THE JOURNAL OF FINANCE • VOL. LIX, NO. 3 • JUNE 2004

A Catering Theory of Dividends

MALCOLM BAKER and JEFFREY WURGLER*

ABSTRACT

We propose that the decision to pay dividends is driven by prevailing investor demand for dividend payers. Managers cater to investors by paying dividends when investors put a stock price premium on payers, and by not paying when investors prefer nonpayers. To test this prediction, we construct four stock price-based measures of investor demand for dividend payers. By each measure, nonpayers tend to initiate dividends when demand is high. By some measures, payers tend to omit dividends when demand is low. Further analysis confirms that these results are better explained by catering than other theories of dividends.

MILLER AND MODIGILIANI (1961) prove that dividend policy is irrelevant to share value in perfect and efficient capital markets. In that setup, no rational investor has a preference between dividends and capital gains. Arbitrage ensures that dividend policy is irrelevant.

Forty-plus years later, the only assumption in this proof that has not been thoroughly scrutinized is market efficiency.[1] In this paper, we argue for a view of dividends that relaxes this assumption. It has three basic ingredients. First, for either psychological or institutional reasons, some investors have an uninformed and perhaps time-varying demand for dividend-paying stocks. Second, arbitrage fails to prevent this demand from driving apart the prices of payers and nonpayers. Third, managers rationally *cater to investor demand*—they pay dividends when investors put higher prices on payers, and they do not pay when investors prefer nonpayers. We formalize this catering view of dividends in a simple model.

The prediction of the model that we focus on in our empirical work is that the propensity to pay dividends depends on a dividend premium (or sometimes

*Baker is from Harvard Business School and NBER; and Wurgler is from the NYU Stern School of Business. For helpful comments, we thank the editor, an anonymous referee, Viral Acharya, Raj Aggarwal, Katharine Baker, Alon Brav, Randy Cohen, Gene D'Avolio, Steve Figlewski, Xavier Gabaix, Paul Gompers, Florian Heider, Laurie Hodrick, Dirk Jenter, Kose John, Steve Kaplan, John Long, Asis Martinez-Jerez, Colin Mayer, Randall Morck, Holger Müeller, Sendhil Mullainathan, Eli Ofek, Lubos Pastor, Lasse Pedersen, Gordon Phillips, Raghu Rau, Jay Ritter, Rick Ruback, David Scharfstein, Hersh Shefrin, Andrei Shleifer, Erik Stafford, Jeremy Stein, Ryan Taliaferro, Jerold Warner, Ivo Welch, Luigi Zingales, and seminar participants at the AFA and EFA annual meetings, Harvard Business School, London Business School, LSE, MIT, the NBER, NYU, Oxford, University of Chicago, University of Michigan, University of Rochester, and Washington University. We also thank Yi Liu, John Long, and Simon Wheatley for data and Ryan Taliaferro for superb research assistance. Baker gratefully acknowledges financial support from the Division of Research of the Harvard Business School.

[1] Allen and Michaely (2002) provide a comprehensive survey of payout policy research.

discount) in stock prices. To test this hypothesis, we use time-variation in four proxies for this dividend premium. The broadest one, which we simply call "the" dividend premium, is the difference between the average market-to-book ratio of dividend payers and nonpayers. The other measures are the difference in the prices of Citizens Utilities' (CU) cash dividend and stock dividend share classes (between 1956 and 1989 CU had two classes of shares which differed in the form but not the level of their payouts); the average announcement effect of recent dividend initiations; and the difference between the *future* stock returns of payers and nonpayers. Intuition suggests that the dividend premium, the CU dividend premium, and initiation effects are positively related to prevailing excess demand for payers. In contrast, the difference in future returns of payers and nonpayers would be negatively related to this demand—if demand for payers is currently so high that they are relatively overpriced, their *future* returns will be relatively low.

We then examine whether the aggregate rate of dividend initiations and omissions are connected to these dividend premium proxies. The results on initiations are the strongest. Each of the four dividend premium proxies is a significant predictor of the initiation rate. The lagged dividend premium variable by itself explains a remarkable 60% of the annual variation in the initiation rate between 1963 and 2000. We also find that when the initiation rate increases by one standard deviation, returns on payers are lower than nonpayers by nine percentage points per year over the next three years. Conversely, the omission rate increases when the dividend premium variable is low, and when future returns on payers are high. The return predictability results are particularly suggestive of a time-varying mispricing associated with dividends.

At face value, these results suggest that *dividends are highly relevant to share price, but in different directions at different times*. Moreover, the dependence of dividend payment decisions on the dividend premium proxies suggests that managers do cater to time-varying investor demand in an effort to maximize current share price. After a review of alternative hypotheses, we conclude that the results are indeed best explained by this catering dynamic. Explanations based on time-varying firm characteristics, such as investment opportunities or profitability, do not account for the results: the dividend premium variable helps to explain the residual "propensity to initiate" dividends that remains after controlling for changing firm characteristics, including investment opportunities, profits, and firm size using the methodology of Fama and French (2001). Many other features of the data are also inconsistent with this explanation. Alternative hypotheses based on time-varying contracting problems, such as agency or asymmetric information, also do not address many key results, such as the connection between dividend payment and the CU dividend premium or future returns.

We then investigate which source of investor demand creates the time-varying dividend premium that attracts caterers. One possibility is traditional dividend clienteles, such as those discussed in Black and Scholes (1974), which are generated by taxes, transaction costs, or institutional investment constraints. One expects these clienteles to be satisfied by changes in the overall level of dividends, not the number of shares that pay them. But the evidence points the opposite way—initiations and omissions are related to the dividend premium, but the aggregate dividend yield, the aggregate payout ratio, and the aggregate rate of dividend increases are not. We also find that the relation between initiations and omissions and the dividend premium is equally apparent after controlling for plausible proxies for clientele demand.

Another possibility is that the dividend premium variables are driven by sentiment. We tentatively endorse this explanation. One possibility is that when the dividend premium is high, investors are seeking firms that exhibit salient characteristics of safety, including dividend payment; when it is low, investors prefer firms with the characteristics of maximum capital appreciation potential, which means no dividends. This view fits the full set of results well. Further evidence that points to sentiment is the positive correlation between the dividend premium and the closed-end fund discount.

In summary, we develop and test a catering view of dividends that relaxes the market efficiency assumption of the Miller and Modigliani dividend irrelevance proof. The theory thus adds to the collection of theories that relax other assumptions of the proof. It also adds to the literature of behavioral corporate finance. Shefrin and Statman (1984) develop behavioral theories of investor preference for dividends based on self-control problems, prospect theory, and regret aversion. Our paper is closer to the line of research that views managerial decisions as rational responses to security mispricing. For example, Baker and Wurgler (2000) and Baker, Greenwood, and Wurgler (2003) view security issuance decisions as responses to perceived mispricing, and Baker and Wurgler (2002a) develop this into a market timing view of capital structure that relaxes the market efficiency assumption of the capital structure irrelevance proof. Shleifer and Vishny (2003) develop a market timing theory of mergers. Morck, Shleifer, and Vishny (1990), Stein (1996), Polk and Sapienza (2002), and Baker, Stein, and Wurgler (2003) study corporate investment in inefficient capital markets. Graham and Harvey (2001) and Jenter (2001) provide further evidence that managers react to mispricing, or at least the perception of mispricing.

Section I develops the catering theory. Section II presents the main empirical results. Section III considers alternative explanations. Section IV discusses some finer points of the catering interpretation. Section V concludes.

I. A Catering Theory of Dividends

The theory has three basic ingredients. First, it posits a source of uninformed investor demand for firms that pay cash dividends. Second, limits on arbitrage allow this demand to affect current share prices. Third, managers rationally weigh the short run benefits of catering to the current mispricing against the long run costs and then make the dividend payment decision.

A. A Simple Model

A simple static model makes the trade-offs precise. Consider a firm with Q shares outstanding. At $t = 1$, it pays a liquidating distribution of $V = F + \varepsilon$ per share, where ε is normally distributed with mean zero. At $t = 0$, managers have the choice of paying an interim dividend $d \in \{0, 1\}$ per share, which reduces the liquidating value by $d(1 + c)$. The risk-free rate is zero. The cost c captures any trade-off between dividend and investment policy, such as costly external finance or taxes. The Miller and Modigliani case has c equal to zero—dividend policy does not affect the cash flows to investors.

There are two types of investors, category investors and arbitrageurs. Both have constant absolute risk aversion. Category investors care about whether the firm pays dividends. They put dividend payers in a separate investment category, in the spirit of Rosch (1978) and Barberis and Shleifer (2003).[2] There are several potential motivations for such investors. First, market imperfections, such as transaction costs, taxes, and institutional investment constraints, cause traditional dividend "clienteles." Black and Scholes (1974) and Allen, Bernardo, and Welch (2000) develop clientele theories. Second, there is a widespread popular belief that dividend payers are less risky. Naïve investors, such as retirees and those who hold dividend paying stocks for "income" despite the tax penalty, are especially likely to fall prey to this bird-in-the-hand argument.[3] Third, some investors may use dividends to infer managers' investment plans. They may interpret nonpayment, controlling for profitability, as evidence that the firm thinks it has strong opportunities, and take dividends as evidence that opportunities are weaker. Fourth, building on ideas in Thaler and Shefrin (1981), Shefrin and Statman (1984) propose that some investors prefer formal dividends to homemade dividends to combat self-control problems. They also motivate a demand for dividends with prospect theory and regret aversion. In sum, the demand for the category of dividend payers arises from many different sources.

We model the demand of category investors through an irrational expectation of the terminal distribution. We also assume they do not recognize the cost c of a dividend. This irrational expectation introduces a category-level uninformed demand. Specifically, category investors categorize because they view nonpayers as growth firms, and they judge the prospects of those firms relative to their own current assessment of growth opportunities. (This emphasizes the third mechanism in the paragraph above; a similar setup can be motivated by any other reason for categorization.) They expect a liquidating distribution of V^D from payers and V^G from nonpayers. For simplicity we assume that they misestimate the mean, but not the distribution around the mean. They have aggregate risk tolerance per period of $\gamma^C = \gamma$. Typically, their net result is to cause V^D and V^G to fall on opposite sides of F.

Arbitrageurs have rational expectations over the terminal distribution, know the long run cost of an interim dividend, and have aggregate risk tolerance per

[2] Mullainathan (2002) studies a more general model of how categorization affects inference.

[3] Hyman (1988) describes investor reaction to Consolidated Edison's 1974 dividend omission. "It smashed the keystone of faith for investment in utilities: that the dividend is safe and will be paid." (p. 109). Baker, Farrelly, and Edelman (1985) survey managers and find strong agreement with the statement that "Investors have different perceptions of the relative riskiness of dividends and retained earnings." See Brav et al. (2003) for a survey of how managers currently view payout policy.

period of γ^A. Thus they correctly expect a liquidating distribution of F if the firm does not pay a dividend and $F - c$ if it does. The risk aversion of arbitrageurs is how we limit arbitrage, and thus why the uninformed demand of category investors will drive prices from fundamentals. There is ample motivation for assuming limited arbitrage; see Shleifer (2000) for a survey of relevant literature.

With arbitrage limited, the misperceptions of category investors cause the relative prices of payers and nonpayers to differ.[4] In particular, investor group k demands

$$D_0^k = \gamma^k \left(E^k(V) - P_0\right). \tag{1}$$

Prices of dividend payers P^D (cum dividend) and growth firms P^G are therefore

$$P_0 = \begin{cases} P_0^D \equiv \dfrac{\gamma}{\gamma+\gamma^A} V^D + \dfrac{\gamma^A}{\gamma+\gamma^A}(F-c) - \dfrac{Q}{\gamma+\gamma^A}, \text{ and} \\[2ex] P_0^G \equiv \dfrac{\gamma}{\gamma+\gamma^A} V^G + \dfrac{\gamma^A}{\gamma+\gamma^A} F - \dfrac{Q}{\gamma+\gamma^A}. \end{cases} \tag{2}$$

Given these prices, the manager chooses whether to pay dividends. The manager is risk neutral and cares about both the current stock price and the value of total distributions. The manager's only influence on the latter is through the cost c. The manager's horizon, or relative weight on current share price versus long run value, is measured as λ. In practice, λ depends on such factors as the amount of equity and options the manager holds, the timing and terms of the future acquisition of such securities, retirement plans, insider trading opportunities, and so forth.[5] In reduced form, the parameters λ and c capture the basic tension facing the manager. In this short-run inefficient market, he needs to decide which of the two prices to maximize: A short-run price affected by category investor demand, and obtained through catering, and a long-run fundamental value determined by investment policy. The decision depends on his horizon and how much of a trade-off there really is between these two objectives.[6] So he solves

$$\max_d \left(1-\lambda\right) P_0 + \lambda(-dc). \tag{3}$$

[4] Barberis, Shleifer, and Wurgler (2003) and Greenwood and Sosner (2003) find evidence that prices are affected by the categories created by stock indexes.

[5] Conditions under which managers will pursue short-run over long-run value are also discussed by Miller and Rock (1985), Stein (1989), Shleifer and Vishny (1990), Blanchard, Rhee and Summers (1993) and Stein (1996).

[6] An example of a setting in which no tradeoff exists is firm names. Cooper, Dimitrov, and Rau (2001) and Rau et al. (2001) document that when investor sentiment favored the Internet (before March 2000), a number of firms added "dot com" to their names, but when sentiment turned away (after March 2000), many firms changed back.

The solution is straightforward. The manager pays dividends if the premium on dividend payers (call this the dividend premium) is positive and exceeds the present value of the long-run cost that he incorporates. That is, when

$$P_0^D - P_0^G \equiv \frac{\gamma}{\gamma + \gamma^A}(V^D - V^G) - \frac{\gamma^A}{\gamma + \gamma^A}c \geq \left(\frac{\lambda}{1-\lambda}\right)c. \tag{4}$$

The first term in the middle is the immediate positive price impact of switching categories. The second is the immediate negative price impact of the arbitrageurs' recognition of the cost c. To induce payment, the net of these must exceed the long-run cost that the manager incorporates, which is the term on the right. Qualitatively, the propensity to pay dividends is increasing in the dividend premium, decreasing in c, decreasing in the prevalence of arbitrage (the relative risk-bearing capacity of arbitrageurs and category investors), and decreasing in the manager's horizon. The announcement effect of an initiation is positive and increasing in the dividend premium.[7]

Note that this two-category version is too simplistic to incorporate key stylized facts about dividend policy, such as the persistence of dividends. That is, equation (4) shares a feature of many theories of dividends (for example, Miller and Rock (1985)) that the decisions to initiate and omit are symmetric. The model also needs to be extended to address the negative announcement effect of omissions documented by Healy and Palepu (1988) and Michaely, Thaler, and Womack (1995).

A natural way to capture these features is to make use of a third category, *former* payers. These stocks tend to lack any of the characteristics noticed by category investors, as they not only pay no dividends, they also have low past earnings growth.[8] Since they attract only arbitrageurs, their price is $P_0^{FD} = F - \frac{Q}{\gamma^A}$. With this third category, the model can incorporate the persistence of dividends. In particular, the decision for growth firms to initiate is still governed by (4), while current payers *continue* when

$$P_0^D - P_0^{FD} \equiv \frac{\gamma}{\gamma + \gamma^A}\left(V^D - \left(F - \frac{Q}{\gamma^A}\right)\right) - \frac{\gamma^A}{\gamma + \gamma^A}c \geq \left(\frac{\lambda}{1-\lambda}\right)c. \tag{5}$$

Like the propensity to initiate, the propensity to continue is decreasing in the long-run cost and increasing in the dividend premium. The new insight is that

[7] Catering to uninformed demand could explain why dividend changes have price impacts while at the same time appear to contain little new information about future earnings (Lintner (1956), Fama and Babiak (1968), Watts (1973), DeAngelo, DeAngelo, and Skinner (1996), and Benartzi, Michaely, and Thaler (1997)).

[8] The low historical earnings growth can be motivated by assuming that former payers' past dividends were not fully replenished by stock issues (perhaps as a result of the same external finance costs represented by c) or, more intuitively, on empirical grounds. Fama and French (2001) report that dividend payers have average (asset) growth rates of 8.78%, while nonpayers average 11.62% and former payers average only 4.67%. These averages are for the 1963 through 1998 full sample. Between 1993 and 1998, the averages are 6.65%, 17.67%, and 7.61% respectively.

continuing may be desirable even when initiating is not. More formally, if γ^A is small, or if c is small *and* V^G and V^D fall on opposite sides of F, then (5) is satisfied whenever (4) is satisfied. Intuitively, former payers are neglected stocks, attracting only arbitrageurs. Even if initiating is undesirable, current payers may want to continue if the price impact to omitting is large. Note that this third category also suggests why some firms might initiate (reinitiate) dividends even when the dividend premium is negative, and why such initiations would still have a positive announcement effect.

A third category is also useful in addressing the stylized fact that the announcement effect of omissions is negative. Specifically, consider an intermediate time period between $t = 0$ and $t = 1$, in which the neglected former payers face a positive probability of being recategorized as growth firms—for example, because of a random earnings shock. In this setup, dividend payers may choose to omit a dividend at $t = 0$ even when (5) is not satisfied. They suffer a short-run negative announcement effect, but the expected value of being recategorized may be worth it. It is straightforward to formally incorporate this effect.

Of course, we note that there are many other ways to explain these facts. Our goal here is to address them in a model that relaxes the market efficiency (more precisely, unlimited arbitrage) assumption of Miller and Modigliani (1961). Such a model predicts that the propensity to pay is increasing in the dividend premium and decreasing in the fundamental cost of dividends. Other predictions involve managerial horizons and arbitrageur risk-bearing capacity. Realistic variants of the model suggest that the decisions to initiate and to continue should be analyzed separately.

B. Testing the Model

Equation (4) makes a robust prediction that the propensity to pay dividends is increasing in the prevailing dividend premium, which displays its variation in the time series. Thus, although the model is static, we can use time-series variation to test it. We construct proxies for the prevailing dividend premium (i.e., investor demand for dividends), and examine their ability to explain time-series variation in the rates of dividend initiation, continuation, and omission. We leave tests of the interesting cross-sectional predictions of the model, which would examine the effect of managerial horizons and/or costs of arbitrage, to future work.

Time variation in the dividend premium proxies is presumed to reflect time variation in category investor demand. As mentioned above, there are several sources of category-level demand for dividend payers. Under traditional clientele arguments, category-level demand would vary with the imperfections that motivate clienteles. For example, the 1974 Employee Retirement Income Security Act (ERISA) may have increased the attractiveness of payers to pension funds (Del Guercio (1996) and Brav and Heaton (1998)). The 1975 advent of negotiated commissions reduced the transaction cost of creating homemade dividends, and tax code changes can also drive changes in clientele demands.

If investors categorize payers because they believe the bird-in-the-hand argument, on the other hand, their demand for payers will vary with their risk tolerance. If investors categorize payers as firms with slower growth aspirations, they may demand payers when they have pessimistic sentiment about general growth opportunities, and demand nonpayers when their sentiment about growth is generally optimistic. Finally, if the germane considerations in the Shefrin and Statman theory vary over time, they might also lead to a time-varying preference for payers.

At any point in time, the dividend premium depends on the net total of all of these sources of uninformed demand for dividends, and perhaps others. After testing the hypothesis that dividend decisions cater to the current dividend premium, we will go deeper to determine whether the dividend premium can be attached to particular, identifiable sources of demand. We will also examine the validity of an emphasis on a discrete categorization of payers and nonpayers as opposed to the size of payouts.

II. Empirical Tests

A. Dividend Payment Variables

Our measures of dividend payment are derived from aggregations of COMPUSTAT data. The observations in the underlying 1962 to 2000 sample are selected as in Fama and French (2001, pp. 40–41): "The COMPUSTAT sample for calendar year t . . . includes those firms with fiscal year ends in t that have the following data (COMPUSTAT data items in parentheses): Total assets (6), stock price (199) and shares outstanding (25) at the end of the fiscal year, income before extraordinary items (18), interest expense (15), [cash] dividends per share by ex date (26), preferred dividends (19), and (a) preferred stock liquidating value (10), (b) preferred stock redemption value (56), or (c) preferred stock carrying value (130). Firms must also have (a) stockholder's equity (216), (b) liabilities (181), or (c) common equity (60) and preferred stock par value (130). Total assets must be available in years t and $t-1$. The other items must be available in t. . . . We exclude firms with book equity below \$250,000 or assets below \$500,000. To ensure that firms are publicly traded, the COMPUSTAT sample includes only firms with CRSP share codes of 10 or 11, and we use only the fiscal years a firm is in the CRSP database at its fiscal year end. . . . We exclude utilities (SIC codes 4900 to 4949) and financial firms (SIC codes 6000 to 6999)."

We count a firm-year observation as a payer if it has positive dividends per share by the ex date, or else it is a nonpayer. To aggregate this firm-level data into useful time series, two aggregate identities are helpful:

$$Payers_t = New\ Payers_t + Old\ Payers_t + List\ Payers_t, \text{ and} \tag{6}$$

$$Old\ Payers_t = Payers_{t-1} - New\ Nonpayers_t - Delist\ Payers_t. \tag{7}$$

The first identity defines the number of payers and the second describes the evolution. *Payers* is the total number of payers, *New Payers* is the number of initiators among last year's nonpayers, *Old Payers* is the number of payers that also paid last year, *List Payers* is the number of payers this year that were not in the sample last year, *New Nonpayers* is the number of omitters among last year's payers, and *Delist Payers* is the number of last year's payers not in the sample this year. Note that analogous identities hold if one switches "*Payers*" and "*Nonpayers*" everywhere. Also note that lists and delists are with respect to our sample, which involves several screens. Thus new lists include both IPOs that survive the screens in their COMPUSTAT debut as well as established COMPUSTAT firms when they first survive the screens. It also includes the established NASDAQ firms that appear in COMPUSTAT for the first time in the 1970s. Likewise, delists include both delists from COMPUSTAT and firms that fall below the screens.

We then define three variables to capture dividend payment dynamics:

$$Initiate_t = \frac{New\ Payers_t}{Nonpayers_{t-1} - Delist\ Nonpayers_t}, \tag{8}$$

$$Continue_t = \frac{Old\ Payers_t}{Payers_{t-1} - Delist\ Payers_t}, \tag{9}$$

$$Listpay_t = \frac{List\ Payers_t}{List\ Payers_t + List\ Nonpayers_t}. \tag{10}$$

In words, the rate of initiation *Initiate* is the fraction of surviving nonpayers that become new payers. The rate at which firms continue paying *Continue* is the fraction of surviving payers that continue paying. It can also be viewed as one minus the rate at which firms *omit* dividends. The rate at which new lists in the sample pay *Listpay* is self-explanatory.

These variables capture the decision whether to pay dividends, not how much to pay. We take this approach for several reasons. First, these are the natural dependent variables in a theory in which investors categorize shares based on whether they pay dividends (e.g., dividends could be taken to indicate "safety," regardless of their size). Second, as an empirical matter, the payout ratio is sensitive to profitability and the dividend yield is sensitive to changes in share prices. The decision to initiate or omit, in contrast, is always a policy decision. Third, Fama and French (2001) document a decline in the number of payers, and no comparable pattern in the payout ratio. However, measures of the level of dividends do turn out to be useful in discriminating among alternative interpretations for the basic results.

Table I lists the aggregate totals and the dividend payment variables. The initiation rate starts out high in the early years, then drops dramatically in

Table I

Measures of Dividend Payment, 1963 to 2000

Dividend payers, nonpayers, and the rates at which subsamples pay dividends. A firm is defined as a dividend payer at time t if it has positive dividends per share by the ex date (Item 26). A firm is defined as a new dividend payer at time t if it has positive dividends per share by the ex date at time t and zero dividends per share by the ex date at time $t-1$. A firm is defined as an old payer at time t if it has positive dividends per share by the ex date at time t and positive dividends per share by the ex date at time $t-1$. A firm is defined as a new list payer if it has positive dividends per share by the ex date at time t and is not in the sample at time $t-1$. A firm is defined as a nonpayer at time t if it does not have positive dividends per share by the ex date. New nonpayers are firms who were payers at time $t-1$ but not at t. Old nonpayers are firms who were nonpayers in both $t-1$ and t. New list nonpayers are nonpayers at t who were not in the sample at $t-1$. The initiation rate *Initiate* expresses payers as a percentage of surviving nonpayers from $t-1$. The rate at which firms continue paying dividends *Continue* expresses payers as a percentage of surviving payers from $t-1$. The rate at which lists pay *Listpay* expresses payers as a percentage of new lists at t.

	Payers				Nonpayers				Payment Rates %		
Year	Total	New	Old	List	Total	New	Old	List	Initiate	Continue	Listpay
1963	529	21	467	41	149	8	123	18	14.6	98.3	69.5
1964	585	17	519	49	154	6	121	27	12.3	98.9	64.5
1965	681	24	565	92	167	3	119	45	16.8	99.5	67.2
1966	821	16	659	146	238	5	145	88	9.9	99.2	62.4
1967	888	13	793	82	288	14	216	58	5.7	98.3	58.6
1968	954	11	849	94	361	19	263	79	4.0	97.8	54.3
1969	1,018	14	908	96	438	16	330	92	4.1	98.3	51.1
1970	1,048	10	946	92	554	54	406	94	2.4	94.6	49.5
1971	1,030	20	951	59	639	75	502	62	3.8	92.7	48.8
1972	1,281	43	953	285	862	52	568	242	7.0	94.8	54.1
1973	1,627	97	1,221	309	1,127	22	719	386	11.9	98.2	44.5
1974	1,719	130	1,535	54	1,044	44	908	92	12.5	97.2	37.0
1975	1,802	118	1,593	91	1,052	65	853	134	12.2	96.1	40.4
1976	1,878	167	1,670	41	941	58	813	70	17.0	96.6	36.9
1977	1,944	146	1,756	42	821	30	721	70	16.8	98.3	37.5
1978	1,956	96	1,747	113	856	53	651	152	12.9	97.1	42.6
1979	1,925	64	1,761	100	1,046	45	708	293	8.3	97.5	25.4
1980	1,854	58	1,735	61	1,137	68	882	187	6.2	96.2	24.6
1981	1,738	48	1,634	56	1,417	90	962	365	4.8	94.8	13.3
1982	1,631	37	1,545	49	1,621	78	1,210	333	3.0	95.2	12.8
1983	1,523	40	1,434	49	1,929	100	1,380	449	2.8	93.5	9.8
1984	1,450	59	1,346	45	2,111	50	1,605	456	3.5	96.4	9.0
1985	1,378	57	1,282	39	2,133	42	1,698	393	3.2	96.8	9.0
1986	1,270	39	1,176	55	2,373	73	1,744	556	2.2	94.2	9.0
1987	1,214	49	1,112	53	2,651	61	1,971	619	2.4	94.8	7.9
1988	1,185	92	1,057	36	2,563	50	2,123	390	4.2	95.5	8.5
1989	1,162	83	1,041	38	2,432	59	2,036	337	3.9	94.6	10.1
1990	1,148	61	1,053	34	2,403	49	2,011	343	2.9	95.6	9.0
1991	1,128	51	1,052	25	2,497	59	2,015	423	2.5	94.7	5.6
1992	1,140	62	1,036	42	2,674	56	2,085	533	2.9	94.9	7.3
1993	1,148	60	1,043	45	3,049	55	2,342	652	2.5	95.0	6.5
1994	1,163	61	1,059	43	3,286	55	2,634	597	2.3	95.1	6.7
1995	1,165	66	1,068	31	3,416	31	2,772	613	2.3	97.2	4.8
1996	1,153	44	1,061	48	3,774	40	2,924	810	1.5	96.4	5.6
1997	1,101	38	1,027	36	3,784	52	3,110	622	1.2	95.2	5.5
1998	1,042	37	978	27	3,501	35	2,997	469	1.2	96.5	5.4
1999	975	27	916	32	3,320	31	2,806	483	1.0	96.7	6.2
2000	871	30	824	17	3,042	50	2,587	405	1.1	94.3	4.0
Mean	1,247	55	1,119	73	1,693	45	1,336	312	6.1	96.2	27.0
SD	402	38	379	63	1,180	24	974	219	5.0	1.7	22.5

the late 1960s, rebounds in the mid-1970s, drops again in the late 1970s and remains low until the end of the sample. The rate at which firms continue paying varies less, as expected. The rate at which lists pay varies the most. As in Fama and French (2001), it has declined steadily over the past few decades.

We do not focus on the level of dividends, but it is useful to get a rough sense of the aggregate economic size of initiations. In the average year in our sample, newly initiated dividends amount to 0.5% of dividends already paid by payers, and 29% of the change in the amount that is paid by payers (in years when this change is positive). That the first number is so small is not surprising. The numerator is small because the rate of initiation is low and the typical initiator is small and starts off with a small dividend. The denominator is high because the persistence of payment is high and the typical surviving payer tends to increase dividends over time. We also caution that the 29% figure is affected by outlying years in which the change in the amount paid by existing payers is barely positive.

B. Stock Market Dividend Premium Variables

We relate dividend payment choices to several stock market-based measures of the uninformed demand for dividend-paying shares (i.e., the dividend premium). Conceptually, we would like to measure the difference between the market prices of firms that have the same investment policy and different dividend policies. In the frictionless and efficient markets of Miller and Modigliani (1961), of course, this price difference is zero. But with limits to arbitrage, as discussed above, uninformed demand causes a price difference, which may vary over time.

Our first measure, which we simply call "the" dividend premium, is motivated by this intuition. It is the difference in the logs of the average market-to-book ratios of payers and nonpayers—that is, the log of the ratio of average market-to-books.[9] We define market-to-book following Fama and French (2001). Market equity is end of calendar year stock price times shares outstanding (COMPUSTAT Item 24 times Item 25).[10] Book equity is stockholders' equity (Item 216) [or first available of common equity (60) plus preferred stock par value (130) or book assets (6) minus liabilities (181)] minus preferred stock liquidating value (10) [or first available of redemption value (56) or par value (130)] plus balance sheet deferred taxes and investment tax credit (35) if available and minus post retirement assets (330) if available. The market-to-book ratio is book assets minus book equity plus market equity all divided by book assets.

[9] Market-to-book ratios are approximately lognormally distributed. As a result, levels of the market-to-book ratio, unlike logs, have the property that the cross-sectional variance increases with the mean. In our context, this means that the absolute size of a premium measured in levels could proxy for a market-wide valuation ratio.

[10] Here we want an aggregate market-to-book measure for a precise point in time, the end of the calendar year. Later in the paper, when we use market-to-book as a firm characteristic, we use the end of fiscal year stock price.

We then take equal- and value-weighted averages of the market-to-book ratios separately for payers and nonpayers in each year. The final dividend premium series are the difference of the logs of these averages. These series are listed in Table II and the value-weighted series are plotted in Figure 1. The figure shows that the average payer and nonpayer market-to-books diverge significantly at short frequencies. Dividend payers start out at a premium in the first years of the sample. Then nonpayers spike in 1967 and 1968 but fall sharply, in relative terms, through 1972. The dividend premium takes another dip in 1974, and for the past two decades payers appear to have traded at a discount. The discount widened in 1999 but closed somewhat in 2000. In Baker and Wurgler (2002b), we draw on academic histories of the capital market and a review of historical articles in the financial press to provide a much more detailed account of the dividend premium's variation over time.

Keep in mind that a 20% "dividend premium" does not mean that there is a 20% share price boost available by simply initiating a dividend. There is a whole profile of firm characteristics that differs across payers and nonpayers, including profitability and size, and the raw dividend premium variable does not control for the valuation of these characteristics. Moreover, these characteristics (unlike dividends) are less under management control. Another potential disadvantage of the dividend premium variable is that it may also reflect the relative investment opportunities of payers and nonpayers. We will consider this alternative explanation at length in our discussion. (To preview, it does not account for the main results.)

Our second measure is the difference in the prices of Citizens Utilities cash dividend and stock dividend share classes. From 1956 to 1989, the Citizens Utilities Company had two classes of shares. The payouts on them were to be of equal value, as set down in an amendment to the corporate charter. In practice, the relative payouts were close to a fixed multiple. Long (1978) describes the case in detail and finds that the relative price of the two shares fluctuates substantially over time. Long (1978), Poterba (1986), and Hubbard and Michaely (1997) conclude that these fluctuations are not explained by traditional theories of dividends.

We measure the CU dividend premium as the difference in the log price of the cash payout share and the log price of the stock payout share. The 1962 through 1972 data were kindly provided by John Long and the 1973 through 1989 data are from Hubbard and Michaely (1997).[11] Table III reports the CU dividend premium series.

By its nature, the CU dividend premium does not reflect anything about aggregate investment opportunities. This is an advantage over the broader dividend premium variable. On the other hand, the CU premium presents a

[11] There are two further adjustments made throughout the 1962 through 1989 series. The annual value that we consider is the log of the average of the monthly price ratios, because the relative prices fluctuate dramatically even within a year. And to control for the fact that cash dividends were quarterly, in practice, while the stock dividends were semiannual, the cash dividends are assumed to be reinvested until the corresponding stock dividend is paid.

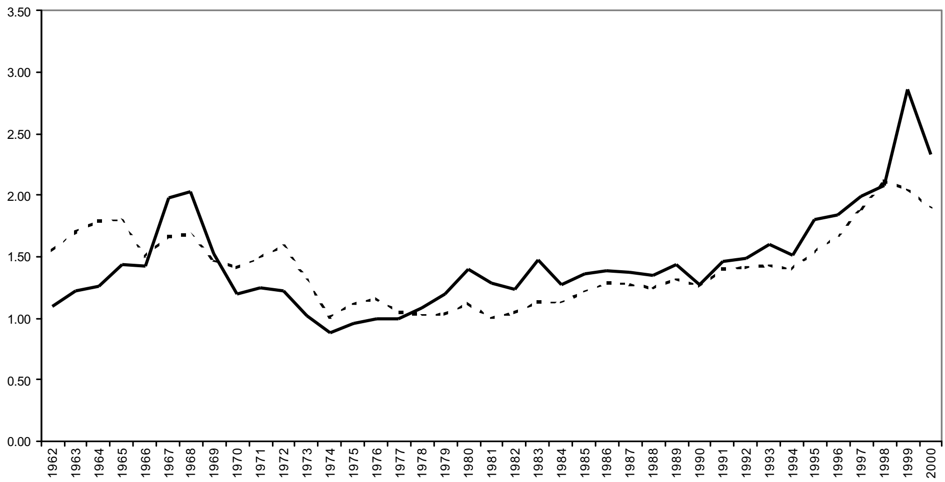
Table II

The Dividend Premium, 1962 to 2000

The market valuations of dividend payers and nonpayers. A firm is defined as a dividend payer at time t if it has positive dividends per share by the ex date (Item 26). The market-to-book ratio is the ratio of the market value of the firm to its book value. Market value is equal to market equity at calendar year end (Item 24 times Item 25) plus book debt (Item 6 minus book equity). Book equity is defined as stockholders' equity (generally Item 216, with exceptions as noted in the text) minus preferred stock (generally Item 10, with exceptions as noted in the text) plus deferred taxes and investment tax credits (Item 35) and post retirement assets (Item 330). The market-to-book ratio reported is an equal-weighed (EW) or value-weighted (VW) average, by book value across dividend payers and nonpayers. These ratios are calculated for the entire sample and for new lists. A firm is defined as a new list if it is not in the sample at time $t-1$. The dividend premium P^{D-ND} is the difference between the logs of the dividend payers' and nonpayers' average market-to-book ratios.

	Payers				Nonpayers				Dividend Premium (P^{D-ND})			
	Total		List		Total		List		Total		List	
Year	EW M/B	VW M/B	EW M/B	VW M/B	EW M/B	VW M/B	EW M/B	VW M/B	EW	VW	EW	VW
1962	1.50	1.55	1.50	1.36	1.19	1.10	1.25	1.12	22.9	34.9	18.6	19.8
1963	1.58	1.70	1.71	1.49	1.30	1.23	1.88	1.71	19.4	32.9	−9.7	−13.8
1964	1.68	1.79	2.09	2.10	1.37	1.26	1.46	1.41	20.1	35.6	35.6	40.0
1965	1.76	1.80	1.60	1.47	1.61	1.43	1.74	1.52	8.8	22.6	−8.5	−3.1
1966	1.52	1.50	1.35	1.20	1.52	1.43	1.55	1.47	0.2	5.4	−14.3	−20.2
1967	1.87	1.66	2.34	1.83	2.36	1.98	3.42	2.65	−23.5	−17.2	−38.0	−36.8
1968	1.99	1.69	2.35	2.89	2.73	2.03	3.32	2.45	−31.7	−18.8	−34.4	16.8
1969	1.60	1.47	1.84	1.67	1.78	1.52	1.90	1.70	−10.4	−3.8	−3.4	−2.1
1970	1.43	1.41	1.51	1.67	1.38	1.20	1.77	1.64	3.1	16.0	−15.6	1.4
1971	1.64	1.50	2.14	2.01	1.48	1.25	2.23	1.90	10.3	18.2	−4.0	5.6
1972	1.62	1.59	1.70	1.74	1.48	1.22	1.84	1.47	9.4	26.6	−8.3	17.0
1973	1.19	1.32	1.27	1.27	1.16	1.02	1.46	1.27	3.2	25.9	−14.1	−0.7
1974	0.93	1.01	1.11	0.91	0.91	0.89	1.08	0.99	2.0	13.2	3.1	−7.6
1975	1.03	1.12	0.90	0.86	1.05	0.95	1.40	1.05	−2.5	15.6	−44.6	−19.9
1976	1.08	1.16	1.37	1.11	1.13	0.99	1.69	1.06	−4.2	15.6	−20.5	4.2
1977	1.06	1.05	1.24	1.23	1.18	1.00	1.32	1.09	−10.7	4.6	−6.3	12.0
1978	1.08	1.03	1.13	1.48	1.34	1.08	1.63	1.24	−22.1	−5.0	−36.5	17.6
1979	1.14	1.04	1.33	0.92	1.75	1.19	2.71	1.61	−43.2	−14.3	−71.6	−55.6
1980	1.25	1.12	1.87	1.20	2.33	1.40	3.86	1.69	−61.9	−22.1	−72.5	−34.2
1981	1.15	1.01	1.46	1.11	1.87	1.29	2.69	1.88	−48.2	−24.9	−61.2	−53.3
1982	1.23	1.05	1.37	1.32	2.03	1.24	3.14	2.05	−50.1	−16.9	−82.6	−44.1
1983	1.41	1.14	1.76	1.21	2.31	1.48	3.18	1.85	−49.3	−26.2	−59.1	−42.9
1984	1.31	1.13	1.72	1.47	1.79	1.28	2.29	1.41	−31.7	−12.5	−28.6	3.5
1985	1.43	1.21	1.64	0.91	2.00	1.36	3.07	1.82	−33.2	−11.0	−62.8	−68.6
1986	1.53	1.29	1.93	1.44	2.27	1.39	3.61	1.74	−39.7	−7.3	−63.0	−18.5
1987	1.47	1.28	1.85	1.53	2.03	1.38	2.83	1.55	−32.4	−7.8	−42.4	−1.4
1988	1.48	1.24	1.47	1.38	1.94	1.35	3.04	1.48	−27.2	−7.8	−72.8	−7.1
1989	1.54	1.32	1.51	1.25	1.97	1.44	3.08	1.61	−24.9	−8.7	−71.2	−25.3
1990	1.39	1.26	1.79	1.80	1.76	1.27	2.27	1.19	−23.5	−1.0	−23.5	41.4
1991	1.59	1.40	1.31	1.24	2.32	1.47	3.45	1.50	−37.8	−4.6	−96.8	−19.4
1992	1.63	1.41	2.03	1.34	2.23	1.49	2.82	1.72	−31.1	−5.3	−32.8	−25.1
1993	1.68	1.43	1.74	1.38	2.33	1.60	2.96	1.82	−33.1	−11.5	−53.2	−27.4
1994	1.55	1.40	1.48	1.47	2.04	1.51	2.59	1.82	−27.6	−7.5	−55.7	−21.7
1995	1.64	1.55	1.83	1.86	2.57	1.80	3.64	2.02	−44.7	−15.1	−68.6	−8.0
1996	1.69	1.67	2.05	1.88	2.41	1.84	3.03	2.09	−35.5	−9.4	−39.0	−10.7
1997	1.86	1.89	1.83	1.52	2.35	1.99	3.02	2.22	−22.9	−4.8	−50.1	−38.0
1998	1.79	2.12	1.98	2.21	2.22	2.09	3.57	2.17	−21.8	1.4	−59.0	1.9
1999	1.68	2.05	1.40	1.34	3.54	2.86	7.97	3.41	−74.9	−33.2	−173.6	−93.0
2000	1.65	1.90	2.18	1.48	2.26	2.33	3.03	1.69	−31.5	−20.6	−33.1	−13.3
Mean	1.48	1.42	1.66	1.48	1.88	1.45	2.64	1.69	−21.1	−0.7	−40.4	−13.6
SD	0.26	0.30	0.35	0.40	0.55	0.41	1.19	0.47	23.7	18.0	36.7	27.8

Panel A. Average market-to-book ratio of dividend payers (dashed line) and nonpayers (solid line)



Panel B. The dividend premium %

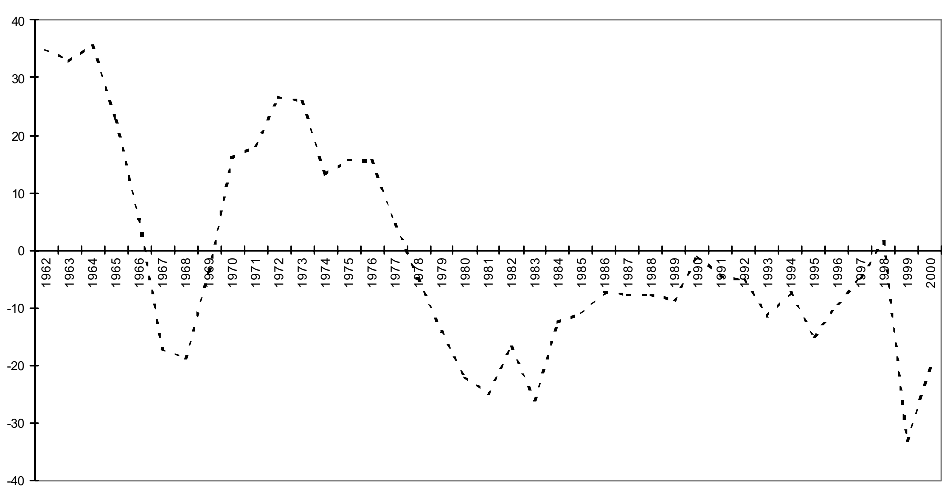


Figure 1. Valuation of dividend payers and nonpayers and the dividend premium, 1962 to 2000. The value-weighted average market-to-book ratio for dividend payers and nonpayers and the dividend premium (the log difference in average market-to-book ratios). A firm is defined as a dividend payer at time *t* if it has positive dividends per share by the ex date (Item 26). The market-to-book ratio is the ratio of the market value of the firm to its book value. Market value is equal to market equity at calendar year end (Item 24 times Item 25) plus book debt (Item 6 minus book equity). Book equity is defined as stockholders' equity (generally Item 216, with exceptions as noted in the text) minus preferred stock (generally Item 10, with exceptions as noted in the text) plus deferred taxes and investment tax credits (Item 35) and post retirement assets (Item 330). The average market-to-book ratios are constructed by value-weighting (by book value) across dividend payers and nonpayers and are plotted in Panel A. Panel B plots the log difference between the market-to-book ratio of payers and nonpayers.

Table III

The Citizens Utilities Dividend Premium and Market Reactions to Dividend Initiations, 1962 to 2000

The Citizens Utilities (CU) price ratio is the log of the ratio of the within-year average cash dividend class share price to the within-year average stock dividend class share price. The 1962 through 1972 data are from Long (1978) and the 1973 through 1989 data are from CRSP. A firm is defined as a new dividend payer at time t if it has positive dividends per share by the ex date (Item 26) at time t and zero dividends per share by the ex date at time $t-1$. We take the first dividend declaration date from CRSP in the 12-month period prior to the fiscal year ending in t. We calculate the sum of the differences between the firm return and the CRSP value-weighted market return for a three-day window $[-1, +1]$ around the declaration date. The announcement effect A scales this return by the standard deviation of the excess returns between 120 calendar days and five trading days before the declaration date. The test statistic from Campbell, Lo, and MacKinlay (1997, equation 4.4.24) is shown in braces and tests the null hypothesis of zero average price reaction in year t.

	CU Dividend Premium	Initiation Announcement Effect			
Year	P^{CU}	N	Excess Return	A	[t-stat]
1962	0.96	1	5.40	1.75	[1.73]
1963	0.98	17	1.94	0.47	[1.92]
1964	1.00	21	1.70	0.41	[1.85]
1965	1.00	21	1.43	0.40	[1.81]
1966	1.00	10	−0.84	−0.23	[−0.73]
1967	0.95	10	0.18	0.06	[0.19]
1968	0.97	7	2.20	0.54	[1.40]
1969	0.97	10	1.82	0.37	[1.16]
1970	1.00	8	5.46	0.85	[2.37]
1971	0.96	19	2.08	0.37	[1.60]
1972	0.93	39	2.17	0.51	[3.14]
1973	0.96	112	3.45	0.70	[7.33]
1974	0.99	94	5.92	0.87	[8.34]
1975	0.96	128	5.21	0.77	[8.59]
1976	0.93	128	4.97	1.05	[11.75]
1977	0.91	114	4.28	1.12	[11.82]
1978	0.90	68	4.02	0.79	[6.43]
1979	0.89	43	3.62	0.70	[4.53]
1980	0.87	35	3.50	0.58	[3.38]
1981	0.92	33	3.57	0.89	[5.08]
1982	0.93	22	3.93	0.62	[2.89]
1983	0.81	25	3.49	0.85	[4.24]
1984	0.89	47	2.13	0.42	[2.85]
1985	0.93	34	1.25	0.35	[2.04]
1986	1.00	31	3.17	0.51	[2.80]
1987	0.92	50	1.38	0.16	[1.15]
1988	0.86	65	2.11	0.48	[3.86]
1989	0.84	50	3.68	0.78	[5.50]
1990	.	46	5.85	0.74	[4.96]
1991	.	31	5.20	0.63	[3.50]
1992	.	46	2.53	0.50	[3.39]
1993	.	42	0.55	0.06	[0.41]
1994	.	51	0.94	0.21	[1.50]
1995	.	44	1.81	0.39	[2.58]
1996	.	18	6.24	0.86	[3.61]
1997	.	20	2.35	0.52	[2.33]
1998	.	19	0.93	0.20	[0.87]
1999	.	17	2.38	0.28	[1.15]
2000	.	10	4.78	0.81	[2.54]
Mean	0.94	41	2.99	0.57	[3.48]
SD	0.05	33	1.75	0.35	[2.87]

long-horizon arbitrage opportunity, so the amount of sentiment that it reflects may be muted. Other disadvantages include the following: CU is just one firm; the stock payout share is more liquid than the cash payout share; there was a one-way, one-for-one convertibility of the stock payout class to the cash payout class, truncating the ability of the price ratio to reveal pro-cash-dividend sentiment; certain sentiment-based mechanisms outlined above involve categorization of firms rather than shares, so a case in which one firm offers two dividend policies may lead to weaker results; and the experiment ended in 1990, when CU switched to stock payouts on both classes.

Our third measure is the average announcement effect of recent initiations.[12] The idea is that if investors are clamoring for dividends, they may make themselves heard through their reaction to initiations. Asquith and Mullins (1983) find that initiations are greeted with a positive return on average, but they do not study variation over time. We define a dividend initiation as the first cash dividend declaration date in CRSP in the 12 months prior to the year in which the firm is identified as a COMPUSTAT New Payer.[13]

Given an initiation in calendar year t, we calculate the cumulative abnormal return over the three-day window from days -1 to $+1$ relative to the CRSP declaration date as the cumulative difference between the firm return and the CRSP VW index. To control for differences in volatility across firms and time (Campbell, et al. (2001)), we scale each firm's three-day excess return by the square root of three times the standard deviation of its daily excess returns. The standard deviation of excess returns is measured from 120 calendar days through five trading days before the declaration date. Averaging these across initiations in year t gives a standardized, cumulative abnormal announcement return A. To determine whether the average return in a given year is statistically significant, we compute a test statistic by multiplying A by the square root of the number of initiations in year t. This is asymptotically standard normal and has more power if the true abnormal return is constant across securities (Brown and Warner (1980) and Campbell, Lo, and MacKinlay (1997)), a natural hypothesis in our context. Table III reports the average standardized announcement effects by year.

Our last measure of the relative stock market valuation of dividend payers is the difference between the *future* returns on value-weighted indexes of payers

[12] One might prefer an announcement effect variable that combines the reactions to initiations and omissions. That is, when demand for dividend payers is high, initiation effects may be particularly positive and omission effects particularly negative. We could not construct this measure, because CRSP data do not provide precise omission announcement dates. However, Yi Liu provided us with annual average announcement effects following omissions *and* reductions in dividends over the period from 1963 through 1999 from Liu, Szewczyk, and Zantout (2003). The correlation between these returns and the subsequent rate of dividend initiation is a statistically significant -0.54, suggesting that our results would be strengthened with a broader sample of announcement effects.

[13] Since COMPUSTAT payers are defined using fiscal years while CRSP allows us to use calendar years, the resulting asynchronicity means that the number of initiation announcements identified in CRSP for year t does not equal the number of COMPUSTAT New Payers in year t. Another difference arises because the required CRSP data are not always available.

Table IV

Statistics for Demand for Dividend Measures, 1962 to 2000

The first column shows the autocorrelation coefficient, the second column shows a Dickey–Fuller test, and the remaining columns show the correlations among the variables. The dividend premium P^{D-ND} is the difference between the logs of the EW and VW market-to-book ratios for dividend payers and nonpayers. The Citizens Utilities dividend premium P^{CU} is the log of the ratio of the cash dividend class share price to the stock dividend class share price. The initiation announcement effect A is the average standardized excess return in a three-day window [−1, +1] around the first declaration dates by new dividend payers. Future relative returns $r_{Dt+1} - r_{NDt+1}$ is the difference in real returns for value-weighted indexes of dividend payers and nonpayers in year t+1. Future relative returns $R_{Dt+3} - R_{NDt+3}$ is the cumulative difference in future returns from year t+1 through t+3. p-values are in brackets.

			Dividend Premium				Future Returns	
	ρ	Unit Root	VW	EW	P_t^{CU}	A_t	$r_{Dt+1} - r_{NDt+1}$	$R_{Dt+3} - R_{NDt+3}$
VW P_t^{D-ND}	0.82	−1.98	1.00					
	[0.00]	[0.29]						
EW P_t^{D-ND}	0.82	−1.58	0.95	1.00				
	[0.00]	[0.49]	[0.00]					
P_t^{CU}	0.61	−2.00	0.60	0.63	1.00			
	[0.00]	[0.28]	[0.00]	[0.00]				
A_t	0.40	−5.18	0.25	0.18	−0.20	1.00		
	[0.02]	[0.00]	[0.13]	[0.27]	[0.31]			
$r_{Dt+1} - r_{NDt+1}$	0.10	−5.31	−0.21	−0.24	−0.28	0.16	1.00	
	[0.54]	[0.00]	[0.20]	[0.15]	[0.14]	[0.35]		
$R_{Dt+3} - R_{NDt+3}$	0.70	−2.52	−0.54	−0.47	−0.28	−0.19	0.63	1.00
	[0.00]	[0.11]	[0.00]	[0.00]	[0.15]	[0.27]	[0.00]	

and nonpayers. Under the stark model outlined in the previous section, managers rationally initiate dividends to exploit an apparent market mispricing. If this is literally the case, a high rate of initiations should forecast low returns on payers relative to nonpayers as the relative overpricing of payers reverses. The opposite should hold for omissions.

Table IV reports the correlations among these variables. We correlate the first three at year t with the excess return on payers over nonpayers $r_D - r_{ND}$ in year $t + 1$ and the cumulative excess return $R_D - R_{ND}$ from years $t + 1$ through t+3. To the extent that they capture a common factor, we expect the dividend premium, the CU premium, and announcement effects to be positively correlated with each other, and negatively correlated with the future excess returns of payers. The actual correlations are as expected, with two exceptions: 1) The CU premium and the initiation effect are negatively (but insignificantly) correlated, and 2) the initiation effect and one-year-ahead excess returns are positively (but insignificantly) correlated. The dividend premium is correlated with each of the others in the expected direction, however. Perhaps the dividend premium is the single best reflection of the common factor. Particularly noteworthy is the strong correlation between the dividend premium and the CU dividend premium. In any case, given that each measure has its own advantages and disadvantages, it is reassuring that they correlate roughly as expected.[14]

Table IV also reports autocorrelations and Dickey–Fuller tests for unit roots. These shed light on the potential for spurious correlations. The textbook case of spurious correlation involves nonstationary variables, and so before one puts weight on the Dickey–Fuller tests it is worth noting the theoretical reasons why these variables are indeed stationary. For example, if the market-to-book ratio is stationary, the dividend premium cannot grow without bound. However, Table IV shows that one cannot reject a unit root in the dividend premium or the CU dividend premium without using this prior information. A similar logic holds for the dividend payment variables: Each one is mathematically bounded between one and zero, yet one cannot formally reject a unit root (unreported). The practical message here is that we should examine the robustness of regression results to the inclusion of time trends and the robustness of return predictability results to the Stambaugh (1999) bias.

C. Time Series Relationships

Here we relate dividend payment to the stock market measures of demand. Figure 2 starts with a plot of the dividend premium against the raw rate of dividend initiation in the following year. The figure reveals a very strong positive relationship, consistent with the prediction of equation (4). On average, the rate of initiation is 11.0% when the dividend premium is positive and only 3.1% when it is negative. In the first half of the sample, the dividend premium and subsequent initiations move almost in lockstep. The premium then submerges in the late 1970s, leading the rate of initiation down once again. The dividend premium has been negative since around 1978, and the initiation rate has also remained low. In unreported results, we have found that a qualitatively similar figure obtains with the rate of initiations by large firms, or small firms, or firms that have been listed for at least five years.

The figure suggests that the relationship has broken down in the most recent period. We show in the next section that this is partly attributable to the changing sample of firms. After 1980, the sample tilts toward small, unprofitable, high market-to-book firms that are unlikely to initiate dividends regardless of market conditions (Fama and French (2001)).[15]

To examine this relationship formally, Table V regresses dividend payment measures on the lagged demand for dividends measures. For example, we run:

$$Initiate_t = a + bP_{t-1}^{D-ND} + cA_{t-1} + d\,P_{t-1}^{CU} + u_t, \tag{11}$$

[14] We have also considered average ex-dividend day returns as a fifth measure of investor demand. Ex-day returns do vary over time (e.g., Eades, Hess, and Kim (1994)). However, they have less of a category-switching interpretation than our other four measures: A dividend payer seems likely to be viewed as a payer before, during, and after the ex-day.

[15] Another explanation for the apparent breakdown is suggested by equation (4), which predicts a monotonic relation between initiations and the dividend premium only when the latter is positive. From 1978 to the end of our sample, the dividend premium has been persistently negative.

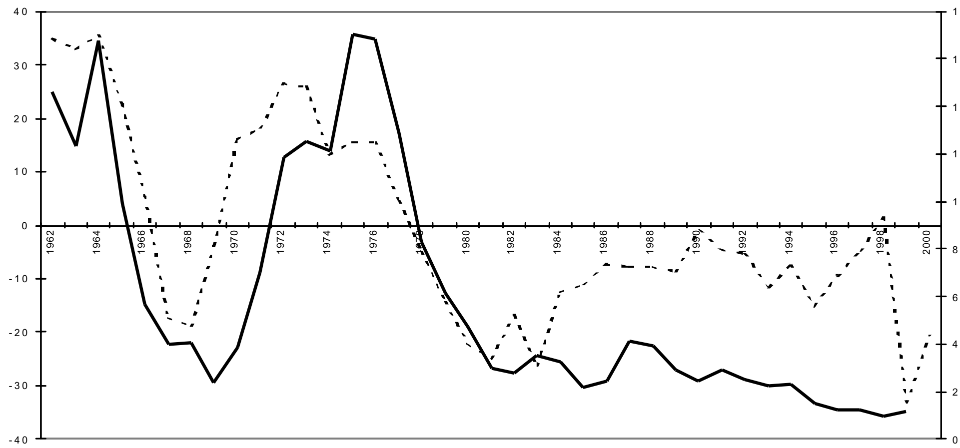


Figure 2. The dividend premium and the rate of dividend initiation, 1962 to 2000. The log difference in the market-to-book ratio of dividend payers and nonpayers (the dividend premium, dashed line—left axis) and one-year-ahead rate of dividend initiations (solid line—right axis). A firm is defined as a dividend payer at time t if it has positive dividends per share by the ex date (Item 26). The initiation rate *Initiate* in $t + 1$ is defined as the percentage rate of new dividend payers at time $t + 1$ among surviving nonpayers from t.

where *Initiate* is the rate of initiation, P^{D-ND} is the market dividend premium (value-weighted or equal-weighted), A is the average initiation announcement effect, and P^{CU} is the Citizens Utilities dividend premium. In the tables, all independent variables are standardized to have unit variance and all standard errors are robust to heteroskedasticity and serial correlation to four lags using the procedure of Newey and West (1987).

Panel A reports the determinants of initiations. The regression in the first column corresponds to Figure 2. It shows that a one-standard-deviation increase in the value-weighted market dividend premium is associated with a 3.90 percentage point increase in the initiation rate in the following year, or roughly three quarters of the standard deviation of that variable.[16] This one variable explains a striking 60% of the time-series variation in the rate of initiation. The second column shows that the effect of the equal-weighted dividend premium is essentially the same.[17] The remaining columns show the effect of other

[16] If nonpayers are trading at a discount to payers, a large number of initiations may mechanically dilute the price of payers and hence lower the premium. This can create the sort of Stambaugh (1999) bias that is described in the Appendix in connection with return predictability. This bias is increasing in the correlation between the errors of the prediction regression in Table V and the errors in an autogression of the dividend premium on the lagged dividend premium. In the case of *Initiate*, these errors have a correlation of less than 0.01, so the bias is inconsequential. In the case of *Continue* and *Listpay*, the correlation is also not statistically significant.

[17] The dependent variable is implicitly an equal-weighted measure, so an equal-weighted independent variable may seem appropriate. On the other hand, the value-weighted premium, which emphasizes larger firms, might be more "visible" to potential initiators. In any case, the two measures perform almost identically in this and future tables.

Table V

Dividend Payment and Demand for Dividends: Basic Relationships, 1962 to 2000

Regressions of dividend initiation and omission rates on measures of the dividend premium. For example, the initiation rate is modeled in Panel A as:

$$Initiate_t = a + bP_{t-1}^{D-ND} + cA_{t-1} + d\,P_{t-1}^{CU} + u_t$$

The initiation rate *Initiate* expresses payers as a percentage of surviving nonpayers from $t - 1$. The continuation rate *Continue* expresses payers as a percentage of surviving payers from $t - 1$. The rate at which listing firms pay *Listpay* expresses payers as a percentage of new lists at t. The dividend premium P^{D-ND} is the difference between the logs of the EW and VW market-to-book ratios for dividend payers and nonpayers. The announcement effects A are the average standardized excess returns in a three-day window [−1, +1] around the declaration dates of new dividend payers. The Citizens Utilities dividend premium P^{CU} is the log of the ratio of the annual average cash dividend class share price to the annual average stock dividend class share price. The independent variables are standardized to have unit variance. t-statistics use standard errors that are robust to heteroskedasticity and serial correlation up to four lags.

	(1)	(2)	(3)	(4)	(5)
			Panel A: $Initiate_t$		
VW P_{t-1}^{D-ND}	3.90				3.80
	[6.56]				[10.74]
EW P_{t-1}^{D-ND}		3.63			
		[5.10]			
P_{t-1}^{CU}			1.70		−0.52
			[2.21]		[−0.82]
A_{t-1}				2.15	1.06
				[2.51]	[1.52]
N	38	38	28	38	28
R^2	0.60	0.52	0.11	0.18	0.70
			Panel B: $Continue_t$		
VW P_{t-1}^{D-ND}	0.85				1.00
	[2.83]				[2.59]
EW P_{t-1}^{D-ND}		0.93			
		[2.96]			
P_{t-1}^{CU}			0.44		−0.25
			[1.02]		[−0.61]
A_{t-1}				0.03	−0.24
				[0.09]	[−0.87]
N	38	38	28	38	28
R^2	0.26	0.30	0.06	0.00	0.25
			Panel C: $Listpay_t$		
VW P_{t-1}^{D-ND}	16.08				10.11
	[6.29]				[2.12]
EW P_{t-1}^{D-ND}		18.15			
		[7.12]			
P_{t-1}^{CU}			14.74		8.16
			[4.68]		[1.64]
A_{t-1}				2.98	−0.28
				[0.58]	[−0.11]
N	38	38	28	38	28
R^2	0.51	0.65	0.47	0.02	0.63

variables, and the results of a multivariate horse race. The lagged initiation announcement effect and the CU premium have significant positive coefficients, as predicted.[18] However, they disappear in a multivariate regression that includes the dividend premium. This is consistent with earlier indications that the dividend premium best captures the common factor.

Panel B reports regressions for the rate of continuation. The dividend premium effect is again as predicted by catering: When dividends appear to be at a stock market *discount*, payers are more likely to *omit* (not continue). The dividend premium effect is smaller here, consistent with the persistence suggested in certain versions of the model. Specifically, a one-standard-deviation increase in the dividend premium increases the continuation rate by 0.85 percentage points. Indeed, since many omissions are forced by low profitability (which we control for in the next section) it may be surprising that the effect is as strong as it is. The other measures do not have explanatory power for the rate of continuation, however.

Panel C shows that the rate at which lists are payers is also positively related to the dividend premium. A one-standard-deviation increase in the dividend premium increases *Listpay* by 16.08 percentage points. The relative size of the coefficient here again reflects the relative variation in the dependent variable. Using a dividend premium variable defined just over recent new lists has at least as much explanatory power (unreported). The CU premium also has a strong univariate effect here, but as before the dividend premium wins a horse race.

Table VI shows the relationship between dividend policy and our fourth demand/dividend premium proxy, future excess returns of payers over nonpayers. In Panel A, the dependent variable is the difference between the returns on value-weighted indexes of payers and nonpayers. Panels B and C look separately at the returns on payers and nonpayers, respectively, to examine whether results for relative returns are indeed coming from the difference in returns, which the theory emphasizes, or payer or nonpayer returns alone. Each panel examines one, two, and three-year ahead returns, and cumulative three-year returns. The table reports ordinary least-squares coefficients as well as coefficients adjusted for the small-sample bias in Stambaugh (1999). The p-values reported in the table represent a two-tailed test of the hypothesis of no predictability using a bootstrap technique described in the Appendix.

Panel A indicates that dividend decisions have strong predictive power for relative returns. A one-standard-deviation increase in the rate of initiation forecasts a decrease in the relative return of payers of eight percentage points in the next year, and 30 percentage points over the next three years. This is a substantial magnitude—arguably, a magnitude worth catering to. The predictive power of the standardized continuation rate is similar. The rate at which lists pay has no predictive power, however, unless a time trend is included, in which

[18] The results are sensitive to the inclusion of the 1963 lagged initiation announcement effect, which is calculated from a single observation. However, a broader measure of market reactions from Liu et al. (2003) that includes dividend omissions and reductions is reliably significant.

Table VI

Dividend Payment and Demand for Dividends: Predicting Returns, 1962 to 2000

Univariate regressions of future excess returns of dividend payers over nonpayers on the initiation rate, the continuation rate, and the rate at which listing firms pay. The dependent variable in Panel A is the difference in real returns between dividend payers r_D and nonpayers r_{ND}. The dependent variable in Panel B is real return of dividend payers r_D. The dependent variable in Panel C is the real return of nonpayers r_{ND}. R_{t+k} denotes cumulative returns from $t+1$ through $t+k$. The initiation rate *Initiate* expresses new payers as a percentage of surviving nonpayers from $t-1$. The continuation rate *Continue* expresses continuing payers as a percentage of surviving payers from $t-1$. The rate at which listing firms pay *Listpay* expresses new COMPUSTAT lists who are payers as a percentage of new COMPUSTAT lists. In the *Listpay* specification, a year trend is included in the regression. The independent variables are standardized to have unit variance. We report OLS coefficients and bias-adjusted (BA) coefficients. Bootstrap p-values represent a two-tailed test of the null hypothesis of no predictability.

		$Initiate_t$				$Continue_t$				$Listpay_t$ (detrended)			
	N	OLS	BA	[p-val]	R^2	OLS	BA	[p-val]	R^2	OLS	BA	[p-val]	R^2
Panel A: Relative Returns													
$r_{Dt+1} - r_{NDt+1}$	37	−7.68	−6.54	[0.15]	0.10	−7.68	−7.97	[0.06]	0.10	−6.13	−6.87	[0.16]	0.07
$r_{Dt+2} - r_{NDt+2}$	36	−13.27	−12.63	[0.01]	0.31	−7.90	−8.20	[0.07]	0.11	−9.47	−9.49	[0.03]	0.15
$r_{Dt+3} - r_{NDt+3}$	35	−8.81	−8.79	[0.06]	0.14	−5.90	−6.13	[0.17]	0.07	−7.08	−7.49	[0.09]	0.08
$R_{Dt+3} - R_{NDt+3}$	35	−30.54	−28.23	[0.06]	0.47	−21.62	−23.63	[0.13]	0.25	−24.88	−23.91	[0.10]	0.28
Panel B: Payer Returns													
r_{Dt+1}	37	−4.06	−4.39	[0.29]	0.06	−2.14	−2.41	[0.46]	0.02	−3.11	−2.15	[0.36]	0.03
r_{Dt+2}	36	−0.95	−1.67	[0.79]	0.00	0.70	0.54	[0.85]	0.00	−3.00	−2.51	[0.31]	0.03
r_{Dt+3}	35	−1.87	−2.28	[0.60]	0.01	1.12	0.92	[0.73]	0.00	−2.88	−2.85	[0.34]	0.02
R_{Dt+3}	35	−8.08	−10.71	[0.39]	0.10	−0.16	−0.23	[0.99]	0.00	−8.19	−4.83	[0.37]	0.09
Panel C: Nonpayer Returns													
r_{NDt+1}	37	3.62	2.26	[0.64]	0.01	5.54	5.76	[0.38]	0.03	3.01	4.93	[0.62]	0.01
r_{NDt+2}	36	12.32	11.02	[0.07]	0.13	8.60	8.73	[0.16]	0.07	6.47	6.83	[0.26]	0.03
r_{NDt+3}	35	6.94	6.54	[0.31]	0.04	7.02	7.27	[0.28]	0.04	4.20	4.80	[0.47]	0.01
R_{NDt+3}	35	22.46	17.45	[0.23]	0.18	21.47	24.81	[0.21]	0.17	16.70	19.88	[0.36]	0.09

case it displays a similar level of predictability to the other dividend policy variables. The bottom panels confirm that the relative return predictability cannot be attributed to just payer or nonpayer predictability. As theory suggests, it is the relative return that matters most.

These tables provide support for the catering theory's main prediction. Firms are more likely to initiate when the dividend premium is apparently positive, and more likely to omit when it is apparently negative.

III. Alternative Explanations

The catering explanation for these results is that dividend payment is, to some extent, a rational managerial response to investor demand pressures that cause a stock market mispricing. While it is often possible to reinterpret an individual empirical relationship, it turns out to be difficult to construct a coherent, non-catering alternative explanation for the full set of results. We discuss a variety of alternative explanations in this section.

A. Statistical Robustness

One question is whether these relationships simply represent a common time trend caused by forces outside the theory. Relative stock returns would not be predictable from a time trend, of course, but the other dividend premium proxies are worth checking. Table VII includes a trend (the calendar year) alongside the broad dividend premium. The coefficient remains strongly significant for initiations. For continuations, however, inclusion of a trend pushes the coefficient to the 10% level of significance, and greatly reduces the size of the coefficient on new lists but does not eliminate its statistical significance.

In unreported results, we include a trend alongside the CU dividend premium and the initiation announcement effect. This changes earlier inferences only in the case of the CU dividend premium—it does not have explanatory power beyond a common trend. We have also considered the raw (unstandardized) average initiation announcement effect, which we did not examine earlier. It turns out to have a positive but insignificant univariate relationship with initiations. However, it is significant in the presence of a trend term.

B. Time-varying Investment Opportunities

We now turn to economics-based alternative explanations. The relationship in Figure 2 could be an artifact of time variation in investment opportunities. That is, nonpayers may be initiating dividends not because they are chasing the relative premium on payers but because their investment opportunities are low in an absolute sense. An inverse relationship between dividends and investment opportunities could follow if external finance is costly, as in Myers (1984) and Myers and Majluf (1984), or if dividends are a response to agency costs of free cash flow, as in Jensen (1986). This is a natural alternative explanation that we consider in detail.

Table VII

Dividend Payment and the Dividend Premium: Other Controls, 1962 to 2000

Regressions of dividend payment rates on measures of the dividend premium, growth opportunities, the personal tax advantage of dividends versus capital gains, and a time trend. For example, the initiation rate is modeled in Panel A as:

$$Initiate_t = a + bP^{D-ND}_{t-1} + c\frac{M}{B_{t-1}} + d\,\frac{D}{P_{t-1}} + eTax_{t-1} + f\,Year_{t-1} + u_t$$

The initiation rate *Initiate* expresses payers as a percentage of surviving nonpayers from $t-1$. The continuation rate *Continue* expresses payers as a percentage of surviving payers from $t-1$. The rate at which listing firms pay *Listpay* expresses payers as a percentage of new lists at t. The dividend premium P^{D-ND} is the difference between the logs of the VW market-to-book ratios for dividend payers and nonpayers. The VW market-to-book ratio M/B is averaged across nonpayers in Panel A, payers in Panel B, and new lists in Panel C. The VW dividend yield D/P is from CRSP. *Tax* is the ratio of after-tax income from a dollar in dividends to a dollar in long-term capital gains. *Year* is the calendar year. All independent variables but *Year* are standardized to unit variance. t-statistics use standard errors that are robust to heteroskedasticity and serial correlation up to four lags.

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: $Initiate_t$			
VW P^{D-ND}_{t-1}	2.83	2.71	2.87	4.19	3.66	3.90
	[5.39]	[5.42]	[5.42]	[6.53]	[7.65]	[4.56]
VW *Nonpayer* M/B_{t-1}	−1.92	−1.34	−1.32			
	[−2.43]	[−2.54]	[−2.32]			
VW D/P_{t-1}				1.63	0.95	0.96
				[3.05]	[1.90]	[2.13]
Tax_{t-1}		1.48	1.72		1.37	1.74
		[3.22]	[2.28]		[2.64]	[2.02]
$Year_{t-1}$			0.03			0.05
			[0.40]			[0.52]
N	38	38	38	38	38	38
R^2	0.70	0.77	0.77	0.70	0.75	0.76
			Panel B: $Continue_t$			
VW P^{D-ND}_{t-1}	0.79	0.57	0.45	0.83	0.56	0.40
	[2.64]	[2.30]	[1.75]	[2.64]	[2.19]	[1.56]
VW *Payer* M/B_{t-1}	0.30	0.50	0.48			
	[1.05]	[2.02]	[2.28]			
VW D/P_{t-1}				−0.16	−0.50	−0.50
				[−0.82]	[−1.85]	[−2.43]
Tax_{t-1}		0.60	0.39		0.68	0.43
		[2.37]	[1.62]		[2.47]	[2.05]
$Year_{t-1}$			−0.03			−0.04
			[−0.78]			[−0.95]
N	38	38	38	38	38	38
R^2	0.29	0.39	0.40	0.27	0.38	0.39

Table VII—*Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel C: $Listpay_t$			
VW P^{D-ND}_{t-1}	16.88	13.86	5.85	16.35	10.92	2.60
	[7.75]	[7.31]	[3.55]	[5.67]	[5.39]	[2.62]
VW $List\ M/B_{t-1}$	2.89	5.36	3.93			
	[0.76]	[2.85]	[3.13]			
VW D/P_{t-1}				1.54	−5.34	−5.54
				[0.47]	[−2.51]	[−5.57]
Tax_{t-1}		12.29	−0.19		13.99	1.05
		[6.50]	[−0.11]		[6.36]	[0.97]
$Year_{t-1}$			−1.67			−1.79
			[−7.37]			[−15.74]
N	38	38	38	38	38	38
R^2	0.53	0.79	0.95	0.52	0.78	0.96

A first point is that this explanation makes the converse prediction that payers will be more likely to *omit* when their investment opportunities are *high*. This would imply a negative relationship between the dividend premium and the rate at which firms continue paying, not positive as found earlier. Therefore, at best, this alternative hypothesis could apply only to initiations.

To examine the investment opportunities explanation (for initiations), a straightforward test is to simply control for the *level* of investment opportunities and see if the dividend *premium* retains residual explanatory power. We consider two potential measures of investment opportunities: the average market-to-book of the firms in question; and the overall CRSP value-weighted dividend yield. The first and fourth columns in Table VII show the results. For initiations, the investment opportunities proxies enter with the predicted signs—nonpayers are less likely to initiate when their average market-to-book is high, and when the overall dividend-price ratio is low. For continuations and new lists, however, these variables enter with the wrong sign. Most important, the dividend premium coefficient is not much affected.

As another check, we look at the oil industry.[19] This offers a unique experiment. The oil price shock in 1973 presumably improved investment opportunities in oil and gas even as it generally reduced them in other industries. The investment opportunities hypothesis therefore suggests a falling rate of initiation in oil and gas firms at this time, even as the overall rate rises. We examine the rate of initiation in the oil industry as defined by Fama and French (1997). Because this time series is more sensitive to idiosyncratic decisions, the overall correlation with the dividend premium is lower at 0.28. But the same pattern emerges. The average oil industry initiation rate is 8.1% following positive dividend premium years and 3.9% following negative dividend premium years.

[19] We thank Randall Morck for this suggestion.

Interestingly, the period following 1973 *reinforces* this pattern: After several years of no new payers, the rate of initiation was 7.7% in 1973, 23.1% in 1974, 12.5% in 1975, and 22.2% in 1976.

Note also that the investment opportunities view makes similar predictions for both repurchases and dividends, while catering involves only the latter. Thus we can examine whether the rate of repurchase is also related to the dividend premium, or only the rate of dividend initiation. We construct aggregate time series measures of the rate of repurchase, defining a repurchase as nonzero purchase of common and preferred stock (COMPUSTAT Item 115). The first useable year is 1972. We find that the rate of repurchase among all firms, and the rate at which firms "initiate" repurchases (new repurchasers in year t divided by surviving nonrepurchasers), have an insignificant negative correlation with the lagged dividend premium (unreported). The dividend initiation rate, by contrast, has a correlation of 0.73 over the same 29-year period.

A related point is that time-varying investment opportunities lead more naturally to variation in the *level* of dividends, not the number of firms paying a dividend as is captured in initiations and omissions. Thus, this alternative hypothesis would predict that the dividend premium should bear an even stronger relationship to the level of dividends, whereas catering to category investors would not necessarily predict a relationship in levels. Consistent with the catering view, we take updated data from Shiller (1989) on earnings and dividends for the S&P 500 and the CRSP value-weighted dividend yield, and find that neither the payout ratio nor the dividend yield is significantly correlated with the lagged dividend premium (unreported). Also note that we control for the dividend yield directly in the last three columns of Table VII. Doing so actually increases the effect of the dividend premium on the initiation rate.

Finally, the investment opportunities hypothesis cannot account for the connection between dividend payment and future relative returns or the CU dividend premium. In summary, it appears that investment opportunities are not driving the results.

C. Correlated Errors in Forecasting Investment Opportunities

The second alternative explanation we consider is a variant of the first and is suggested by the referee. Perhaps managers and investors make correlated errors in their forecasts. That is, investors sometimes get excited about growth prospects and bid up the price of nonpayers, who they feel are better suited to exploit new opportunities. Managers, rather than catering to this sentiment, are equally smitten and choose to invest all available resources rather than paying dividends. This story is better than the rational expectations version outlined above in that it can address the return predictability results, but otherwise it has the same list of shortcomings.

D. Time-varying Characteristics

Another possibility is that our demand measures are somehow related to the cross-sectional *distribution* of dividend-relevant characteristics *within* payer and nonpayer samples. This is a more general version of the investment opportunities explanation discussed above. As a contrived example, suppose the variance of investment opportunities among nonpayers increases (for some reason) whenever the dividend premium increases. Then an increasing initiation rate could indicate that a relatively high fraction of nonpayers do not need to retain cash, rather than nonpayers as a group catering to the dividend premium. In such a situation, the average investment opportunities of nonpayers are held constant, so the time series exercises in Table VII would still mistakenly attribute the effect to the dividend premium.

We evaluate this explanation by controlling directly for sample characteristics. In particular, we examine whether the dividend premium helps to explain the residual variation in dividend decisions after controlling for the characteristics studied by Fama and French (2001). They model the expected probability that a firm is a payer as a function of four variables:

$$\Pr(Payer_{it} = 1) = \text{logit}\left(a + bNYP_{it} + c\frac{M}{B_{it}} + d\frac{dA}{A_{it}} + e\frac{E}{A_{it}}\right) + u_{it}, \quad (12)$$

where size *NYP* is the NYSE market capitalization percentile, that is, the percentage of firms on the NYSE having smaller capitalization than firm i in that year. Market-to-book M/B is measured as defined previously, with the slight modification that here we use the fiscal year closing stock price (COMPUSTAT Item 199) instead of the calendar year close. Growth dA/A in book assets (COMPUSTAT Item 6) is self-explanatory. Profitability E/A is earnings before extraordinary items (18) plus interest expense (15) plus income statement deferred taxes (50) divided by book assets. The error term u is the residual propensity to pay dividends for a given firm-year.

The tests proceed in two stages. In the first stage, we follow Fama and French in estimating firm-level logit regressions using these firm characteristics. As before, we examine dividend payment separately among surviving nonpayers, surviving payers, and new lists. We also follow them in estimating specifications that exclude M/B—they suggest that the degree to which this variable measures investment opportunities may change over time, and indeed we have been arguing that this variable is affected by uninformed investor demand.

In the second stage, we regress the average annual prediction errors, or the aggregate "propensity to pay," on the value-weighted dividend premium. For example, naming $\widehat{PTI}$ the residual rate of initiation or the "propensity to initiate," we estimate:

$$\begin{aligned} \widehat{PTI}_t &= f + gP_{t-1}^{D-ND} + v_t, \text{ where} \\ \widehat{PTI}_t &\equiv \frac{1}{N}\sum_i u_{it}. \end{aligned} \quad (13)$$

Explanatory power for the propensity to initiate (or, analogously, the propensity to continue $\widehat{PTC}$ or propensity to list as a payer $\widehat{PTL}$) means that the dividend premium is not affecting dividend policy through the average *or* the cross-sectional distribution of any of these four characteristics.[20] The regression in (13) is analogous to our earlier time series regressions, such as equation (11), but now the effect of varying characteristics has been explicitly removed. Note that the two-stage approach gives deference to the firm characteristics variables by allowing the dividend premium to explain only residual variation. Also, in terms of statistical power, note that the dividend premium is using only 38 data points to fit, not thousands like the characteristics.

Table VIII shows the results of this exercise. The first stage results indicate that size and profitability have the most robust effects on the propensity to pay, as Fama and French find. The right column shows the second stage results. In general, controlling for characteristics directly, the dividend premium retains statistically significant explanatory power for most subsamples. Comparing these coefficients to our earlier results in Table V, we see that controlling for firm characteristics barely affects the initiation rate coefficient. This is compelling evidence that the dividend premium effect is not working through a background correlation with the level *or* the distribution of characteristics.

Figure 3 shows that controlling for characteristics also improves the post-1980 correlation between initiations and the dividend premium. This suggests that the raw rate was depressed in recent decades by the influx of small, unprofitable, high market-to-book firms noted by Fama and French. Within the language of the model, firms with these characteristics would tend to have high fundamental costs of paying dividends. Controlling for characteristics may better reveal the partial effect of the dividend premium.

Interestingly, the only period where the rate of initiation is sharply lower than would be expected from the dividend premium is the early 1970s. This gap was apparent from a close inspection of Figure 2, but is particularly clear here in Figure 3. A possible explanation for this is the dividend component of Nixon's wage and price ceilings. Dann (1981) and Bagwell and Shoven (1989) point out that repurchases spiked in 1973 and 1974. Some firms who might have initiated dividends were apparently persuaded to repurchase instead. But once the controls were removed, initiations shot back up as indicated by the dividend premium.

Controlling for characteristics does tend to reduce the effect of the dividend premium among the other samples, however. That characteristics would help to explain omissions might be expected given that they are known to be associated with such firm-level characteristics as low profitability. Nevertheless, the dividend premium approaches statistical significance even in this sample, and remains significant in the new list sample.

The methodology in Table VIII also confirms once again that our empirical results, like our theory, are mainly about the decision *whether* to pay dividends,

[20] Including the dividend premium directly in equation (12) and estimating the coefficients in a panel regression gives qualitatively similar results to our two-stage procedure (unreported). A panel regression is necessary in that specification because the dividend premium does not vary within a year, as the Fama–MacBeth procedure requires.

Table VIII

Dividend Payment and the Dividend Premium: Firm Characteristics Controls, 1963 to 2000

Two-stage regressions of dividend payment on firm characteristics and the dividend premium. The first stage performs a set of Fama–MacBeth logit regressions of dividend payment on firm characteristics for each of three samples: surviving payers, surviving nonpayers, and new lists:

$$\Pr\left(Payer_{it} = 1\right) = \text{logit}\left(a + bNYP_{it} + c\frac{M}{B_{it}} + d\frac{dA}{A_{it}} + e\frac{E}{A_{it}}\right) + u_{it}$$

The second stage regresses the average annual prediction errors (actual policy minus predicted policy) from the logit regressions on the dividend premium. For example, for the sample of surviving nonpayers we estimate the residual rate of initiation, or the "propensity to initiate dividends" *PTI*, and model it as:

$$P\tilde{T}I_t = f + g\,P_{t-1}^{D-ND} + v_t, \text{ where } P\tilde{T}I_t \equiv \frac{1}{N}\sum_i u_{it}$$

The first two rows examine the propensity to initiate *PTI* and so restrict the sample to surviving nonpayers. The next two rows examine the propensity to continue *PTC* and so restrict the sample to surviving payers. The last two rows examine the propensity to list as a payer *PTL* and so restrict the sample to new lists. The firm characteristics are the NYSE percentile *NYP*, the market-to-book ratio M/B (which is excluded in some specifications), asset growth dA/A, and profitability E/A. The NYSE percentile is the percentage of firms listed on the NYSE that are equal to or smaller in terms of market capitalization. The market-to-book ratio is the ratio of the market value of the firm to its book value. Market value is equal to market equity at calendar year end (Item 24 times Item 25) plus book debt (Item 6 minus book equity). Book equity is defined as stockholders' equity (generally Item 216) minus preferred stock (generally Item 10) plus deferred taxes and investment tax credits (Item 35) and post retirement assets (Item 330). Asset growth is the change in assets (Item 6) over assets. Profitability is earnings before extraordinary items (Item 18) plus interest expense (Item 15) plus income statement deferred taxes (50) over assets. The dividend premium P^{D-ND} is the difference between the logs of the VW market-to-book ratios for dividend payers and nonpayers. t-statistics in the second stage regression use standard errors that are robust to heteroskedasticity and serial correlation up to four lags.

	NYP_t		M/B_t		dA/A_t		E/A_t		$\text{VW}\,P_{t-1}^{D-ND}$	
	b	[t-stat]	c	[t-stat]	d	[t-stat]	e	[t-stat]	g	[t-stat]
PTI	1.54	[10.65]	−0.85	[−6.62]	0.21	[1.13]	9.60	[10.00]	3.35	[6.66]
PTI	0.90	[6.13]			−0.23	[−1.48]	6.47	[12.42]	3.76	[6.71]
PTC	4.57	[10.09]	0.33	[1.31]	1.50	[5.15]	15.06	[5.87]	0.34	[1.73]
PTC	4.61	[10.63]			1.37	[4.96]	14.20	[6.01]	0.32	[1.56]
PTL	4.56	[40.95]	−0.78	[−15.86]	−0.84	[−6.44]	10.76	[11.67]	11.20	[5.51]
PTL	3.88	[37.16]			−1.19	[−8.64]	7.80	[13.06]	12.78	[6.99]

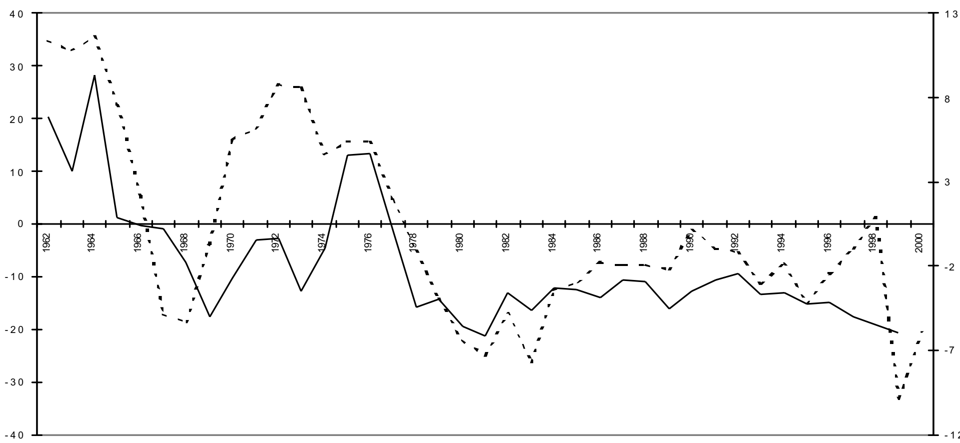


Figure 3. The dividend premium and the propensity to initiate, 1962 to 2000. The log difference in the market-to-book ratio of dividend payers and nonpayers (the dividend premium, dashed line—left axis) and one-year-ahead propensity to initiate dividends (solid line—right axis). A firm is defined as a dividend payer at time t if it has positive dividends per share by the ex date (Item 26). The propensity to initiate $\widetilde{PTI}$ in $t + 1$ (the rate of initiation controlling for prevailing firm characteristics in t+1) is estimated following Fama and French (2001). First, we estimate a set of annual Fama–MacBeth logit regressions of dividend payment on firm characteristics over the sample of surviving nonpayers, using as firm characteristics the NYSE market capitalization percentile *NYP*, the market-to-book ratio M/B, asset growth dA/A, and profitability E/A. Second, we estimate the propensity to initiate as the annual average prediction error (actual initiation rate minus predicted rate) of these regressions.

not *how much* to pay. That is, we have constructed a time series of the raw rate of dividend *increases* and found that it has a significant positive correlation with the dividend premium (unreported), but that this result comes entirely from changing characteristics like profitability. When these characteristics are accounted for using the two-stage procedure, there is no relationship between the residual propensity to increase dividends and the dividend premium (unreported).

Finally, we can also ask whether the average annual prediction errors from Table VIII predict the relative returns of payers and nonpayers. In other words, whether non-characteristics-related variation in dividend payment, which is presumably more likely to be motivated by catering, also predicts returns. We confirm that the average prediction errors have comparable or greater predictive power than the raw dividend payment measures (unreported).

E. Time-varying Contracting Problems

Another class of alternative explanations involves time-varying contracting problems, such as adverse selection or agency. With regard to adverse selection, one might propose that when nonpayers trade at a low value, this is a particularly important time for them to signal their investment opportunities. Initiating dividends serves as a signal in the models of Bhattacharya (1979), Hakansson (1982), John and Williams (1985), and Miller and Rock (1985). Once again, a natural way to evaluate this hypothesis is to control for the level of nonpayer market-to-book directly, and examine whether the dividend premium has residual explanatory power. Table VII shows that it does. Moreover, it is hard to imagine a rational expectations equilibrium model in which dividend payment decisions predict future stock returns, or would have any natural reason to depend on the CU dividend premium.

Agency costs may also vary over time, with high agency costs requiring dividend payments. For example, La Porta et al. (2000) find that dividend policy varies across countries according to the degree of investor protection. If the dividend premium were a simple time trend, this could be a more compelling explanation. As it stands, this explanation requires governance to improve briefly in the late 1960s, deteriorate, and then improve again.[21] Of course, it is possible that variation in investment opportunities and profits might affect agency costs, but this would be addressed in Table VIII. Given those results, one must imagine agency problems that arise and vary over independent of firm characteristics. And again, this hypothesis does not address the time to the CU dividend premium connections between dividend payment and future returns.

IV. Discussion

Process of elimination, plus the close connection between the results and predictions of the model, suggests that managers are, at least to some extent, catering to investor demand. In this section we try to flesh out the picture. We discuss the nature of managerial motivations and the source of the investor demand behind the dividend premium.

A. Are Managers Smart?

Facts such as strong return predictability, and indeed the very existence of the CU dividend premium, suggest that there is a market mispricing at play here. However, it does not necessarily follow that managers are "smart" in the stark sense of the model. Our results are consistent with an explicit attempt to capture a perceived mispricing, as in the model, but also with a somewhat softer view of catering. Managers may just cater to, or even be forced by proxy vote to meet, extreme investor demands in general, and mispricing is merely a symptom of extreme investor demand. In this interpretation managers are not knowingly outwitting the market. Their decisions will still generate return predictability, but they are not explicitly designed to capture mispricing.

It is obviously difficult to distinguish between these two closely related interpretations. In the ".com" name changes studied by Cooper et al. (2001) and Rau et al. (2001), it is fairly easy to believe that managers were motivated by

[21] One interpretation is that *perceptions* of agency costs vary over time, for example in a naïve response to rather than in a rational anticipation of corporate misconduct. But in this view, dividend initiation is an attempt to cater to investor sentiment rather than to solve genuinely time-varying contracting problems.

mispricing. The long-run returns evidence on equity issues and repurchases also has a compelling market timing flavor, and indeed the essence of the CFO's job is to minimize the cost of capital. In the setting of dividends, on the other hand, the smart managers view is harder to envision. So for now, we take both views as related cases of an underlying catering dynamic, and leave the task of distinguishing between them to future work.

B. Are Managers Catering to Clienteles?

We turn to the *source* of time-varying demand for dividend payers. Who are managers catering to? Black and Scholes (1974) suggest that uninformed demand for dividends can result from dividend clienteles, which in turn derive from such imperfections as taxes, transaction costs, or institutional investment constraints.[22] Of course, in the Miller (1977)- style equilibrium they describe, managers compete so aggressively that a nontrivial dividend premium or discount never arises. As a result, for any given firm, dividend policy is kept irrelevant at the margin. The evidence we have seen so far has more of a disequilibrium flavor, of course, but it may still reflect catering to traditional clienteles, albeit of a time-varying structure.

One important observation here is that rational clienteles would be satisfied by a supply response in the aggregate level of dividends, not the number of dividend-paying shares. Also, if they are diversified, rational clienteles will not care about how the supply response is distributed across firms. Indeed, Marsh and Merton (1987) point out that current dividend payers, with high financial slack and modest investment opportunities, are probably the lowest marginal cost suppliers of dividends. Thus if the dividend premium were being driven by changes in the structure of rational clienteles, it should have a closer connection to the level of dividends than to the number of payers. We find the opposite.

Another approach is to see if we can directly match up the dividend premium with any plausible proxies for clienteles or dividend payment behavior. A natural proxy for tax clienteles, for example, is the relative tax advantage of dividend income versus capital gains. Figure 1 suggests that the 1986 Tax Reform Act, which should have shrunk the anti-dividend tax clientele, had no visible effect on the dividend premium.[23] Similarly, Hubbard and Michaely (1997) study the reaction of the CU dividend premium to the 1986 reform. They conclude that tax-motivated clienteles do not seem to affect that variable. On the other hand, the evidence from Graham and Kumar (2003) suggests some degree of tax-related clientele trading activity.

Table VII contains a more formal test of whether firms are catering to tax clienteles. The personal tax advantage for dividends (typically a net disadvantage) is measured as the ratio of the after-tax income from a dollar of dividends

[22] Miller and Scholes (1978) propose that tax code changes could have no influence, because taxes on dividends can be postponed indefinitely. However, Peterson, Peterson, and Ang (1985) find empirically that most investors do not avoid taxation.

[23] As an aside, the lack of a differential reaction to the reform by payers and nonpayers also seems inconsistent with dividend tax capitalization.

to a dollar of long-term capital gains—that is, one minus the average marginal income rate, divided by one minus the average marginal long-term capital gains rate. The tax rates in this calculation are weighted average rates across shareholder groups as calculated by the NBER TAXSIM model. They are reported at www.nber.org/~taxsim/mrates/mrates2.html and described by Feenberg and Coutts (1993). Table VII shows that if anything, the initiation rate is positively related to this variable, not negatively related, and in any case its inclusion does not much affect the dividend premium coefficient. (In Panel C, the large *t*-statistic on taxes disappears when a trend is included because of trends in both the rate at which new lists pay and the tax advantage variable.)

Transaction costs also vary over time, changing the cost of homemade dividends. Perhaps this induces changes in demand by transaction cost clienteles. Black (1976) dismisses this argument, pointing out that there are simple institutional solutions to the problem of the small investor's transaction costs. On the other hand, Jones (2001) shows that transaction costs have declined dramatically since the mid-1970s, which coincides with a reduction in the rate of initiation.[24] Jones's Figure 4 shows the average annual one-way transaction cost for the NYSE, or one half of the bid-ask spread plus commissions. This series is strongly positively correlated with the rate of initiation, though this comes mostly from a common time trend; the correlation between the detrended variables is not significant (unreported). More importantly, in regressions that include both variables, the dividend premium has more statistical significance than transaction costs in explaining the initiation rate (unreported).

Another possibility is that dividend clienteles are motivated by institutional investment constraints. For instance, the 1974 Employee Retirement Income Security Act (ERISA) may have increased the pro-dividend clientele by creating a vague "prudent man" rule for pension funds. The law was revised in 1979 to allow pension funds to provide venture capital, thus erasing any doubt that nonpayers were acceptable investments and perhaps shrinking the dividend clientele. Figure 2 could be broadly consistent with these institutional shifts. However, the dividend premium seems to anticipate the law, peaking in 1972 and beginning its drop in 1977. Perhaps ERISA is part of the story in this period. Gompers and Metrick (2001) document the general rise in institutional ownership in the 1980s and 1990s. However, we are not aware of any specific investment constraints that could explain the dividend premium over the 1960s and early 1970s.

Finally, the rational clientele explanations for the dividend premium face some difficulty accounting for the magnitude of the return predictability effects. Even under limited arbitrage, in equilibrium the marginal clientele investor should still be indifferent to leaving the clientele or taking advantage of the mispricing that his colleagues presumably induce. But the marginal clientele investor's savings on transaction costs or taxes, for example, would seem

[24] The rise of mutual funds roughly coincides with these falling transaction costs, potentially lowering an individual investor's cost of monetizing capital gains further still.

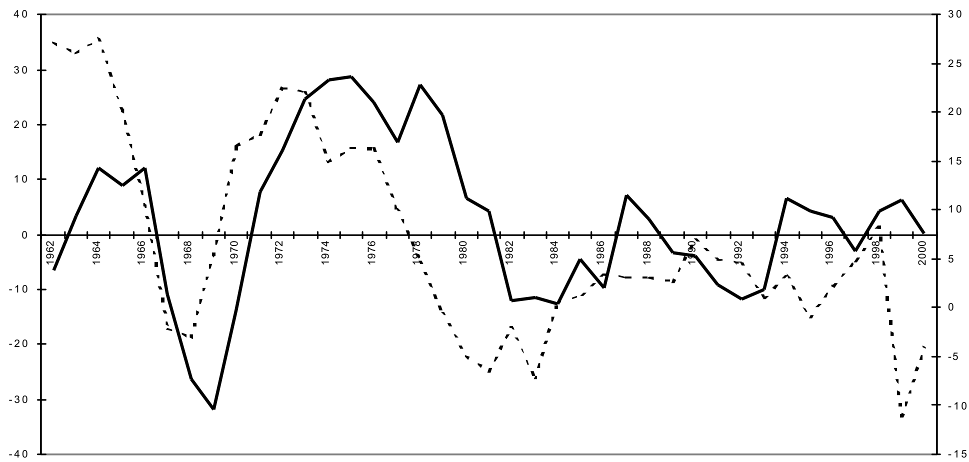


Figure 4. The dividend premium and the closed-end fund discount. The log difference in the market-to-book ratio of dividend payers and nonpayers (the dividend premium, dashed line—left axis) and the closed-end fund discount (solid line—right axis). The value-weighted closed-end fund discount uses data on net asset values and market prices for general equity and convertible funds from Neal and Wheatley (1998) for 1962 to 1993, from CDA/Wiesenberger for 1994 to 1998, and from *The Wall Street Journal* for 1999 to 2000.

unlikely to be worth a sacrifice of nine percentage points per year in pre-tax expected returns.

C. Are Managers Catering to Investor Sentiment?

The discussion above suggests that traditional clienteles are not driving the dividend premium proxies. This leaves sentiment as the remaining possibility. Of course, economists are just beginning to study sentiment, and so such explanations are harder to reject by construction. Here we attempt to provide a few rejectable hypotheses. We start by comparing the dividend premium to the closed-end fund discount. Zweig (1973) and Lee, Shleifer, and Thaler (1991) view the closed-end discount as a measure of investor sentiment. Of course, whether the discount reflects risk tolerance, expectations for growth stocks, or both, has not been established. But documenting a connection is still useful because such a connection would not be predicted by any of the nonsentiment explanations we have considered.

We gather value-weighted discounts on closed-end stock funds for 1962 through 1993 from Neal and Wheatley (1998), for 1994 through 1998 from CDA/Wiesenberger, and for 1999 and 2000 from the discounts on stock funds reported in *The Wall Street Journal* in the turn-of-the-year issues. Figure 4 shows the relationship between the dividend premium and the closed-end fund discount. They are not perfectly synchronous, but they are visibly related. The contemporaneous correlation is 0.37 with a p-value of 0.02.

This provides some initial support for a sentiment-based explanation. To tie it back to our basic results, Table IX uses the closed-end fund discount as an instrumental variable for the dividend premium. The table also uses

Table IX

Dividend Payment and the Dividend Premium: Instrumental Variables, 1962 to 2000

Instrumental variables estimates of the effect of the dividend premium on dividend payment rates. For example, the initiation rate is modeled in Panel A as:

$$Initiate_t = a + b\widehat{P}_{t-1}^{D-ND} + u_t, \text{ where}$$

$$\widehat{P}_{t-1}^{D-ND} = c + dCEFD_{t-1} + e_1 r_{CGt-1} + e_2 r_{CGt-2} + e_3 r_{CGt-3} + f\left(R_{Dt+3} - R_{NDt+3}\right) + v_{t-1}$$

The initiation rate *Initiate* expresses payers as a percentage of surviving nonpayers from $t-1$. The continuation rate *Continue* expresses payers as a percentage of surviving payers from $t-1$. The rate at which listing firms pay *Listpay* expresses payers as a percentage of new lists at t. The dividend premium VW P^{D-ND} is the difference between the logs of the value-weighted average market-to-book ratios of dividend payers and nonpayers. We instrument for the dividend premium with the value-weighted closed-end fund discount, past nominal capital gains on the value-weighted CRSP index, and future cumulative relative returns of payers over nonpayers. The value-weighted closed-end fund discount uses data on net asset values and market prices for general equity and convertible funds from Neal and Wheatley (1998) for 1962 to 1993, from CDA/Wiesenberger for 1994 to 1998, and from *The Wall Street Journal* for 1999 to 2000. r_{CGt-k} denotes capital gains in year $t-k$. R_{t+k} denotes cumulative future returns from $t+1$ through $t+k$. The OLS t-statistics use standard errors that are robust to heteroskedasticity and serial correlation up to four lags. The 2SLS t-statistics use standard errors that are robust to heteroskedasticity.

		2SLS			
	OLS	CEFD	Past Capital Gains	Future Returns	All
			Panel A: $Initiate_t$		
VW P_{t-1}^{D-ND}	3.90	8.31	6.25	5.51	5.42
	[6.56]	[3.86]	[3.85]	[4.71]	[6.60]
N	38	38	38	36	36
R^2	0.60	–	–	–	–
			Panel B: $Continue_t$		
VW P_{t-1}^{D-ND}	0.85	1.63	0.65	1.60	1.15
	[2.83]	[1.82]	[1.25]	[2.71]	[3.52]
N	38	38	38	36	36
R^2	0.26	–	–	–	–
			Panel C: $Listpay_t$		
VW P_{t-1}^{D-ND}	16.08	8.72	26.99	11.47	17.69
	[6.29]	[1.10]	[4.28]	[1.97]	[5.34]
N	38	38	38	36	36
R^2	0.51	–	–	–	–

lagged capital gains and future relative returns on payers and nonpayers as instruments. The logic for using future relative returns is that they are arguably a purer (though perhaps noisier) measure of sentiment for dividend payers (and firms with associated characteristics) than the dividend premium itself.[25] Recent capital gains on the market could reflect various sentiment mechanisms—after a crash, unsophisticated investors may tend more toward the "bird in the hand" rationale, for example, or they may become convinced that general growth opportunities are bleak. By either logic, they would be more inclined to hold dividend payers.

Table IX shows that the instrumental variables coefficients are, in general, about as strong as the basic OLS coefficients. For the specification that uses future returns as an instrument, the coefficient table merely puts the earlier predictability results in units of the dividend premium. For the other specifications, the results are more novel. At a minimum, they confirm that that the specific component of the dividend premium associated with these variables helps to explain rates of initiation and omission, thus casting doubt on generic "omitted third factor" alternative explanations. To the extent that the instruments pick up investor sentiment, the results provide affirmative support for a sentiment interpretation.

V. Conclusion

We propose a view of dividends that is based on relaxing the market efficiency assumption of the dividend irrelevance proof. It adds to the collection of dividend theories that relax other assumptions of the proof. The essence of the catering theory is that managers give investors what they currently want. In the case of dividends, catering implies that managers tend to initiate dividends when investors put a relatively high stock price on dividend payers, and tend to omit dividends when investors prefer nonpayers. A simple model formalizes the key trade-offs involved and offers testable predictions.

Our empirical work focuses on the prediction that the rates of dividend initiation and omission depend on the current "dividend premium," or the difference between the current stock prices of payers and nonpayers. We test this prediction by forming four stock price-based proxies for the dividend premium. We find that the aggregate initiation rate is significantly positively related to all four of our proxies. (In one case, this does not amount to more than a common trend.) In addition, the rate of omission is significantly negatively related to two of the four dividend premium proxies. After reviewing alternative hypotheses, we conclude that catering is the most natural explanation. The results suggest that dividends are highly relevant to share value, but in different directions at different times. Moreover, managers apparently recognize and cater to shifts in investor demand for dividend payers.

We then inquire about the source of investor demand for dividends. We do not find much evidence that it springs from traditional dividend clienteles.

[25] We thank Lubos Pastor for suggesting that we use past capital gains in this manner.

Instead, sentiment appears to be a key factor. This is suggested in the connection between the closed-end fund discount and the dividend premium variable, and in instrumental variables estimates of the effect of the dividend premium on dividend payment. In Baker and Wurgler (2002b), we review academic histories of the capital markets and historical news articles to better understand why investor attitudes toward dividends have changed over time.

Finally, we remind the reader that both the model and the empirical results are about the discrete decision whether to pay dividends, not how much to pay. Once dividends are initiated, increases and decreases appear to be governed more by firm-level profitability than by the relative valuations of payers and nonpayers. Thus in terms of aggregate economic significance, catering explains the number of payers but not the total payouts by existing payers. As a result, we suggest that catering be taken as a building block in an overall descriptive theory of dividend policy.

Appendix

This appendix describes the simulations that generate the bias-adjusted coefficients and p-values reported in Table VI. As discussed by Stambaugh (1999), a small-sample bias arises when the explanatory variable is persistent and there is a contemporaneous correlation between innovations in the explanatory variable and stock returns. For example, in the following system

$$R_t = a + bX_{t-1} + u_t \tag{A1}$$

$$X_t = c + dX_{t-1} + v_t, \tag{A2}$$

the bias is equal to

$$E[\hat{b} - b] = \frac{\sigma_{uv}}{\sigma_v^2} E[\hat{d} - d], \tag{A3}$$

where the hats represent OLS estimates. Kendall (1954) shows the OLS estimate of d has a negative bias. The bias for OLS b is therefore of the opposite sign to the sign of the covariance between innovations in dividend policy and returns.

For us, the sign of this covariance is not obvious *a priori* (unlike when the predictor is a scaled-price variable). To address the potential for bias and conduct inference, we use a bootstrap estimation technique. The approach is identical to Baker and Stein (2002) and is similar to that used in Vuolteenaho (2000), Kothari and Shanken (1997), Stambaugh (1999), and Ang and Bekaert (2001). For each regression in Table VI, we perform two sets of simulations.

The first set generates a bias-adjusted point estimate. We simulate (A1) and (A2) recursively starting with X_0, using the OLS coefficient estimates, and drawing with replacement from the empirical distribution of the errors u and v. We throw out the first 100 draws (to draw from the unconditional distribution of X), then draw an additional N observations, where N is the size of the original sample. (For the cumulative three-year regressions, the number of additional draws is one third the size of the original sample, since it contains overlapping returns.) With each simulated sample, we re-estimate (A1). This gives us a set of coefficients b^*. The bias-adjusted coefficient BA reported in Table VI subtracts the bootstrap bias estimate (the mean of b^* minus the OLS b) from the OLS b.

In the second set of simulations, we redo everything as above under the null hypothesis of no predictability—that is, imposing b equals zero. This gives us a second set of coefficients b^{**}. With these in hand, we can determine the probability of observing an estimate as large as the OLS b by chance, given the true $b = 0$. These are the p-values in Table VI.

REFERENCES

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

Allen, Franklin, and Roni Michaely, 2002, Payout policy, Working paper, University of Pennsylvania.

Ang, Andrew, and Geert Bekaert, 2001, Stock return predictability: Is it there? NBER Working paper 8207.

Asquith, Paul, and David W. Mullins, Jr., 1983, The impact of initiating dividend payments on shareholders' wealth, *Journal of Business* 56, 77–96.

Bagwell, Laurie Simon, and John B. Shoven, 1989, Cash distributions to shareholders, *Journal of Economic Perspectives* 3, 129–140.

Baker, H. Kent, Gail E. Farrelly, and Richard B. Edelman, 1985, A survey of management views on dividend policy, *Financial Management* 14, 78–84.

Baker, Malcolm, and Jeremy Stein, 2003, Market liquidity as a sentiment indicator, *Journal of Financial Markets*, forthcoming.

Baker, Malcolm, Robin Greenwood, and Jeffrey Wurgler, 2003, The maturity of debt issues and predictable variation in bond returns, *Journal of Financial Economics* 70, 261–291.

Baker, Malcolm, Jeremy Stein, and Jeffrey Wurgler, 2003, When does the market matter? Stock prices and the investment of equity-dependent firms, *Quarterly Journal of Economics* 118, 969–1006.

Baker, Malcolm, and Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns, *Journal of Finance* 55, 2219–2257.

Baker, Malcolm, and Jeffrey Wurgler, 2002a, Market timing and capital structure, *Journal of Finance* 57, 1–32.

Baker, Malcolm and Jeffrey Wurgler, 2002b, Why are dividends disappearing? An empirical analysis, *Journal of Financial Economics*, forthcoming.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2003, Comovement, *Journal of Financial Economics*, forthcoming.

Benartzi, Shlomo, Roni Michaely, and Richard Thaler, 1997, Do changes in dividends signal the future or the past? *Journal of Finance* 52, 1007–1034.

Bhattacharya, S., 1979, Imperfect information, dividend policy, and 'the bird in the hand fallacy,' *Bell Journal of Economics* 10, 259–270.

Black, Fischer, 1976, The dividend puzzle, *Journal of Portfolio Management* 2, 5–8.

Black, Fischer, and Myron S. Scholes, 1974, The effects of dividend yield and dividend policy on common stock prices and returns, *Journal of Financial Economics* 1, 1–22.

Blanchard, Olivier, Chanyong Rhee, and Lawrence Summers, 1993, The stock market, profit, and investment, *Quarterly Journal of Economics* 108, 115–136.

Brav, Alon, John R. Graham, Campbell Harvey, and Roni Michaely, 2003, Payout policy in the 21st century, Working paper, Duke University.

Brav, Alon, and J. B. Heaton, 1998, Did ERISA's prudent man rule change the pricing of dividend omitting firms? Working paper, Duke University.

Brown, Stephen, and Jerold Warner, 1980, Measuring security price performance, *Journal of Financial Economics* 8, 205–258.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk, *Journal of Finance* 56, 1–44.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, (Princeton University Press, Princeton, NJ).

Cooper, Michael J., Orlin Dimitrov, and P. Raghavendra Rau, 2001, A rose.com by any other name, *Journal of Finance* 56, 2371–2388.

Dann, Larry Y., 1981, Common stock repurchases: An analysis of returns to bondholders and stockholders, *Journal of Financial Economics* 9, 113–138.

DeAngelo, Harry, Linda DeAngelo, and Douglas J. Skinner, 1996, Dividend signaling and the disappearance of sustained earnings growth, *Journal of Finance* 40, 341–371.

Del Guercio, Diane, 1996, The distorting effect of the prudent-man laws on institutional equity investments, *Journal of Financial Economics* 40, 31–62.

Eades, Kenneth M., Patrick J. Hess, and E. Han Kim, 1994, Time-series variation in dividend pricing, *Journal of Finance* 49, 1617–1638.

Fama, Eugene F., and Harvey Babiak, 1968, Dividend policy: An empirical analysis, *Journal of the American Statistical Association* 53, 1132–1161.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60, 3–44.

Feenberg, Daniel, and Elizabeth Coutts, 1993, An introduction to the Taxsim model, *Journal of Policy Analysis and Management* 12, 189–194.

Gompers, Paul, and Andrew Metrick, 2001, Institutional investors and equity prices, *The Quarterly Journal of Economics* 116, 229–259.

Graham, John R., and Campbell R. Harvey, 2001, The theory and practice of corporate finance: Evidence from the field, *Journal of Financial Economics* 60, 187–244.

Graham, John R., and Alok Kumar, 2003, D. dividend clienteles exist? Evidence on dividend preferences of retail investors, Working paper, Duke University.

Greenwood, Robin, and Nathan Sosner, 2003, Trade and the comovement of stock returns: Evidence from Japan, Working paper, Harvard University.

Hakansson, Nils H., 1982, To pay or not to pay dividends, *Journal of Finance* 37, 415–428.

Healy, Paul M., and Krishna G. Palepu, 1988, Earnings information conveyed by dividend initiations and omissions, *Journal of Financial Economics* 21, 149–176.

Hubbard, Jeff, and Roni Michaely, 1997, Do investors ignore dividend taxation? A reexamination of the Citizens Utilities case, *Journal of Financial and Quantitative Analysis* 32, 117–135.

Hyman, Leonard, 1988, *America's Electric Utilities: Past, Present, and Future* (Public Utility Reports: Arlington, VA).

Jensen, Michael C., 1986, Agency costs of free cash flow, corporate finance and takeovers, *American Economic Review* 76, 323–329.

Jenter, Dirk, 2001, Managerial portfolio decisions and market timing, Working paper, MIT Sloan School.

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

Jones, Charles M., 2001, A century of stock market liquidity and trading costs, Working paper, Columbia University.

Kendall, M. G., 1954, Note on bias in estimation of auto-correlation, *Biometrika* 41, 403–404.

Kothari, S. P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time series analysis, *Journal of Financial Economics* 44, 169–203.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 2000, Agency problems and dividend policies around the world, *Journal of Finance* 55, 1–33.

Lee, Charles M., Andrei Shleifer, and Richard Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75–110.

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

Liu, Yi, Samuel Szewczyk, and Zaher Zantout, 2003, An examination of the long-run market reaction to announcement of dividend omissions and reductions, Working paper, Drexel University.

Long, John B., 1978, The market valuation of cash dividends: A case to consider, *Journal of Financial Economics* 6, 235–264.

Marsh, Terry A., and Robert C. Merton, 1987, Dividend behavior for the aggregate stock market, *Journal of Business* 60, 1–40.

Michaely, Roni, Richard H. Thaler, and Kent L. Womack, 1995, Price reactions to dividend initiations and omissions: Overreaction or drift?, *Journal of Finance* 50, 573–608.

Miller, Merton H., 1977, Debt and taxes, *Journal of Finance* 32, 261–275.

Miller, Merton H., and Franco Modigliani, 1961, Dividend policy, growth and the valuation of shares, *Journal of Business* 34, 411–433.

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

Miller, Merton H., and Myron Scholes, 1978, Dividends and taxes, *Journal of Financial Economics* 6, 333–364.

Morck, Randall, Andrei Shleifer, and Robert Vishny, 1990, The stock market and investment: Is the market a sideshow? *Brookings Papers on Economic Activity* 2, 157–215.

Mullainathan, Sendhil, 2002, Thinking through categories, Working paper, Massachusetts Institute of Technology.

Myers, Stewart, 1984, The capital structure puzzle, *Journal of Finance* 39, 575–592.

Myers, Stewart, and Nicholas Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.

Neal, Robert, and Simon M. Wheatley, 1998, Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis* 33, 523–547.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Peterson, Pamela, David Peterson, and James Ang, 1985, Direct evidence on the marginal rate of taxation on dividend income, *Journal of Financial Economics* 14, 267–282.

Polk, Christopher, and Paola Sapienza, 2002, The real effects of investor sentiment, Working paper, Northwestern University.

Poterba, James M., 1986, The market valuation of cash dividends: The Citizens Utilities case reconsidered, *Journal of Financial Economics* 15, 395–405.

Rau, P. Raghavendra, Ajay Patel, Igor Osobov, Ajay Korana, and Michael J. Cooper, 2001, The game of the name: Value changes accompanying dot.com additions and deletions, Working paper, Purdue University.

Rosch, Eleanor, 1978, Principles of categorization, in Eleanor Rosch and Barbara B. Lloyd, eds.: *Cognition and Categorization* (Lawrence Erlbaum Associates, Hillsdale, NJ).

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

Shiller, Robert J., 1989, *Market Volatility* (MIT Press, Cambridge, MA).

Shleifer, Andrei, 2000, *Inefficient markets: An introduction to behavioral finance* (Oxford University Press, Oxford, UK).

Shleifer, Andrei, and Robert W. Vishny, 1990, Equilibrium short horizons of investors and firms, *American Economic Review Papers and Proceedings* 80, 148–153.

Shleifer, Andrei, and Robert W. Vishny, 2003, Stock market driven acquisitions, *Journal of Financial Economics* 70, 295–312.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

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

Stein, Jeremy C., 1996, Rational capital budgeting in an irrational world, *Journal of Business* 69, 429–455.

Thaler, Richard H., and Hersh M. Shefrin, 1981, An economic theory of self-control, *Journal of Political Economy* 89, 392–406.

Vuolteenaho, Tuomo, 2000, Understanding the aggregate book-to-market ratio and its implications to current equity-premium expectations, Working paper, Harvard University.

Watts, Ross, 1973, The information content of dividends, *Journal of Business* 46, 191–211.

Zweig, Martin E., 1973, An investor expectations stock price predictive model using closed-end fund premiums, *Journal of Finance* 28, 67–87.