agency problems, catering motives, or clientele effects, although to keep the discussion

manageable we will not make such comparisons.

A.   *Surveys*

Dividend policy is an explicit choice of executives (more precisely, the board). The

proposition that their behavior may be guided not by their own hands but by an unseen higher

market force, and therefore survey evidence should be disregarded, is inappropriate here. We

therefore view the fact that our model is consistent with what managers say about dividend

policy as an important success.

The strongest results fit nicely with the reference point setup. For example, as noted in

the Introduction, the Brav et al. (2005) survey of 384 executives revealed strong agreement that

shareholders will react negatively to cuts in the dividend, whereas the reward for increases is

modest. Executives believe that dividends convey information. As a result, they strive to keep a

stable dividend policy. These are straightforward predictions of the model. It is intrinsically

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

Finally, it is notable that standard signaling theories do not naturally focus on dividends per share. Moreover, standard signaling 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*

As Miller (1986) hypothesized, the Lintner (1956) model can be given behavioral foundations. Given reasonable parameter values, our model predicts that dividends follow a partial-adjustment policy and are, more generally, 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. As an empirical matter, all of this adds up to Lintner-like 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. Lambrecht and Myers (2010)'s model is highly compatible with the Lintner model and smoothing. It is not a signaling model so we do not include it in our horse race. It has many appealing features; on the other hand, it seems unrealistic that thousands of large, established, public U.S. firms smooth their payouts because a coalition of habit-formation managers prefer to smooth out their stealing. The required magnitude of rent extraction alone could be too great.

C. *Announcement Effects*

Even if executives disavow standard signaling models, it is clear that shareholders care about dividends. Aharony and Swary (1980) examine cases in which dividend announcements occur separately from earnings announcements. The average cumulative abnormal return in a 21-

day window surrounding a dividend decrease was on the order of five percentage points. The average cumulative abnormal return surrounding a dividend increase was closer to one percentage point. See also Charest (1978) and subsequent papers.

That dividend cuts would be received especially poorly is a direct 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. We would suggest that signaling with dividends as reference points explains the empirical facts reviewed here as well as or better than standard signaling models.

V.   New Predictions and Tests

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

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

A.   Salience, Ease of Recall, and Repetition

What makes a number like dividends per share memorable? As noted earlier, round numbers are easier to remember. Quarterly repetition of the same dividend helps as well. In our main sample of U.S. dividend payers, we examine whether dividends and dividend changes

concentrate on round numbers, whether the use of salient numbers affects how the market reacts

to dividend changes, and whether repetition increases the strength of a particular dividend level

as a reference point as measured by the market reaction to changes. After this, we consider a

sample of ADRs and how dividend reference points fail to translate across currencies.

B.   Main Sample

Our primary sample of U.S. dividend payments is summarized in Table 1. We obtain

dividend data from the Center for Research on Security Prices (CRSP) database. We start with

all records in the event database with a distribution code (CRSP:DISTCD) equal to 1232. These

are ordinary taxable dividends paid at a quarterly frequency. We further limit the sample to firms

with a share code (CRSP:SHRCD) of 10 or 11. This restricts our attention to ordinary common

shares and eliminates most companies incorporated outside the U.S., Americus Trust

Components, closed-end funds, and REITs. Such firms have dividend policies that may have

reference points denominated in non-dollar currencies or have regulatory or contractual

restrictions on dividend policy. We also eliminate dividend payments of 0, dividend payments

greater than $2.00 per share, and dividends for which there is no declaration date

(CRSP:DCLRDT). Otherwise, the entire CRSP database ending in 2009 appears in Table 1. The

data here start with the beginning of the CRSP file in 1926, but restricting attention to more

recent periods does not change the economic or statistical conclusions below.

Our main variable of interest is the raw dividend payment per share. It is easier to think

about dividend policy as clearing a threshold or creating a new reference point in raw, rather than

split-adjusted terms. The median dividend payment is $0.195. For changes in dividends, we have

a choice. We can examine only changes where no split has occurred since the last dividend or we

can look at split adjusted differences. For the CRSP sample, where quantity of data is not a

problem, we look only at pure changes, where no split has occurred. Even with the more restrictive definition of pure changes, we have almost 250,000 observations. The majority of quarterly dividend changes are zero, and only a small fraction 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 being not uncommon.

To measure the market reaction to a dividend announcement, we compute a three-day abnormal stock return around the declaration date. This is the simple return (CRSP:RET) for the firm in the day before, the day of, and the day after a dividend declaration minus the return of the CRSP value-weighted index over the same window. On average, a dividend declaration is met with approximately a 20 basis point abnormal return. The median abnormal return is also zero. This is a sample of firms that did not omit a dividend, so a slightly positive average is not surprising. We also measure volume over the period from the dividend declaration through three days after. We normalize this volume by taking the log difference between the average daily declaration date volume and the average daily volume in the previous 90 calendar days.

C.   *Round Numbers and Reaching for Thresholds*

Dividends are paid in round numbers. This is apparent in a simple histogram of dividends per share in Panel A of Figure 4. A nickel per share is the modal dividend, a dime the second most common amount, and a quarter is very nearly the third most common amount. There are

notable spikes at other round multiples of $0.05. Panel B shows that the most common values for the second and third digits are 0.050, followed by 0.000, and to a much lesser extent 0.025 and 0.075. Other round multiples of 0.01 are somewhat less common, and non-round values are rare.

Dividend changes are also made in round numbers. Figure 5 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 2 exactly: the very large mass at exactly zero, a discontinuity between negative and positive changes even when the mass point at zero is removed, and considerably more clustering just above zero than just below zero. Panel A shows the discontinuity at zero. Little else is even perceptible because the density at zero renders the rest of the distribution inconsequential. When we remove zero changes in Panel B, clustering just to the right of zero is much more apparent. Again, the density above zero is many times the density at equivalent changes below zero. When we further split the sample into increases and decreases, it is also clear that the left tail of the distribution of dividend cuts in Panel D is longer than the right tail of dividend increases in Panel C, as we observed in Panel C of Figure 2.

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 2. To be appreciated for raising the dividend, firms must do so in more than a trivial way.

Figure 6 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 4 that firms tend to pay dividends in round numbers. We also check situations where a firm is not starting at a $0.025 threshold in Panels D, E, and F. The same pattern appears. Clearly, boards think of communicating dividend policy in an easy to recall dollar per share quantity, rather the alternative of deciding on a very specific dollar amount and dividing this equally among shareholders.

D.   *Market Reaction*

Figure 7 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 strikingly similar the pattern predicted by the model in Panel D of Figure 2. 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 somewhat more formally in Table 2, where we estimate piecewise linear regressions of the market reaction on the change in dividends per share to trace

out the patterns in Figure 7 in a regression framework. We are particularly interested in the shift in slope below and above zero. The first regression is a simple linear regression. Each $0.01 change in dividends leads to a 9 basis point market reaction.

This obscures a highly nonlinear relationship where changes around zero are much more important than larger movements. The second, piecewise linear regression shows that small cuts in dividends up to $0.025 are greeted with a market reaction of 71 basis points for each $0.01 change. Small increases in dividends up to $0.025 are greeted with a market reaction of 35 basis points, or approximately half the slope that we observe in dividend cuts. There are similarly large differences in the next increments, though the reaction per $0.01 of dividend change drops off quickly. As a summary test, we compare the sum of the three coefficients between -$0.10 and zero to those between zero and +$0.10. The slope for dividend cuts is larger both economically and statistically, as one might have guessed from Figure 7. 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 evidence that dividend cuts do 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 667 basis points, or six percent more than normal volume. Similar dividend increases are also associated with higher volume but

28

the rate is somewhat smaller at 570 basis points. Once again, the joint test of the differential sensitivity of volume to dividend changes above and below zero is statistically strong: The market reacts with greater volume following a dividend cut.

We also look for patterns around threshold dividend changes. Instead of sorting the sample around zero dividend change, we sort it around the next $0.10, $0.05, or $0.025 threshold. This means implicitly that we are capturing both the threshold above and the threshold below the current level of dividends in our sorts in Figure 8 and our piecewise linear regressions in Table 4, which themselves use threshold breakpoints. Figure 8 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 7, and changes that cross a threshold from above have a slightly stronger negative reaction than in Figure 7. The numerical results in Table 4 show this somewhat more clearly. The change in market reaction per $0.01 change in dividends below the lower threshold is 100 basis points, versus 71 basis points in Table 2. The same comparison for thresholds of $0.05 is 66 basis points versus 57 basis points. For dividend increases the differences are smaller at 39 versus 35 basis points, and for a $0.05 increase, the effect is actually somewhat smaller at 22 versus 27 basis points. In short, threshold effects are important on the downside, suggesting that round number thresholds are important reference points.

Another, and perhaps cleaner, test of reference points is to examine dividend streaks. If memory is an important part of reference point formation, then repeated dividends of the same amount per share may be stronger reference points. Hence, cutting or raising a dividend after a long streak may have stronger market reactions. 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 9, 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 only annually, so there is a drop in clustering after a streak of 4, 8, 12, and 16. Removing the impact of annual periodicity, the conclusion is strengthened.

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 10 and once with a piecewise linear regression in Table 5. The results are as expected. The no streak sample has essentially no difference between the effect of a $0.01 decrease and a $0.01 increase in the neighborhood of zero. Meanwhile, the patterns are successively stronger for the short and long streak samples.

For example, consider the long streak sample. The market reaction to dividend cuts is stronger than gains, and also stronger than the unconditional coefficients in Table 2 at 93 basis points per $0.01 change in dividends just below zero, versus 71 basis points in Table 2. The market reaction to a dividend increase is also larger at 52 versus 35 basis points in Table 2, despite being half of the reaction to a dividend cut in Table 4. Taken together, these results suggest that repetition increases the strength of a reference point.

E.   *BP-Amoco: A Case Study*

Our last set of tests involves dividends on American Depository Receipts (ADRs). An ADR allows U.S. investors to purchase shares in a company that is incorporated abroad and

listed on a foreign exchange, but without executing a transaction on a foreign exchange in a different currency. Because of foreign exchange volatility, the dividend policy of a firm with an ADR trading in the U.S. is by definition unable to create a reference point in two different currencies simultaneously.

The case of BP-Amoco shown in Figure 11 provides a fascinating demonstration of how the reference point is set to appeal to the relevant investor base. In December of 1998, British Petroleum acquired Amoco to form BP-Amoco. BP was listed on the London Stock Exchange but also traded through an ADR. Panel A shows that prior to the merger, Amoco had increased dividends by $0.025 each year for the prior four years. BP had increased dividends by £0.0125 semiannually for the previous two years. Not surprisingly, the dollar dividend on the ADR was hardly so regular.

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

F.   ADR Sample

For a broader analysis, we start with a list of ADRs and matched parents from Datastream over the period from 1990 through 2009 described in Table 6. We restrict the sample to firms with an ADR traded on the New York Stock Exchange, the NASDAQ, and other U.S. OTC exchanges. This gives us a preliminary list of 4,916 Datastream codes for ADRs and their

31

parents. Despite this large initial number of potential firms, the coverage and quality of Datastream dividend data is much lower than CRSP, so our tendency in forming a sample is one of inclusion rather than restricting attention to the cleanest situations. Some of the parents appear more than once, meaning that there is more than one ADR for a given parent firm. We treat these as separate observations.

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

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

The dividends per share for the parents are paid in a wide range of currencies, from Yen to Euro, so the levels of dividends per share are sometimes an order of magnitude higher. The median for ADRs is $0.194. The median dividend payment for the parent sample is 1.2, which includes many small dividend payments in more valuable currencies and many large ones in less valuable currencies. There is no unambiguous way to put all of these currencies on level terms, without losing the essential nature of a reference point analysis, so we leave them in raw terms.

Our specific tests involve the market reaction around zero dividend changes for ADRs. We compute announcement returns for the ADR sample by merging declaration dates from Worldscope (Datastream: DECQ1-DECQ4) to Datastream return indexes (Datastream:RI) for the five-day window surrounding the declaration date. Dividend payments are matched to declaration dates that occur for up to three previous months in an attempt to increase coverage. We also use a slightly larger window to capture the lower quality of Worldscope's declaration dates. The average dividend announcement return is approximately 110 basis points, though the median is again zero.

G.   *A Placebo Test*

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

Figure 12 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 5. Part of this is because we broadened the sample as much as possible, perhaps at the cost of including some special or liquidating dividends, and part of this may be because the Datastream data is lower quality.

We examine the market reaction to these changes in Table 7. The first observation is that the relationship between dividend changes and the market reaction is everywhere less economically and statistically significant. The R-squared drops from 0.0127 to 0.0004. Moreover, there is no clear pattern in the neighborhood of zero. The reaction to cuts is on average almost the same as the reaction to increases. Together, these results suggest that neither ADR boards nor investors view past dividends – paid in dollars – as an important reference point. The corollary from this placebo test is that changes in dividend policy are important because of an endogenously 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 of value remains. The money burning takes the form of tax-inefficient distributions, foregone profitable investment, or costly external finance. The

executives who actually set dividend policy overwhelmingly reject these ideas—yet, at the same time, are equally adamant that "dividends are a signal" to shareholders and that cutting them has negative consequences.

We develop a more realistic signaling theory. We use core features of the prospect theory value function to create a model in which past dividends are reference points against which future dividends are judged. The theory is consistent with several important aspects of the data, including survey evidence, patterns of market reaction to dividend announcements, and dividend smoothing and the Lintner partial-adjustment model. We also find support for its novel prediction that dividends are formed in ways that make them memorable and thus stronger reference points.

References

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

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

Arkes, Hal, David Hirshleifer, Danling Jiang, and Sonya S. Lim, 2008, Reference point adaptation: Tests in the domain of security trading, *Organizational Behavior and Human Decisions Processes* 105, 67-81.

Atkinson, R. C., and R. M. Shiffrin, 1968, Human memory: A proposed system and its control processes, In *The Psychology of Learning and Motivation (Volume 2)*, K.W. Spence and J.T. Spence, eds., New York: Academic Press, 89-195.

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

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

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

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

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.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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