THE JOURNAL OF FINANCE • VOL. LV, NO. 6 • DEC. 2000

A Theory of Dividends Based on Tax Clienteles

FRANKLIN ALLEN, ANTONIO E. BERNARDO, and IVO WELCH*

ABSTRACT

This paper explains why some firms prefer to pay dividends rather than repurchase shares. When institutional investors are relatively less taxed than individual investors, dividends induce "ownership clientele" effects. Firms paying dividends attract relatively more institutions, which have a relative advantage in detecting high firm quality and in ensuring firms are well managed. The theory is consistent with some documented regularities, specifically both the presence and stickiness of dividends, and offers novel empirical implications, e.g., a prediction that it is the tax difference between institutions and retail investors that determines dividend payments, not the absolute tax payments.

ALTHOUGH A NUMBER OF THEORIES have been put forward in the literature to explain their pervasive presence,[1] dividends remain one of the thorniest puzzles in corporate finance. In a frictionless world without taxes or transaction costs, dividends and share repurchases are equivalent. If dividends are taxed more heavily than capital gains, as is the case in the United States and many other countries, share repurchases are apparently superior to dividends. Nevertheless, dividends continue to be a substantial proportion of earnings—and personal dividend taxes continue to be a substantial source of income for the I.R.S. For the 1973 to 1983 period, dividends for the largest 1,000 firms in the United States averaged 44 percent of earnings whereas repurchases averaged only 6 percent (see Allen and Michaely (1995)). Although, as Bagwell and Shoven (1989) have stressed, repurchases increased significantly in 1984 and have remained high, repurchases were not a substitute for dividends. From 1984 to 1988, repurchases increased from 6 percent to 38 percent of earnings, but dividends still increased from 44 percent to 51 percent. The I.R.S. *Statistics of Income* publication documents that taxable dividends in adjusted gross income amounted to $82 billion in about

*Franklin Allen is from the Wharton School at the University of Pennsylvania. Antonio E. Bernardo is from the Anderson School at UCLA. Ivo Welch is from the School of Management at Yale University. The authors thank Laurie Hodrick, Shlomo Benartzi, Charles Calomiris, Mark Grinblatt, Kathleen Fitzgerald, Avraham Kamara, Jiang Luo, Hamid Mehran, Avanidhar Subrahmanyam, Richard Vines and Chris Hayden (Georgeson Investor Services) and seminar participants at Northwestern, NYU, UC/Davis, Columbia, UC/Riverside, Wharton, Yale, Atlanta, Duke, Chicago, USC, Stanford, Washington, and Wisconsin for helpful comments.

[1]Because the dividend literature is large, we delay the review of earlier work to Section IV.

25 million individuals' tax returns in 1994.[2] A second puzzling fact about dividends concerns the inclination of firms to "smooth" them, as originally documented in Lintner (1956).

Our paper presents an intuitive and novel explanation that is consistent with these two puzzles: the inclination of firms to pay dividends rather than to repurchase shares, and the inclination of firms to smooth dividends. Our hypothesis is based on two assumptions.

Our first assumption is that there are groups of investors who are taxed differently and who have different incentives to become informed about corporate affairs. In the United States, for example, public and corporate pension funds, colleges and universities, labor unions, foundations, and other corporations are either fully or largely exempt from taxes. The proportion of stocks owned by these groups has been significant for many years and has increased substantially since 1980 (see Allen and Santomero (1998)). For simplicity, we assume there are just two clienteles and call them "untaxed institutions" and "taxed individuals." Because of their scale, the (untaxed) institutions have greater incentives to become informed about the firm. Institutions are more likely than retail investors to conduct "due diligence" to find out whether a particular firm is well run or poorly managed. They are also more likely to facilitate mechanisms by which potential shortcomings are corrected. Such institutional activities can be intrinsic, such as their ability to make a quick sale of a large block of shares to a potential raider. But they can also be proactive. In recent years, institutions have become increasingly involved in corporate governance. Their ability to vote in large blocks and influence corporate leaders has had an influence on such corporate choices as changes in the corporate charter and board composition and membership. (Section IV.C justifies this assumption in more detail.) In economic terms, the presence of institutional shareholdings can be associated with higher firm value because of signaling effects, agency effects, or both. In Section I, we assume that better managers take advantage of the detection ability of institutions; in Section II, we assume that institutions add value by expending resources to monitor the corporation.

Our second assumption is that dividends are one way of attracting institutions. This can be justified by directly appealing to common *institutional charter* and *prudent man rule* restrictions that make it more difficult for many institutions to purchase investments with low dividend payouts (Brav and Heaton (1998)). Yet, many tax-exempt institutions such as universities and charities do not have direct restrictions and still hold significant amounts of dividend-paying stock. We show that the dividend-paying firms' in-equilibrium market prices make them a relatively better purchase for institutions than for retail investors, because dividends are taxed for individuals

[2] Although this I.R.S. number includes dividends on privately held shares, and although it does not measure the marginal dividend tax rate, it does indicate that dividend taxes are nontrivial. For comparison, the equivalent figure for taxable personal interest income was $126 billion.

but untaxed for institutions. This comparative advantage results in an endogenously higher fraction of ownership by institutions for dividend-paying stocks.[3]

Our two assumptions imply that firms can attract more institutions as shareholders by paying dividends and that dividend-paying firms will perform better than otherwise equal non-dividend-paying firms.

In our asymmetric information model, taxable dividends exist to signal that firms' management is "good," because paying dividends increases the chance that firm quality will be detected by the institutions. Bad firms dislike attracting institutions because institutional presence increases the probability that quality will be revealed. They will not find it worthwhile to incur the dividend tax costs to imitate higher-quality firms. Good firms do not fear detection and are willing to have their shareholders incur dividend taxes in order to signal quality.

In our agency model, taxable dividends exist to attract informed institutions whose presence ensures that the firm will remain well run. If management underperforms, then the institutions will, for example, facilitate takeovers by selling large blocks of stock—or they may even become directly involved in the corporate governance process. This model is also consistent with dividend smoothing. Because institutional actions are valuable for the same firms period after period, the same firms are inclined to continue "purchasing" them. Further, dividend reductions would indicate a desire to reduce institutional ownership and the implicit oversight of the firm's affairs. Institutional shareholders are exactly the kind of shareholders able to dislodge the management in such a situation. Thus, it is especially those managers whose firms pay high dividends and have high institutional ownership who would suffer the most dramatic consequences from cutbacks in dividends. In contrast, issuers who choose share repurchases attract a clientele that will be less able to muster resistance to possible future payout cutbacks. Our model is thus consistent with both big puzzles in the dividend literature: why dividends are paid and why they are "sticky."

Our paper now proceeds as follows: Section I introduces the model and notation. The main task of the model is to demonstrate how financial policy can determine the composition of shareholders and the relationship of this clientele composition to firm pricing. Consistent with the tax code, we assume that large, institutional shareholders receive favorable dividend tax treatment. We show that this induces them to tilt their allocation from the optimally diversified portfolio in favor of firms for which they enjoy a comparative tax advantage. The institutional presence in turn enhances the release of information about the firm, which is itself reflected in the equilibrium price. In Section I, better firms use the ability of larger shareholders to

[3] In our model, risk aversion is the counterbalancing force that prevents institutions from completely allocating themselves towards dividend-paying firms. This is realistic: the institutional sector favors dividend-paying stocks, but not at the exclusion of all other investment opportunities. See also Keim (1985).

sometimes detect firm quality to separate themselves. In this version, institutional shareholders add value by allowing better managers to signal their quality. In Section II, large tax-advantaged shareholders provide some direct services to the firm, for example, facilitation of quick and negotiated control block sales to potential raiders, by more informed voting and board communication and representation. This part is consistent with an agency view of the world. Section III summarizes the empirical implications of our model. In brief, our model offers a rich set of empirical implications, some consistent with earlier work, others new, to allow an empiricist to distinguish it from earlier work. Section IV discusses existing theories of dividends and also the two critical components of the model in light of existing evidence: that dividends attract relatively more institutional shareholdings and that the presence of institutional block shareholdings can add value to the firm. Finally, Section V summarizes.

I. A Signaling Model

In this section, we develop a model in which managers have inside information about the quality of their firm. We assume that institutions are better able to discover this information than retail investors, for example, by conducting "due diligence." Dividends attract institutions via clientele effects, which makes it more likely that the true firm quality will be revealed. Thus, dividends can induce separation between low-quality and high-quality firms.

Firm Characteristics

The economy consists of N firms, which can either be of high (H) or low (L) type. The random pretax payoffs for firms of type $j = H,L$, denoted $\tilde{V}_j$, are assumed to be normally distributed with mean μ_j and variance σ^2 with $\mu_H > \mu_L$. Consequently, the expected pretax payoffs of high-quality firms are greater than those of low-quality firms. For simplicity, we assume that the payoff variance is identical for all firms and that the random payoffs of all firms are uncorrelated.[4] Once the cash flows are realized, the firm is liquidated and the cash flows are distributed to the equity holders. All firms have no debt.

Managers

Managers are responsible for choosing the firm's dividend policy. We assume that managers maximize p, the firm's current share price.[5] If managers of low-quality firms are later revealed to have wasted firm funds by

[4] Section II.F shows that when firms have correlated payoffs, they de facto have to compete for institutional attention. In equilibrium, this leads to *higher* dividends.

[5] Managerial compensation may be aligned (increase) with share price, but is assumed to be small enough not to enter investors' wealth function in a significant manner. More generally, managers could also like higher share prices in the future. As long as there is sufficient weight on short-term performance, our intuition that dividends help managers signal and obtain a higher share price today remains.

paying dissipative dividends to try to obtain a higher share value, they would regret having had their investors pay taxes on dividends in addition to receiving only the low-quality firm value. Managers of the firm know the quality of the firm and are assumed to choose a dividend policy which maximizes expected share price. We denote the expected share price obtained by a manager of true type $t1$ who chooses dividends indicating that he is type $t2$ as $S_{t1,t2}$. We will show that, to signal the firm's quality, managers may want to pay out dividends even though capital gains have preferential tax treatment for all investors.

Dividends, Repurchases, and Investor Types

The total payout, which is the sum of dividends and repurchases, is held constant. The dividend decision consists of deciding how much of the total payout should be in the form of dividends. The remainder is paid out as a repurchase. Without loss of generality, we assume that the capital gains tax rate is zero. Hence, if the payout is in the form of share repurchases no taxes are incurred. Dividends, however, are subject to personal tax: if dividends D are paid out, then the total after-tax payoff to shareholders is reduced by $\tau_i \cdot D$, where τ_i is the marginal tax rate on dividend income for investor type i. (Corporate taxes are identical for repurchases and dividends, and are thus ignored w.l.o.g.) There are two groups of investors in our model, called "institutional" (I) and "retail" (R) investors. Each group consists of many identical investors, all of whom have constant absolute risk aversion preferences. The aggregate risk aversion for the group of institutional (retail) investors is denoted γ_I (γ_R). Type i investors have aggregate cash wealth $W_{o,i}$ and aggregate share endowment $\bar{\theta}_{ij}$ in firm $j = 1,\ldots,N$ and we normalize the total share endowment in each of the firms to unity, that is, $\bar{\theta}_{Ij} + \bar{\theta}_{Rj} = 1 \forall j$.

Our model explores two important differences between these investors:

Detection: The higher the fraction of the firm owned by institutions, the more likely it is that the firm's true type is consequently revealed/discovered. Section II replaces this assumption with an ability of large shareholders to facilitate a takeover if managerial performance later falls short.

Dividend Taxes: Institutional investors face a lower tax rate on dividends than retail investors, $\tau_I < \tau_R$.

Model Structure

After the firm's dividend policy is announced, investors trade shares at prices determined in a competitive market. The model time line is illustrated in Figure 1. Algebraic symbols used in the model are summarized in Table I.

Investor Utility Maximization

The investors' problem is to allocate their wealth among the N stocks and a riskless asset that yields a certain gross return that we normalize to unity.

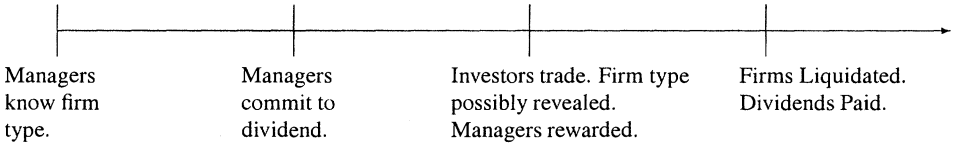

Figure 1. Signaling model time line.

Because all investors in the same group are identical, we consider the decision problem for a representative investor of each group. Wealth for representative investor i is given by

$$\tilde{W}_i = \sum_{j\in[H,L]} \theta_{ij}(\tilde{V}_j - \tau_i D_j) + \left[W_{o,i} - \sum_{j\in[H,L]} (\theta_{ij} - \bar{\theta}_{ij})p_j\right], \tag{1}$$

where θ_{ij} is the demand by an investor of type i in stock of type j, $\bar{\theta}_{ij}$ is their pretrade endowment of shares, D_j is the dividends paid by type j, and p_j is the share price for firms of type j. Investors maximize their expected utility of wealth conditional on the information signalled by the firm's dividend policy. Thus, the representative investor solves the following optimization problem:

$$\max_{\theta_{ij}} E[-e^{-\gamma_i \tilde{W}_i} | \{D_j\}]. \tag{2}$$

In a separating equilibrium, $\tilde{W}_i$ is conditionally normally distributed, in which case equation (2) is equivalent to

$$\max_{\theta_{ij}} - e^{-\gamma_i E(\tilde{W}_i) + (\gamma_i^2/2)\,\mathrm{Var}(\tilde{W}_i)}. \tag{3}$$

A. *The Separating Equilibrium*

We focus attention on a separating equilibrium in which high-quality firms pay a dissipative dividend to attract more institutional monitors and signal their higher quality.

An important part of our definition of equilibrium is how investors' beliefs about firm quality are related to the firm's dividend policy: let $\mu^c(D_j)$ denote investor beliefs about the expected cash flow of firm j as a function of the announced dividend, D_j. We give the following definition of a separating equilibrium:

Definition 1: A separating equilibrium is a collection of demand functions, prices, and dividend policies $\theta_{Ij}(D), \theta_{Rj}(D), p_j(D), D_H^*, D_L^*$ such that:

(1) Investors' inference about firm quality is correct: $\mu^c(D_H^*) = \mu_H$; $\mu^c(D_L^*) = \mu_L$.

Table I

List of Common Symbols

Variables introduced after page 2509 are used only in the agency model.

Abbreviation	Page	Description
		Exogenous Variables: Firms and Managers
N	2502	Number of firm types. (Two in the signaling model; one (for each firm) in the monitoring model.)
$\tilde{V}_j$	2502, 2510	Before-tax firm value. In signaling model, $\tilde{V}_j \sim N(\mu_j, \sigma^2)$, i.e., distributed with mean μ_j and variance σ^2. In the monitoring model, $\tilde{V}_j \sim N[\mu_j + f(M_j), \sigma^2]$, i.e., a function of the presence of institutions.
μ_j, σ_j	2502, 2510	Mean and variance of $\tilde{V}_j$.
j	2502	Firm type (quality). Unspecified in the monitoring model. $j \in \{H,L\}$, i.e., *H*igh or *L*ow, in the signaling model.
$S_{t1,t2}$	2503	Expected share price obtained by a manager of type $t1$, claiming to be a manager of type $t2$.
$\pi(\cdot)$	2507	Probability that firm type is publicly and verifiably revealed to/discovered by investors. It is a function of institutional shareholdings, i.e., $\pi \equiv \pi(\theta_I)$. In the separating equilibrium, it increases in paid dividends, i.e., $\pi \equiv \pi[\theta_I(D)]$.
c	2512	Cost Per Unit of Monitoring.
α	2512	Effectiveness Per Unit of Monitoring. Monitoring enhances firm value by $f(M) = (1/\alpha)M_j^{\alpha}$
		Exogenous Variables: Investors
$i \in I,R$	2503	Shareholder type, where I denotes institutional shareholders, R denotes retail shareholders.
$\bar{\theta}_{ij}$	2503	(Aggregate) endowments of holdings in firm j invested by investors of class i.
$W_{0,i}$	2503	Initial aggregate cash endowment by investors of type i
τ_i	2503	Tax rate on dividends faced by investors of class i. Assumed: $\tau_I < \tau_R$.
W_i	2503	Aggregate wealth of investors of class i.
γ_i	2503	Risk aversion of investors by class i.
$\bar{\tau}$	2506	Preference weighted average tax rate, $(\tau_I\gamma_R + \tau_R\gamma_I)/(\gamma_R + \gamma_I)$.
κ	2506	Inverse of the sum of risk tolerances, $[(1/\gamma_I) + (1/\gamma_R)]^{-1}$.
		Endogenous Variables
θ_{ij}	2504	(Aggregate) percent of holdings in firm j by investors of class i. (Note: θs add to 1 across all investors for each firm, but not across all firms for each investor.)
p_j	2504	Share price for firms of type j.
D_j	2503, 2510	Paid dividend by firms of type j. D^* is the notation for dividend paid by high-quality firm that satisfies the self-selection constraint in equilibrium. Low-quality firms pay $D = 0$.
M_j	2510	Amount of monitoring in Firm j.

(2) Given dividends D and prices p, investor demands $\theta_{Ij}(D)$ and $\theta_{Rj}(D)$ solve their respective optimization problem.
(3) Markets clear: $\theta_{Ij} + \theta_{Rj} = 1 \forall j$.
(4) Managers optimally reveal their quality: $S_{H,H} \geq S_{H,L}$ and $S_{L,L} \geq S_{L,H}$.

Note that in any "reasonable" separating equilibrium, low-type investors pay the lowest possible dividend, $D_L^* = 0$, to minimize their tax obligation.

B. Equilibrium in the Trading Stage

The separating equilibrium is computed using backward induction. In the trading stage, we take the dividend policies of high- and low-quality firms as given. Suppose that investors conjecture that all firms paying dividends D^* are high-quality and firms paying $D = 0$ are low-quality. Given these conjectures, investors believe (1) their after-tax payoffs for a dividend-paying firm are normally distributed with mean $\mu_H - \tau_i D^*$ (where $\tau_I < \tau_R$) and variance σ^2; and (2) their after-tax (and pretax) payoffs for low-quality firms are normally distributed with mean μ_L and variance σ^2. Given these conjectures:

PROPOSITION 1: *Equilibrium shareholdings and prices for* dividend paying *(type-H) firms are:*

Institutional Shareholdings:

$$\theta_{IH}^* = \frac{\gamma_R}{\gamma_I + \gamma_R} + \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + \gamma_R)\sigma^2}\right] D^*. \tag{4}$$

Retail Shareholdings:

$$\theta_{RH}^* = \frac{\gamma_I}{\gamma_I + \gamma_R} - \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + \gamma_R)\sigma^2}\right] D^*. \tag{5}$$

Equilibrium Price:

$$p_H = \mu_H - \kappa\sigma^2 - \bar{\tau} D^*, \tag{6}$$

where $\bar{\tau}$ is the preference-weighted average tax rate (i.e. $\bar{\tau} \equiv (\tau_I \gamma_R + \tau_R \gamma_I)/(\gamma_I + \gamma_R)$) and $\kappa \equiv [(1/\gamma_I) + (1/\gamma_R)]^{-1}$ is the inverse of the sum of the risk tolerances.

Equilibrium shareholdings and share prices for non-dividend-*paying (type-L) firms are:*

Institutional Shareholdings:

$$\theta_{IL}^* = \frac{\gamma_R}{\gamma_I + \gamma_R}, \tag{7}$$

Retail Shareholdings:

$$\theta_{RL}^{*} = \frac{\gamma_I}{\gamma_I + \gamma_R}, \tag{8}$$

Equilibrium Price:

$$p_L = \mu_L - \kappa\sigma^2. \tag{9}$$

(All proofs are in the appendix.)

Shareholdings of high-quality firms consist of the usual optimal risk-sharing term and a novel clientele term, which depends on the relative institutional tax advantage and dividend level. Because of the tradeoff between after-tax returns and risk sharing, the clientele effect is dampened when investors are more risk averse and investments are more risky. Finally, share prices are reduced by the in-equilibrium tax loss.

Because institutional and retail investors are risk averse, they typically hold all securities. As a result, both are marginal investors in all types of stocks. The retail investors value dividend-paying stocks less than institutional investors on the margin because of the taxes. As a result, dividend-paying stocks have a lower price (for a fixed μ) than non-dividend-paying stocks. This in turn makes the dividend-paying stocks relatively more attractive to the institutional investors because they offer a higher pretax return.

C. Optimal Dividend Policies

The first stage of the game—managerial dividend choice—is now analyzed, given the competitive equilibrium in the trading stage, which assumed that dividend policies of the two types are perfectly revealing. For this equilibrium to exist, managers of high-quality firms must prefer to pay $D^{*}(> 0)$, and managers of low-quality firms must prefer not to mimic high-quality types and pay zero dividends, $D = 0$.

For low-quality managers, paying a dividend will lead to a higher share price if institutions fail to detect their true quality. However, if they are detected, the share price will be lower than if they did not pay a dividend because the dividend tax is dissipative. Let $\pi(\theta_I(D))$ denote the probability that a manager of a low-quality firm is revealed to be a low-quality type for a given institutional holding, which will depend on dividend level D. We assume that $\pi(\theta_I(D))$ is differentiable with $\pi_D = \partial\pi/\partial\theta_I \cdot \partial\theta_I/\partial D > 0$ because $\partial\pi/\partial\theta_I > 0$ by assumption and $\partial\theta_I/\partial D = (\tau_R - \tau_I)/[(\gamma_I + \gamma_R)\sigma^2] > 0$ (see equation (4)).

Consider high-quality managers. If they paid $D = 0$, they would avoid the dissipative dividend, but the market would believe them to be of low-quality, unless later revealed otherwise. Their expected share price will be

$$S_{H,L} = [1 - \pi(\theta_I(0))](\mu_L - \kappa\sigma^2) + \pi(\theta_I(0))(\mu_H - \kappa\sigma^2). \tag{10}$$

If they pay dividends $D = D^*$, the expected share price will be

$$S_{H,H} = (\mu_H - \bar{\tau}D^* - \kappa\sigma^2). \tag{11}$$

Thus, a necessary condition for a separating equilibrium is

$$S_{H,H} \geq S_{H,L} \Leftrightarrow D^* \leq \frac{[1 - \pi(\theta_I(0))]\cdot(\mu_H - \mu_L)}{\bar{\tau}}, \tag{12}$$

which states that high-quality firms must not pay too large a dissipative dividend. The upper bound on the dividend is increasing in the difference in average quality, because being recognized as high-quality becomes more valuable. Furthermore, the upper bound on the dividend is decreasing in $\bar{\tau}$, because as the weighted-average tax rate increases, it becomes more costly to signal quality with dividends.

Now consider the low-type managers' problem. If they do not pay a dividend, that is, $D = 0$, the share price reflects the fact that the market believes the firm is low quality. In this case,

$$S_{L,L} = (\mu_L - \kappa\sigma^2). \tag{13}$$

If they pay dividends $D = D^*$, the firm will be revealed to be low quality with probability $\pi(\theta_I(D^*))$, so

$$S_{L,H} = [1 - \pi(\theta_I(D^*))](\mu_H - \bar{\tau}D^* - \kappa\sigma^2) + \pi(\theta_I(D^*))(\mu_L - \bar{\tau}D^* - \kappa\sigma^2). \tag{14}$$

The necessary condition for a separating equilibrium is

$$S_{L,L} \geq S_{L,H} \Leftrightarrow D^* \geq \frac{[1 - \pi(\theta_I(D^*))]\cdot(\mu_H - \mu_L)}{\bar{\tau}}. \tag{15}$$

If equation (12) holds with equality, then it is also true that equation (15) holds (with inequality). In other words, $\bar{D} = (1 - \pi(\theta_I(0)))[(\mu_H - \mu_L)/\bar{\tau}]$ supports a separating equilibrium. This choice of dividend, however, is unnecessarily costly. In the best separating equilibrium, high-quality managers want to choose the lowest (dissipative tax-minimizing) D that satisfies both equations (12) and (15). This optimal level of dividend payouts that supports a separating equilibrium satisfies equation (15) with equality, and thus depends on the (unspecified) $\pi(\theta_I(D))$ function.

PROPOSITION 2: *The D^* satisfying equation* (15) *with equality is unique and also satisfies equation* (12).

It does not require a large difference between high- and low-quality firms to produce a reasonably large dividend yield. For example, with $\mu_H = \$100, \mu_L = \$99, \sigma_L^2 = \sigma_H^2 = \$1$ (the type uncertainty is roughly equivalent to the within-type uncertainty), $\gamma_I = \gamma_R = 2, \pi(\theta_I(D)) = \theta_I(D) - 0.5$, and $\tau_I = 0\%, \tau_R = 50\%$, the solution is that high-quality firms pay \$2.67 in dividends and trade for $p_H = \$98.33$ (predividend taxes); low-quality firms pay no dividends and trade for $p_L = \$98$ today. The dividend-paying firm offers a pretax expected rate of return of 1.7 percent and a posttax rate of return of 0.3 percent, whereas the non-dividend-paying firm offers a rate of return of 1.0 percent. Consequently, investors *self-sort*: institutions hold 83.3 percent and retail investors hold 16.7 percent of the high-quality firm. The aggregate tax loss is about \$0.22 (much less than the \$1.33 implied by the statutory tax rate). Thus, a high-quality firm which is perceived by investors ex-ante to be worth either \$99 or \$100 must waste about one fifth of its value difference to signal its true quality.

PROPOSITION 3: *The optimal dividend choice for high-quality firms*

(1) increases in $\mu_H - \mu_L$*;*
(2) decreases in $\bar{\tau}$*, holding* $(\tau_R - \tau_I)$ *and* γ*s fixed;*
(3) decreases in $(\tau_R - \tau_I)$*, holding* $\bar{\tau}$ *fixed;*
(4) increases in σ^2*;*
(5) may increase or decrease in γ_I*; and*
(6) may increase or decrease in γ_R*.*

These comparative statics are not unusual within the context of a signaling model. The dividend level that is necessary to support separation is higher when a low-quality firm has more to gain by imitating (high $\mu_H - \mu_L$ and low $\bar{\tau}$), and when the relative tax penalty $(\tau_R - \tau_I)$ is smaller. If there were no difference between tax rates, no amount of dividend would induce more institutional share ownership. Higher risk aversion coefficients dampen the portfolio clientele tilt, but also influence the preference-weighted average tax rate.[6]

Yet, if the costs to separation become sufficiently high, although a literal interpretation of *any* signaling model would predict more "signal" (dividends) to be sent, it is more reasonable to presume that other real-world institutions might evolve that permitted information to be communicated in a cheaper fashion. Instead of the separating equilibrium, which can be justified with a Cho and Kreps (1987) argument, a pooling equilibrium in which no firm pays any dividends becomes more plausible.

II. The Agency Model

In the preceding section, institutions create value by allowing firms to use large shareholders' due diligence to separate themselves. But the very same large shareholders cannot only uncover information about the firms' quali-

[6] A model extension that introduces a free-riding institution also shows that this free-riding institution makes separation more difficult, and thus increases the dividend paid in equilibrium.

ties, they can also use this information to help the firm control agency problems. For example, institutional shareholders are more likely to provide large blocks to potential raiders more quickly and more reliably (Grossman and Hart (1980)), to vote in an informed manner, and generally to influence the board. Shleifer and Vishny (1986, Section 5) similarly link dividend policy to block shareholder retention.

A. The Model

We now substitute an ability by institutions to monitor for the (previously assumed) ability to differentiate between existing high- and low-quality firms. Consequently, V, μ, and p lose their quality (H/L) subscripts. Instead, we introduce M, an endogenous amount of monitoring by institutional investors. Retail investors suffer from free-riding problems, and are thus presumed unable to monitor. M units of monitoring increases firm value by M^{α}/α, but one unit of monitoring also costs c dollars. By assuming away both any persistent effect of monitoring and costs, and by assuming total payouts (i.e., dividends and repurchases) are constant at a level that maintains the firm at the same size, we can repeat the one-period model (without further loss of generality).

Firms

The economy consists of N firms, with uncertain future pretax payoff for a single firm j denoted by $\tilde{V}_j$. We assume that $\tilde{V}_j$ is normally distributed with mean μ_j and variance σ_j^2. In our basic model, we assume that the $\tilde{V}_j$ are independent across j. We will relax this assumption in Section II.F. Once the cash flows are realized, the firm is liquidated and the cash flows are distributed to the equity holders.

Dividends, Repurchases and Investor Types

The treatment of dividends, repurchases, and investors is identical to that assumed in the previous section. The sum of dividends and repurchases is constant. The capital gains tax rate is zero so repurchases do not lead to any taxes. Unlike repurchases, dividends are subject to personal taxation: if dividends D_j are paid out, then the total after-tax payoff (dividends plus capital gains) to a shareholder is reduced by $\tau_i \cdot D_j$, where τ_i is the applicable marginal tax rate on dividend income received from firm j for investor i.

The two groups of investors in our model remain institutional (I) and retail (R) investors. Each group consists of identical investors, all of whom have constant absolute risk aversion preferences and care only about their own after-tax returns. The aggregate risk aversion for the group of institutional (retail) investors is denoted γ_I (γ_R). Type i investors have aggregate cash wealth $W_{0,i}$ and aggregate share endowment $\bar{\theta}_{ij}$ in firms $j = 1, \ldots, N$ and we normalize the total share endowment in each of the firms to unity, that is, $\bar{\theta}_{Ij} + \bar{\theta}_{Rj} = 1$ for all j.

The differences in shareholders are:

Institutional Shareholder Monitoring: Only institutions are large enough to provide services. This assumption is plausible because retail investors suffer from more severe free-rider problems: the retail clientele consists of many investors, each of whom has only a small stake in the firm and thus little incentive to expend personal resources to provide monitoring services that benefit all shareholders. In contrast, institutions (typically less than a handful of owners with more than five percent holdings in most Fortune 500 firms) have more of a stake in seeing the company perform well. They can sell large share blocks to potential raiders more quickly and more cheaply (Grossman and Hart (1980)), influence the board, become well informed and vote against management, or simply make "suggestions" for improvements to management. These suggestions can lead to changes in operating policy and managerial effort that increases the value of the firm. (Section IV offers more detailed justifications for the monitoring and limited free-riding assumption.)

Dividend Taxes: Institutional investors face a lower tax rate on dividends than retail investors, $\tau_I < \tau_R$.

Model Structure

Our model consists of four stages. In the first stage, managers choose a dividend policy to maximize share price in the trading stage.[7] In the second stage, trade between institutional and retail investors occurs. The institutional (retail) shareholdings (and market-clearing prices) depend on the firm's dividend policy, because of the differential tax treatment of dividends to the two types of investors and the inferred monitoring activity. In the third stage, institutional investors choose an optimal amount of monitoring, M, given their shareholdings. In the final stage, the firm is liquidated and shareholders are paid off. The model time line is illustrated in Figure 2. Algebraic symbols used in the model are summarized in Table I.

B. Monitoring Stage

We begin by considering the institutions' monitoring decision. Institutions are assumed to be able to coordinate their actions[8] and choose the amount of monitoring in stage three that yields maximal expected profits for the institutions in aggregate, given their shareholdings after the trad-

[7] Monitoring enhances firm value. Even if sometimes undesirable ex-post, monitoring is sought after by managers ex-ante. Furthermore, the model could be extended to allow managers to receive a weighted average of the current and future share price. The model would become far more complex (then having to deal with managerial uncertainty, rather than the certain share price), but all of the insights of the current model would remain intact.

[8] In the main model, we assume that institutions do not face a coordination problem in their monitoring activity. In Section II.G, we show that the introduction of free-riding institutions reduces the desire of other institutions to monitor, but that firms still choose dividends that are positive.

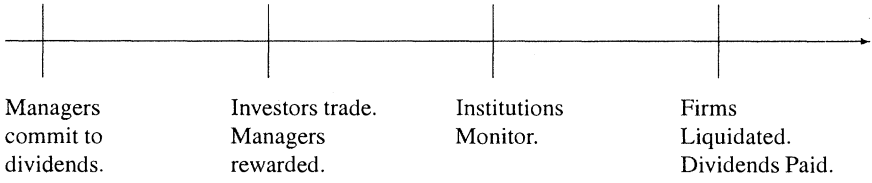

Figure 2. Monitoring model time line.

ing round in stage two. Let θ_{Ij} denote the aggregate shareholdings in firm j by the institutions in stage two. Monitoring is costly, at c dollars per unit of M, but M units of monitoring increases firm value by the sure amount $f(M) = (1/\alpha)M_j^{\alpha}$, where $0 < \alpha < \frac{1}{2}$.[9] In sum, the predividend tax, postmonitoring payoffs for firm j are distributed as $N[\mu_j + f(M_j), \sigma_j^2]$. Institutions hold θ_{Ij} share of the firm. Thus institutions solve

$$\max_{M_j} \theta_{Ij} \cdot \frac{1}{\alpha} M_j^{\alpha} - c \cdot M_j \quad \Rightarrow \quad M_j^* = \left(\frac{\theta_{Ij}}{c}\right)^{1/(1-\alpha)}. \tag{16}$$

Notice that the optimal level of monitoring increases in the institutional shareholdings, θ_{Ij}; increases in the efficiency of monitoring, α; and decreases in the cost of monitoring, c. This is the mechanism by which dividends can enhance firm value: they attract institutions that then voluntarily provide value-enhancing services.

C. *Trading Stage*

In stage two, institutions and retail investors must allocate their wealth among the N stocks plus a riskless asset that yields a certain gross return that we normalize to unity. Because all investors in the same group are identical, we consider the decision problem for a representative investor of each group. We assume that both institutions and retail investors behave competitively in the trading round. This assumption is made for tractability reasons, and is relaxed in Section II.E.[10]

[9] The assumption $0 < \alpha < \frac{1}{2}$ ensures that the benefit of optimal monitoring, $f(\cdot)$, is concave not only in M, but also in institutional shareholdings (θ_I).

[10] Assuming that institutions behave competitively in the trading decision and coordinated at the monitoring stage can be justified. There are many tax-exempt institutions that could compete to potentially become a block investor for a given firm, but in any given firm, the 100 percent of shares available can only end up as block holdings for less than, say, a dozen institutions. Typically, there are about a handful of institutions with block holdings in any given firm. Furthermore, Section II.E shows that dividends may actually be *larger* if we assume that institutions understand their own subsequent impact on future monitoring in the trading stage.

All investors understand that institutions will choose M_j^* units of monitoring in stage three. Thus, they correctly predict the effects that aggregate monitoring will have on firm value. Institutions solve the following problem in stage two:

$$\max_{\theta_{Ij}} E[-e^{-\gamma_I \tilde{W}_I}], \tag{17}$$

where $\tilde{W}_I = \sum_{j=1}^{N}[\theta_{Ij}(\tilde{V}_j + (1/\alpha)(M_j^*)^{\alpha} - \tau_I D_j)] + [W_{0,I} - \sum_{j=1}^{N} p_j(\theta_{Ij} - \bar{\theta}_{Ij})]$, and p_j and D_j are the price and dividend of firm j. As before, institutional investors solve

$$\max_{\theta_{Ij}} \gamma_I \cdot E(\tilde{W}_I) - \left(\frac{\gamma_I^2}{2}\right)\mathrm{Var}(\tilde{W}_I). \tag{18}$$

By the same reasoning, retail investors solve

$$\max_{\theta_{Rj}} \gamma_R \cdot E(\tilde{W}_R) - \left(\frac{\gamma_R^2}{2}\right)\mathrm{Var}(\tilde{W}_R), \tag{19}$$

where $\tilde{W}_R = \sum_{j=1}^{N}[\theta_{Rj}(\tilde{V}_j + (1/\alpha)(M_j^*)^{\alpha} - \tau_R D_j)] + [W_{0,R} - \sum_{j=1}^{N} p_j(\theta_{Rj} - \bar{\theta}_{Rj})]$. Finally, we close the model in the trading stage with the market-clearing conditions $\theta_{Ij} + \theta_{Rj} = 1$ for all firms j.

PROPOSITION 4: *In the trading stage, the equilibrium is*

$$\theta_{Ij} = \frac{\gamma_R}{\gamma_I + \gamma_R} + \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + \gamma_R)\sigma_j^2}\right] D_j, \tag{20}$$

$$\theta_{Rj} = \frac{\gamma_I}{\gamma_I + \gamma_R} - \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + \gamma_R)\sigma_j^2}\right] D_j, \tag{21}$$

$$p_j = \mu_j + \frac{1}{\alpha}\left(\frac{\theta_{Ij}}{c}\right)^{\alpha/(1-\alpha)} - \kappa\sigma^2 - \bar{\tau} D_j, \tag{22}$$

where $\kappa = \left((1/\gamma_I) + (1/\gamma_R)\right)^{-1}$ *is the inverse of the sum of the risk tolerances and* $\bar{\tau} = (\tau_I \gamma_R + \tau_R \gamma_I)/(\gamma_I + \gamma_R)$ *is the preference-weighted average tax rate.*

As before, institutional shareholdings consist of an optimal risk-sharing component, equal to $\gamma_R/(\gamma_I + \gamma_R)$, and a clientele component that is induced by different marginal tax rates. When tax-rate differentials are zero, investors' equilibrium holdings cannot be changed by paying dividends. Assuming institutions have a lower marginal tax rate than individuals, the institu-

tional shareholdings in the high-dividend-paying firms increase *linearly* in the tax differential and in the paid dividend. Moreover, the dividend tax deadweight loss reduces the in-equilibrium share price.

D. Dividend Choice

In stage one, we assume that the firm's managers choose the dividend policy, D_j, that maximizes the share price, p_j, in stage two. Managers are fully aware of the implications of their dividend choice on institutional shareholdings, monitoring, and share prices. An increase in the dividend has two countervailing effects. Because $\tau_R > \tau_I$, an increase in the dividend increases institutional shareholdings and the level of value-enhancing monitoring of the firm. On the other hand, the dividend tax reduces firm value. The optimal dividend equates the marginal benefit of monitoring with the marginal tax cost.

PROPOSITION 5: *Under mild parameter restrictions,*[11] *the optimal dividend in stage one is*

$$D_j^* = \left[\frac{1}{(1-\alpha)\bar{\tau}}\right]^{(1-\alpha)/(1-2\alpha)} \left[\frac{(\tau_R - \tau_I)}{c\sigma_j^2(\gamma_I + \gamma_R)}\right]^{\alpha/(1-2\alpha)} - \frac{\gamma_R \sigma_j^2}{(\tau_R - \tau_I)}. \tag{23}$$

PROPOSITION 6: *The optimal dividend is greater when (1) the cost of monitoring c is smaller; (2) the difference in tax treatments* $(\tau_R - \tau_I)$ *is greater, holding* $\bar{\tau}$ *fixed; (3) the preference-weighted average tax rate* $\bar{\tau}$ *is smaller, holding* $(\tau_R - \tau_I)$ *fixed; (4) monitoring efficiency* α *is greater; and (5) the variance of the stock payoff* σ_j^2 *is smaller.*

To a firm, the marginal cost of dividends increases with the effective dissipative tax rate $\bar{\tau}$. The marginal benefit of dividends depends on two factors: the effectiveness of the dividends to increase institutional shareholdings, and the effectiveness of institutional shareholdings on value-enhancing monitoring. Firms find it is easier to attract more institutions when σ^2 is low and $\tau_R - \tau_I$ is high. Given a level of institutional shareholdings, value-enhancing monitoring increases when monitoring is cheap (low c) and effective (high α).

E. Extension: Market Power

In the basic model we assumed that there were many small institutions, each of which traded as if they had negligible impact on the stock price via their own trading and their own monitoring (although they did take the

[11] To guarantee that the first term is larger than the second term, that is, that the dividend is positive, one must rule out the region in which c, σ^2, and the risk aversion coefficients γ_I and γ_R are too high. Similar constraints would apply to guarantee institutional shareholding between zero and one, and a positive price. To violate these constraints (in which case D^* would be zero), one has to assume extreme and unreasonable parameter values.

aggregate monitoring into account). In this subsection, we assume one large institution that understands both the subsequent amount of monitoring it will do[12] and its effect on the price of shares through its own trading. To get a closed-form solution we must assume that $f(M) = \ln(M)$ and $\bar{\theta}_{Ij} = 0$. The latter assumption is made for tractability and states that the institutional investor does not start out with a toehold (Admati, Pfleiderer, and Zechner (1994)).

PROPOSITION 7: *In the model in which institutions anticipate their in-equilibrium monitoring and the impact of their trades on equilibrium prices, the trading stage results in*

$$\theta_{Ij} = \frac{\gamma_R}{\gamma_I + 2\gamma_R} + \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + 2\gamma_R)\sigma_j^2}\right] D_j - \frac{1}{(\gamma_I + 2\gamma_R)\sigma_j^2}, \tag{24}$$

$$\theta_{Rj} = \frac{\gamma_I + \gamma_R}{\gamma_I + 2\gamma_R} - \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + 2\gamma_R)\sigma_j^2}\right] D_j + \frac{1}{(\gamma_I + 2\gamma_R)\sigma_j^2}, \tag{25}$$

$$p_j = \mu_j + \ln\left(\frac{\theta_{Ij}}{c}\right) - \left(\frac{\tau_R\gamma_R + \tau_I\gamma_R + \tau_R\gamma_I}{2\gamma_R + \gamma_I}\right) D - \frac{(\gamma_R + \gamma_I)\gamma_R\sigma^2 + \gamma_R}{(2\gamma_R + \gamma_I)}. \tag{26}$$

Having some market power, the institution "shaves" its optimal holdings to incorporate its own price impact and subsequent monitoring costs, thus reducing the equilibrium price in the trading stage.

PROPOSITION 8: *If* $f(M) = \ln(M)$, *the optimal dividend in stage one when the institutional shareholders have limited market power is given by:*

$$D_j^* = \frac{\gamma_I + 2\gamma_R}{\tau_R\gamma_I + \tau_I\gamma_R + \tau_R\gamma_R} - \frac{\gamma_R\sigma_j^2}{(\tau_R - \tau_I)} + \frac{1}{(\tau_R - \tau_I)}. \tag{27}$$

Compared to the base case, the institution has relatively lower holdings in the dividend-paying firm (and more so if the dividend is larger). With lower institutional holdings, the marginal monitoring benefit from raising dividends is larger. However, it is more difficult to attract the institutional clientele, which reduces the marginal benefit of raising the dividends. Thus, the optimal dividend can either be greater or smaller than in the base model.

F. Extension: A Second Firm with Correlated Value

In this subsection, we assume that there are two firms with *correlated* values. Firm 1 has random liquidation value of $\tilde{V}_1$, which is normally distributed with mean μ_1 and variance σ_1^2. Firm 2 has random liquidation

[12] Technically, this changes the previously fixed M_j^* term in the institutional wealth objective in the trading stage into a term that depends on θ_I (according to equation (16)) and subtracts the cost of monitoring.

value of $\tilde{V}_2$, which is normally distributed with mean μ_2 and variance σ_2^2. The covariance (correlation) between the liquidation values of Firm 1 and Firm 2 is σ_{12} (ρ).

Consider the equilibrium in our model for Firm 1, taking the dividend policy of Firm 2 as given. The monitoring stage of the model is identical to the basic model.

PROPOSITION 9: *The equilibrium for the shares of firm 1 in the trading stage of the two-firm model is*

$$\theta_{I1} = \frac{\gamma_R}{\gamma_I + \gamma_R} + \left[\frac{(\tau_R - \tau_I)}{(\gamma_R + \gamma_I)\sigma_1^2(1-\rho^2)}\right]D_1 - \left[\frac{\sigma_{12}}{\sigma_1^2}\frac{(\tau_R - \tau_I)}{(\gamma_R + \gamma_I)\sigma_2^2(1-\rho^2)}\right]D_2, \tag{28}$$

$$\theta_{R1} = \frac{\gamma_I}{\gamma_I + \gamma_R} - \left[\frac{(\tau_R - \tau_I)}{(\gamma_R + \gamma_I)\sigma_1^2(1-\rho^2)}\right]D_1 + \left[\frac{\sigma_{12}}{\sigma_1^2}\frac{(\tau_R - \tau_I)}{(\gamma_R + \gamma_I)\sigma_2^2(1-\rho^2)}\right]D_2, \tag{29}$$

$$p_1 = \mu_1 + \frac{1}{\alpha}\left(\frac{\theta_{I1}}{c}\right)^{\alpha/(1-\alpha)} - \kappa(\sigma_1^2 + \sigma_{12}) - \bar{\tau}D_1. \tag{30}$$

Notice that the equilibrium price in the second stage is identical to the one-firm model except the term multiplying κ is $(\sigma_1^2 + \sigma_{12})$ instead of σ_1^2.

PROPOSITION 10: *The optimal dividend in stage one of the two-firm model with correlated values is*

$$D_1^* = \left[\frac{1}{(1-\alpha)\bar{\tau}}\right]^{(1-\alpha)/(1-2\alpha)}\left[\frac{(\tau_R - \tau_I)}{c\sigma_1^2(\gamma_I + \gamma_R)(1-\rho^2)}\right]^{\alpha/(1-2\alpha)} - \frac{\gamma_R\sigma_1^2(1-\rho^2)}{(\tau_R - \tau_I)} + \left(\frac{\sigma_{12}}{\sigma_2^2}\right)D_2 \tag{31}$$

PROPOSITION 11: *If $\sigma_{12} > 0$, the optimal dividend D_1^* is increasing in (1) σ_{12} and (2) D_2.*

As σ_{12} increases, the two firms become closer substitutes from the perspective of the institution. Thus, each firm in effect competes for institutional attention. The equilibrium outcome is a *higher* level of dividends than it is in the base model.

G. Extension: A Second, Free-riding Institution

In this subsection, we assume that there are two institutions: the second institution has aggregate risk aversion parameter of γ_F. Both institutions pay the same tax rate on dividends, but we assume that one type of institution does no monitoring and free rides on the monitoring of the other institution.[13] The optimal monitoring problem for the remaining monitor in the third stage is identical to the one-firm case.

PROPOSITION 12: *The equilibrium shareholdings for the monitoring institution i*

$$\theta_{Ij} = \frac{\gamma_F \gamma_R}{\gamma_F \gamma_R + \gamma_I \gamma_R + \gamma_I \gamma_F} + \left[\frac{\gamma_F(\tau_R - \tau_I)}{(\gamma_F \gamma_R + \gamma_I \gamma_R + \gamma_I \gamma_F)\sigma_j^2} \right] D_j, \tag{32}$$

and the equilibrium price is

$$p_j = \mu_j + \frac{1}{\alpha}\left(\frac{\theta_{Ij}}{c}\right)^{\alpha/(1-\alpha)} - \hat{\kappa}\sigma^2 - \hat{\tau}D_j, \tag{33}$$

where $\hat{\kappa} = \left((1/\gamma_I) + (1/\gamma_F) + (1/\gamma_R)\right)^{-1}$ *is the inverse of the sum of the risk tolerances and* $\hat{\tau} = \left((\gamma_F \gamma_R \tau_I + \gamma_I \gamma_R \tau_I + \gamma_I \gamma_F \tau_R)/(\gamma_F \gamma_R + \gamma_I \gamma_R + \gamma_I \gamma_F)\right)$ *is the preference-weighted average tax rate.*

PROPOSITION 13: *The optimal dividend in the model with two institutions is*

$$D_j^* = \left[\frac{1}{(1-\alpha)\hat{\tau}}\right]^{1-\alpha/1-2\alpha} \left[\frac{\gamma_F(\tau_R - \tau_I)}{c\sigma_j^2(\gamma_I \gamma_R + \gamma_I \gamma_F + \gamma_R \gamma_F)}\right]^{\alpha/(1-2\alpha)} - \frac{\gamma_R \sigma_j^2}{(\tau_R - \tau_I)}. \tag{34}$$

Compared to the base model (23), free-riding may increase or decrease the optimal dividends. The intuition is the same as for proposition 8. (This is also the case if we hold the aggregate risk tolerance constant.) The intuition that firms pay dividends to induce monitoring remains. For example, presume that $\sigma^2 = \$2$, $\alpha = 0.25$, $c = 0.02$, $\tau_R = 0.5$, $\tau_I = \tau_F = 0$, and thus $\bar{\tau} = 0.0833$. Even if the free-riding institution(s) are relatively less risk averse than the monitor, for example, with $\gamma_I = \gamma_R = 2$ and $\gamma_F = 0.5$, the monitoring institution(s) still hold 27.2 percent of the firm in equilibrium and provide

[13] A taxed institution that does not help the firm (e.g., a trust) is equivalent to a retail investor in our model. Thus, this section covers the presence of all four combinations of monitoring and tax status.

0.444 units of monitoring at a cost of \$0.89. The firm receives \$3.26 in benefits from this monitoring, and optimally pays \$5.06 in dividends to attract this monitoring. In equilibrium, the deadweight dividend tax loss is about \$0.42.[14]

H. A Multiperiod Extension: Smoothing

We can also extend the model to a multi-period setting to explain the well-known fact that firms tend to choose a smooth level of dividends over time. Consider an infinite horizon version of the model in which the firm chooses its dividend policy each period and institutional monitoring increases firm value only in that period. The total payout each period is such that the size of the firm remains constant. In this case, the one-period model above is representative of every period in the infinite horizon model. If the parameters $\{\gamma_I, \gamma_R, \tau_I, \tau_R, \sigma^2, c\}$ remain constant each period, managers would choose the same dividends each period. The model used to derive this smoothing result is special in a number of ways. However, the reason that the level of dividends is independent of the realization of the firm's payoff is that dividends are chosen ex-ante to attract institutional investors to motivate managers. The smoothing result is thus likely to hold in a much wider range of circumstances than in the constant parameter scenario.

In a more realistic dynamic setting, there is another effect that reinforces smoothing. If a firm pays dividends in the first period, it will have attracted an institutional clientele. In contrast, a firm that does not pay dividends but instead repurchases shares will have attracted a retail clientele. If a firm that paid dividends in period one were to decide to cut its dividends in period two (e.g., as out-of-equilibrium behavior), its shareholder base would be precisely the clientele that can punish managers for poor behavior and correct the problem. In contrast, if a firm that repurchased shares in period one were to decide not to repurchase in period two, its clientele would be too diffuse to create difficulties for management. This clientele smoothing pressure that arises from the multiperiod dynamics reinforces the single-period clientele smoothing choice that arises from the assumed constant parameters. Consistent with the nonimmediate clientele rearrangement found in empirical studies, DeAngelo and DeAngelo (1990) find that reluctance to reduce dividends in response to a difficult situation increases, *ceteris paribus*, when dividends have been long standing.

III. Empirical Implications

Our theory is designed to explain an empirical puzzle, and its constructs are intrinsically close to empirical proxies. The theory links tax differentials, activism, and services by large, tax-exempt shareholders to company perfor-

[14] The desired dividend yield and relative firm value enhancement can be calibrated by choosing a suitable base firm value μ.

mance and company dividends. Still, as with all theories, empirical tests are confronted with a more complex reality, in which *ceteris paribus* is difficult to achieve. Although the theory itself was sketched in *levels*, more promising tests of our theory are likely to come from tests in *differences*. For example, our theory argues that tax-exempt firms' dividends induce activist institutions to tilt towards them, but omitted considerations—such as unmeasurable institutional beliefs (e.g., about stock return performance by characteristics), information differences, and preferences—that may create institutional shareholdings benchmark levels that are different for each firm. Consequently, firm-specific, omitted variables are better controlled for by relating the effect of *changes* in dividends on *changes* in shareholdings, which nets out many firm-specific and investor-specific omitted considerations. Further, although dividend yields may be the most reasonable measure of payout, an empiricist must be aware that time-series variation in stock prices (the denominator) probably overwhelms time-series variation in the manager's choice variable, dividends (the numerator).

Our first two implications derive from the premise of our theory:

1. Firms with more severe agency or inside information problems ex-ante are more likely to pay dividends in order to control them. Firms experiencing increases (decreases) in agency or information problems are more likely to increase (decrease) their dividends.
2. When other shareholders (e.g., KKR or Berkshire Hathaway) appear who can mitigate the agency/information concerns but who do not value dividends, we would expect target companies to substitute reinvestment, share repurchases, or other means of payout for dividend payments.

Our next implications exploit cross-sectional heterogeneity in dividends, clientele, and performance.[15]

3. Firms initiating/increasing their dividends attract new institutional clientele. Firms omitting/decreasing their dividends lose institutional clientele. Such tests would be particularly powerful if shareholders with preference for dividends (e.g., tax exempts) and with activist tendencies could be identified. (For more detail, see Section III.B.)
4. Firms experiencing increases in large, activist shareholders experience increases in future performance. (For more detail, see Section III.C.)
5. Firms increasing dividends have unexpected future performance increases.

[15] In levels, implications 3 to 5 would be stated as: firms paying dividends have more institutional clientele; firms with more active shareholders perform relatively better in the future; firms paying dividends perform relatively better in the future. As stated above, *ceteris paribus* is difficult to accomplish in level tests.

Our theory is also consistent with dividend smoothing and can relate smoothing to institutional shareholdings:

6. Firms are reluctant to cut dividends (or to reduce the growth rate of dividends). Dividends are "stickier" or "smoother" than share repurchases. Firms with more large, activist, institutional, tax-exempt shareholders are *especially* reluctant to reduce dividends, and especially eager to increase dividends.[16] Dividend smoothing implies that empirical tests which relate dividends to past performance are likely to find a positive relationship; and tests which relate dividend levels to institutional holding levels and future performance levels are likely to be weak. Tests relying on the effects of *unexpected changes* in dividends are likely to be more powerful.

In addition to firm-specific tests, an empiricist can also use the theory to predict aggregate time-series and international patterns.

7. Dividends change with the tax differential between institutions and individuals over time. (In John and Williams (1985), it is not the tax *differential* that matters, but the taxes per se.)
8. Aggregate dividends paid (relative to aggregate repurchases) relate to the size of the institutional (tax-exempt) sector. The effect of increases in the institutional sector on dividends depends on the other parameters of the model.
9. Our theory can be tested not only in time series, but also in international cross section. Adjusting for other control mechanisms available in each country, we would expect to see use of dividends in countries in which there are low-tax investors who are more inclined to act as monitors and in which dividends are relatively tax disadvantaged. Dewenter and Warther's (1998) abstract states that "Japanese firms, particularly keiretsu-member firms, face less information asymmetries and/or agency conflicts than U.S. firms, and . . . information asymmetries and/or agency conflicts affect dividend policy. Japanese firms experience smaller stock price reactions to dividend omissions and initiations, they are less reluctant to omit and cut dividends, and their dividends are more responsive to earnings changes" (i.e., there is less smoothing).

Finally, our theory has a natural relation to debt instruments:

10. If there are institutions (e.g., banks) with an ability to monitor, and which tend to (or are forced only to) hold firms' debt instruments, debt payments—which are also more difficult to shelter by retail investors—can serve the same role as dividend payments in our model.

[16] On one hand, firms with a large institutional activist shareholder are more expected to increase dividends, which reduces the unanticipated shareholder response. On the other hand, such dividend increases are likely to be especially valuable to investors if they can retain and attract the appropriate investor clientele.

> Still, dividends and debt payments may not be exactly alike. An institution that holds debt may be able to detect lousy management, but it would not be able to displace it easily, unless the firm goes bankrupt. This may provide an additional set of implications: under some assumptions about the effects of monitoring in different states, firms paying dividends should also pay interest. Thus, our theory could be used to explain the association of the presence of banks with better performance and firms paying dividends. Better firms may want to attract both types of institutions, because they offer a signal/monitoring for different states of the world.

Of course, in reality, our theory coexists with other rationales for dividends, institutional ownership, and performance. For example, the tax code provides for cases under which dividends are treated as nontaxable return of funds. For firms whose dividends are not fully taxable, one would expect less of an institutional comparative preference. Utilities may have had regulatory requirements for high dividend payouts, independent of any advantages to attracting institutions. Similarly, the managerial and owner incentives of individuals may be different in some firms. Ownership may be so concentrated (e.g., Microsoft's Bill Gates) that share repurchase payouts are as difficult to shelter from taxes (by tax-loss selling) as are dividend payouts. In other cases, the presence of executive stock options may push managers of some firms towards share repurchases rather than towards dividends.[17] And a more powerful test could rely on the holdings of the single largest, external, activist shareholder, rather than simply aggregate institutional shareholdings. Finally, an empiricist should consider comparing dividend-paying companies not necessarily to non-dividend-paying companies, but to equivalent share repurchasing companies.

The aforementioned long-run time-series implications relating aggregate dividends to aggregate sector institutional holdings and performance are also very susceptible to confounding effects. For example, before the emergence of institutions, a naïve view would prescribe no reason for dividends. However, wealthy individuals may have played the role of institutions. Prior to the *Tax Reform Act of 1986*, the tax code had numerous loopholes by which wealthy individuals (unlike their poorer cousins) could shelter income. These individuals would have also had more concentrated holdings and taken a more active role in the oversight and governance of the firm. In essence, wealthy individuals would have assumed the role of institutions hypothesized in our paper. Similarly, one can take a broader view of institutions: other corporations as shareholders also receive tax-favored dividend treatment and may similarly enhance the value of the firm by oversight or other synergistic linkages.

[17] Executive stock options may thus be a direct substitute of institutional monitoring/dividends. Executive stock options give management incentives to reduce dividends and thus its institutional oversight clientele.

IV. Existing Evidence

Our theory is not the first to offer an explanation of dividends. Section IV.A reviews the most related theories and points out how our theory differs from these. For a comprehensive review of the dividend literature see Allen and Michaely (1995). Further, our theory is not the first concerned with our two main assumptions: (1) that investors are taxed differently and invest rationally, so that dividends can induce specific clientele changes; and (2) that the presence of an institutional clientele can increase the value of the firm. Both of these assumptions have been explored in many other papers, and Sections IV.B and IV.C review the related evidence.

A. Existing Theories of Dividends

The closest paper to our own is Shleifer and Vishny (1986, Section 5), which also recognizes that dividends can be a mechanism to compensate institutional investors. However, dividends can function effectively only in certain circumstances in their model. Small shareholders pay the minimal (or zero) dividends to keep a large institutional shareholder just indifferent between keeping and divesting an existing toehold. An institutional shareholder has a fixed probability of finding a value improvement. Prices are exogenous, and no dividend is paid if the expected return on the investment is high. In our model, corporations maximize share price by paying dividends to actively tilt ex-ante investments in equilibrium. Our model also considers the problem of *limited* in-equilibrium self-sorting to derive equilibrium holdings given only a *relatively* advantageous dividend tax treatment (Shleifer and Vishny (1986) rely on four tax rates), and permits value uncertainty and diversification. While Shleifer and Vishny (1986) focus on the role of toeholds in possibly permitting takeovers, we focus on the in-equilibrium clientele tilting argument. (Our model also works, e.g., with a signaling argument.) Finally, we link the institutional investor holding to dividend smoothing.

John and Williams (1985) present a theory where dividends are a costly signal precisely because they are taxed more heavily than share repurchases. Shareholders sell shares to satisfy liquidity needs. If the firm is undervalued their holdings will be diluted. By using taxed dividends as a costly signal they can prevent this dilution. Although John and Williams' model has many attractive features, it does not provide a fully satisfactory resolution of the dividend puzzle. It is not obvious that its empirical implications are consistent with the stability and smoothing of dividends that has been documented by Lintner (1956), and many subsequent authors. The best way to extend the John and Williams model over a longer time is not entirely clear. If firms' prospects do not change over time, then once a firm has signaled its type, no further dividend payments will be necessary and payouts can be made through share repurchases. If firms' prospects are constantly changing, which seems more plausible, and if dividends signal these, we would expect the dividends to constantly change, also. By contrast, in the agency version of our model the firm requires agency services every period.

Bernheim (1991) also provides a theory of dividends where signaling occurs because dividends are taxed more heavily than repurchases. In his model the firm controls the amount of taxes paid by varying the proportion of the total payout that is in the form of dividends as opposed to repurchases. A good firm can choose the optimal amount of taxes to provide the signal. As with the John and Williams model, this model does not provide a good explanation of dividend smoothing.

Kumar (1988) provides a theory of dividend smoothing. In his model the managers who make the investment decision know the true productivity type of the firm but the outside investors do not. Also the managers want to invest less than the outside investors because they are less diversified. They will try to achieve lower investment by underreporting the firm's productivity type. Kumar shows that a fully revealing equilibrium where dividends perfectly signal productivity cannot exist. If it did, the shareholders' could deduce the firm's true productivity type, but this is inconsistent with managers underreporting. A coarse signalling equilibrium can exist, though. Within an interval of productivity it is shown that it is optimal for the different types to cluster at a corresponding dividend level. This theory is consistent with smoothing, because small changes in productivity will not usually move a firm outside the interval and so its dividend will not change. Unfortunately, it does not explain why share repurchases, which are taxed less, are not used instead of dividends.

Building on work by Ofer and Thakor (1987) and Barclay and Smith (1988), Brennan and Thakor (1990) present another theory about why repurchases have a disadvantage compared to dividends. When some shareholders are better informed about the prospects of the firm than others, they will be able to take advantage of this information when there is a repurchase. They will bid for stock when it is worth more than the tender price but will not bid when it is worth less. Similarly, uninformed buyers will be worse off. When money is paid out in the form of dividends, the informed and the uninformed receive a pro rata amount; they will do equally well. As a result, uninformed shareholders prefer dividends to repurchases; this preference will persist even if dividends are taxed more heavily than repurchases, provided the tax disadvantage is not too large. On the other hand, the informed will prefer repurchases because this allows them to profit at the expense of the uninformed. Brennan and Thakor argue that the method of disbursement chosen by firms will be determined by a majority vote of the shareholders. If the uninformed have more votes than the informed, dividends will be used, but if the informed predominate, repurchases will be chosen. The Brennan and Thakor model is an intriguing explanation of the preference that firms appear to have for dividends. It answers the question of why firms prefer to use dividends even though they are taxed more heavily. Unlike John and Williams' theory, it is consistent with dividends being smoothed. However, the range of tax rates for which dividends are preferred to repurchases because of adverse selection is usually small.

Chowdhry and Nanda (1994) and Lucas and McDonald (1998) also consider models where there is a tax disadvantage to dividends and an adverse selection cost to repurchases. In their models, managers are better informed than shareholders and it is shown how payout policy depends on whether managers think the firm is overvalued or undervalued relative to the current market valuation. Both models provide interesting insights into the advantages and disadvantages of dividends and repurchases. However, the stability and smoothing of dividends is difficult to explain in this framework unless firms remain undervalued or overvalued relative to their market value through time.

Two other models that are related to ours are Hausch and Seward (1993) and Zwiebel (1996). Hausch and Seward (1993) model the distinction between dividends and repurchases as being one between deterministic and stochastic payment methods since the repurchase price is uncertain but the level of dividends is not. They show firms can signal the level of their internally generated cash by an appropriate choice between the two methods. This theory does not explain smoothing. Zwiebel (1996) develops a model of dynamic capital structure where managers choose debt to credibly constrain their misuse of free cash flow. His analysis also has implications for payout policy because dividends constrain managers similarly to debt. He suggests dividends are superior to repurchases because they can be paid regularly whereas repurchases must be irregular to be taxed at a lower rate than dividends. However, repurchases could be done continuously, so it is not clear this is the case.

B. Taxes, Dividends and External (Block) Shareholdings

Tax-Induced Clientele Effects

Miller and Modigliani (1961) lay out their famous conditions under which dividend payout is irrelevant. In their framework, different shareholders do not add value differentially to the underlying firm. As in our own paper, Miller and Modigliani (1961) invoke the tax clientele effect, arguing that there are no problems in having institutions assume *all* dividend-paying shares. Unlike our paper, their intent is to invoke it to argue that dividend taxes can de facto be *costlessly* avoided. Taking this clientele-tax avoidance argument one step further, Miller and Scholes (1978) argue that tax-exempt insurance policies can tax-shelter *all* income (and specifically dividend income) which is not consumed. In effect, they argue that the "marginal" shareholder is a tax-exempt entity. Studying ex-dividend price reactions, Elton and Gruber (1970) (and others) find empirically that clientele effects are present and serve to reduce (but not eliminate) the aggregate dividend tax burden.

Our own model is consistent with these views, in that clientele effects can reduce the tax burden. Unlike Miller and Scholes (1978), however, we employ a counterveiling force—a desire to diversify (and costs to providing ser-

vices to the firm)—that prevents the tax-exempt sector from assuming *all* dividend-paying stocks and thus eliminating all dividend taxes economy-wide. Aside, our model would work just as well with a smaller counterveiling force (which can be achieved by assuming a low institutional risk-aversion coefficient γ_I), in which case institutions could perfectly sort themselves to dividend-paying stocks (and then perform monitoring services). Yet, allowing for imperfect separation seems to be more realistic. As stated in the introduction, dividend taxes raised $82 billion in 1994 alone. It seems that dividend taxes are not being fully avoided.

Finally, a debt-related argument in Miller (1977) could be transfered to dividends. Aggregate taxes are important, but investor clientele effects render each individual firms' choice irrelevant. When one firm increases its debt, another firm can decrease it to rebalance the optimal aggregate amount of dividend taxes paid.[18] Our own model considers different firms to benefit differently from different shareholders. It thus identifies which firms should pay and which firms should not. Further, our paper produces an algebraic and solvable model that clarifies a number of the issues raised in Miller (1977).

Non-Tax-Induced Clientele Effects

Although our work assumes that institutional investment choices are unconstrained, one could justify an institutional preference for dividend-paying stocks directly by appealing to common institutional charter restrictions (which force institutions to invest only in stocks that pay dividends), or to the *prudent man rule*. In this context, the prudent man rule forces many institutional managers to invest overwhelmingly in stocks with high dividend yields. Of course, it is reasonable to argue that such investment rules survive because they make sense: tax-exempt institutions *should* sort themselves towards stocks with higher dividend yields because of their higher pretax expected returns. Moreover, to the extent that other large, potentially taxable investors would sort themselves preferentially towards dividend-paying stocks, our hypothesis would apply even in the absence of any tax advantages.[19]

Empirical Evidence on Clientele Changes

Michaely, Thaler, and Womack (1995) and others have examined volume changes around dividend changes as indicators of clientele rearrangements. Such tests offer little power, given the high variance of volume. Dhaliwal,

[18] Our model could equally well be used to argue that the presence of debt induces clientele effects, which in turn provides firms with appropriate monitors. But there is no puzzle about debt to explain. Debt is tax-deductible for firms, and therefore possibly tax efficient. The monitoring benefits of debt are just "gravy."

[19] Instead of relating dividends to tax differences, an empiricist could relate dividend payouts and firm performance to the prevalence of charter restrictions, or simply dividend payouts to the observed presence of institutions in firms.

Erickson, and Trezevant (1998) examine changes in institutional shareholdings around dividend initiation dates directly. In a sample of 133 initiators of dividends from the 1982 to 1995 period, they find that 80 percent of their firms experience increases in institutional shareholdings over the three to nine months following the initiation. The average firm's institutional shareholdings increase from about 25 percent to about 30 percent within one year (about five to ten more institutional holders of record). The shareholder increases are highly statistically significant, and significant in competition with factors like time trends, beta, prior stock and accounting performance, and size—all of which tend to be insignificant. Dhaliwal et al. (1998) also find that clientele effects weakened after *The Tax Reform Act of 1986*, which decreased the relative tax rate on dividends for individuals. In sum, Dhaliwal et al. (1998) find good evidence in favor of dividend-/tax-clientele effects: in fact, their *only* reliable predictor of institutional ownership changes seems to be the initiation of dividends. Brav and Heaton (1998) find that many institutional investors stopped holding dividend-omitting firms after the 1974 ERISA subjected private pension fund managers to the prudent man rule, and dividend-omitting firms underperformed. After the reinitiation of dividends, both effects reversed.[20]

There is also ample anecdotal evidence of the relation between dividends and institutional involvement. For example, in the *Bank of America* Roundtable (*Journal of Applied Corporate Finance* 10(2), Summer 1997, p. 57), the CFO of Bank of America, Mike O'Neill, articulates that "We've got a lot of institutional investors, and a number of them continue to have dividend requirements that we just try to meet. Many of our institutional investors will not invest in a company that doesn't have at least a 2 percent dividend yield . . . We think there is a value to having a broad investor base . . ."

Conclusion

In our model, by tilting their portfolios in favor of dividend-paying stocks, tax-exempt institutions gain higher rates of return but choose to incur a loss of some diversification benefits and some costs of monitoring. Thus, in equilibrium, clientele effects reduce but do not eliminate dividend taxes paid. By varying the sector risk aversion parameters, the model can be consistent with a scenario in which institutions sort themselves perfectly to dividend-paying stocks or with a scenario in which dividends induce only mild clientele effects.

[20] Strickland (1997) offers more ambiguous evidence. Although tax-exempts do seem to sort themselves towards dividend-paying stocks, the effect becomes statistically insignificant after controlling for other firm characteristics. Still, Strickland (1997) has only four years of data and his D/P ratios may measure more variations in price than variations in smoothed dividends. In contrast, Dhaliwal et al. (1998) and Brav and Heaton (1998) focus on changes in dividends, not dividend yields.

C. External (Block) Shareholdings, Corporate Control and Performance

Our model assumes that (institutional) block-owner presence increases the actual or inferred value of the firm. Attributing a special role to large (institutional) investors is either an assumption made or conclusion drawn in so many previous finance and nonfinance papers that we can only give the reader a flavor of the existing literature.[21]

In their aforementioned well-known study, Shleifer and Vishny (1986) point out that there are several mechanisms through which blockholders can add value. These monitoring activities can be almost zero cost. For example, institutions could vote more actively, or formally or informally publicize problems in firms, or simply publicize their reasons for divesting of large share blocks. Perhaps more importantly, they could sell share blocks to a potential raider. Hostile takeover attempts are more likely to occur if there are large, unaffiliated blockholders (Shivdasani (1993)).[22] Indeed, it would be difficult, costly, or perhaps outright impossible for a potential raider to quickly accumulate a meaningful block from dispersed retail investors' shares (Bagwell (1991, 1992)).[23] Consequently, raiders are more likely to appear and succeed if there are large shareholders who are either willing to sell their shares or willing to vote with them against incumbent management.

There are also more costly options available to large activist shareholders. For example, they could open a dialogue with management and the board; they could designate a representative or third party as external candidates for board seats; they could sell shares in bulk at the first sign of trouble; or they could take sides or become a player in drawn out and sometimes public battles to remove management.

Free-riding Considerations

An important question is the extent to which block shareholders can coordinate on the more expensive monitoring/detection activities. Section IV.G shows that free-riding reduces dividends in the monitoring model and increases dividends in the signaling model, but the basic intuition that dividends can "purchase" institutional ownership remains. Admati et al. (1994) and Maug (1998) argue that large shareholders tend to monitor more than

[21] For example, the reader may also want to consult volume 20 of the *Journal of Financial Economics*, which focuses on issues of power among corporate managers, shareholders, and directors. Similarly, the legal literature, too, argues that long-term institutional investors have particularly strong influence on firms (e.g., Ayres and Cramton (1994)).

[22] Stulz (1988) argues that management may seek to attract a "friendly" clientele, that will vote shares in favor of management in control contests. Depending on managerial shareholdings and the trade-off between incentive effects and power, this can increase or decrease total firm value. Pound (1988) documents that such reallocation of shares can decrease the value of the firm. (The increased institutional shareholdings to the announcement of dividends [and the positive share price reaction] could be considered the flip side of this relationship.)

[23] For example, after a management buyout offering is announced, the total outside percentage of shares held in blocks decreases (Peck (1996)).

small shareholders, even when smaller shareholders successfully free ride. Still, free riding remains a concern, especially for those monitoring activities that are expensive and for those firms in which institutional holdings are spread over many different investors.

The largest institutional shareholder is likely to play a significant role. Shleifer and Vishny (1986) find that in about half of Fortune 500 firms, *the largest shareholder is a pension fund, profit-sharing plan, or institution* (the other half are firm and family holdings). The single largest shareholder typically holds about 15 percent of the votes.[24] The widely acknowledged leader in shareholder activism (and the only one to publish its voting records) is CalPERS, the California Pension Employee Retirement System, which manages about US-$125 billion in assets as of early 1998. The College Retirement Equities Fund (TIAA-CREF) is similarly active, but so are many smaller pension funds.[25] Because pension funds tend to be disproportionately long-term investors and because public pension funds, in particular, tend to be unaligned with firms' management, they are also more likely to be able to play the role attributed to them in our paper.[26] Smith (1996) examines 51 firms targeted by CalPERS from 1987 to 1993. CalPERS typically held large stakes in the aforementioned 51 firms, which had experienced poor stock price performance. Seventy-two percent of these firms eventually adopted proposed CalPERS changes. Del Guercio and Hawkins (1999) study shareholder activism and find that funds are more successful at monitoring and promoting change in target firms than previously recognized. Further, our model predicts a trade-off between diversification and clientele tilt. Del Guercio and Hawkins (1999) find that CalPERS and CREF seem to be 80 percent indexed, but devote the remaining 20 percent to large-stake investments in just a few firms in which they also take a very active interest.

To alleviate free-riding problems, institutions may also have evolved coordination mechanisms and coalitions (see, e.g., Zwiebel (1995)). For example, there are at least three companies (Institutional Shareholder Services (ISS), Investor Responsibility Research Center, and Proxy Monitor) which either make recommendations that are by-and-large followed by their client institutions or even vote their clients' shares directly. (Most institutions have a fiduciary responsibility to vote.) ISS advises about 350 institutions (plus

[24] Institutions only have to file with the SEC if they own at least 5 percent of the outstanding shares. On average, institutions tend to hold about 25–30 percent in aggregate, which is certainly enough to be a concern for poorly performing management.

[25] The *Wall Street Journal* (1998a) describes CREF's ongoing battle to appoint a largely independent board for Disney, and mentions CREF's ongoing attempts at other companies. CalPers is by no means alone. The *Wall Street Journal* (1998b) reports that even "small" pension funds can concentrate on a small set of firms and pressure them into changes. Lens, a Portland, Maine, pension fund pressured directors of Waste Management and Steon & Webster to replace several chief executives and board members, and pressured Sears, Roebuck to divest its brokerage, real-estate, and insurance operations.

[26] Pension funds also need not exert pressure on firms directly—about 60 percent of their assets are allocated to external managers, who in turn are replaced if they perform poorly. And in many cases, pension funds reserve the right to vote shares purchased by investment advisors.

many trustees, custodians, and corporations) for about 10,000 shareholder meetings each year. CalPERS also was a force in the creation of the Council of Institutional Investors, representing 80 institutional investors with $600 billion in assets as of 1993, designed to encourage shareholder activism. Carleton, Nelson, and Weisbach (1998) document that TIAA-CREF (another large pension fund) targets companies and gets these targeted companies to implement changes 95 percent of the time. More informal coordination devices could involve a tit-for-tat leadership by one institution in one case and by another institution in another case.

Empirical Studies of Aggregate Institutional Behavior

There are some empirical studies that investigate the role of institutional and block share ownership in the determination of subsequent firm policy and performance.[27]

Pound (1988) describes that in about half of dissident proxy contests, outsiders win. Brickley, Lease and Smith (1988) find a positive relation between institutional holdings (especially foundations, public-employee pension funds, and mutual funds) and the percentage of votes cast against management-sponsored antitakeover amendments. He also argues that other firms' holdings (unlike other institutional investors) tend to be aligned with management, and are thus not effective monitors. Denis, Denis, and Sarin (1997) find that the median management owned 6.3 percent of shares, and institutions held 32.9 percent. There is a statistically significant higher unusual turnover rate if a firm has at least one 5 percent owner—with an implied turnover probability influence for the "institutional presence" variable which is of about the same influence as the prior year's stock market performance! Further, the presence of an outside block holder increases the sensitivity of executive turnover to firm performance. (Naturally, this is confounded by the ownership stake of the management team, and potentially by the alignment of external blocks with management.) Denis and Serrano (1996) find that management turnover following unsuccessful control contests is concentrated among poorly performing firms in which outside blockholders acquire an ownership stake. They conclude that monitoring by active outside investors facilitates valuable internal control efforts. The frequency of turnovers is startling: only 21.6 percent of 37 poor precontest performers are removed in firms without an unaffiliated block, whereas 65 percent of 23 such managers are removed in firms with an unaffiliated block. Shivdasani (1993) finds that the probability of hostile takeovers in a sample rises from 19 percent to 36 percent if unaffiliated blockholders own 23 percent rather than 0 percent of the firm's equity. Bethel, Liebeskind, and Opler's (1998) study is the only one to explicitly focus on differences among external blockholders. They develop a classification into activists (e.g., Carl Icahn, Irwin Jacobs),

[27] Unfortunately, empirical institutional ownership measures are only proxies for the identification of tax-exempt and activist investors (see Scholes, Wilson, and Wolfson (1992)).

financial blockholders (which includes pension funds) and (friendly) strategic blocks. They find that both active and financial blocks are associated with subsequent operating improvements, even when there is no subsequent M&A activity. They conclude that "the market for partial corporate control can play an important role in reducing the agency costs that result from the separation of ownership and control in U.S. corporations." (p. 631)

In response to activist shareholders, many corporations have developed significant shareholder relations groups. Many of them also contract with such firms as The Carson Group and Georgeson, which provide consulting services that identify and communicate with institutional shareholders and attempt to forecast the institutional and general response to shareholder proposals. Georgeson also reports that among 99 sponsors of corporate governance proposals in 1997, only three were sponsored by other, possibly taxed, entitities.[28]

Conclusion

There is a long tradition of attribution of a special role to institutional shareholders in both empirical and theoretical papers. Despite a relevant concern that free riding reduces the effectiveness of monitoring, especially for the subset of expensive monitoring activities, there is ample empirical evidence that block holders and institutions can play a role in the facilitation of control of corporations. To the extent that free-rider issues are important, firms may need to attract a larger number of share blocks than derived in our paper, that is, by paying even more dividends. Naturally, if these costs become prohibitive, firms may shift from institutional oversight to other (unidentified) control mechanisms altogether.

Finally, it is worthwhile to point out that although academics sometimes attempt to classify shareholder activities into either an ability to detect (signaling) and an ability to monitor or dislodge management (agency), shareholder activism often accomplishes both and may be empirically difficult or impossible to decompose.

D. Dividends and Performance

There is a large event-study literature that shows that stock prices respond to both dividend and share repurchase announcements in a positive fashion. Michaely et al. (1995) find that dividend initiators experience positive return drift and dividend omitters experience negative return drift. Part of this long-run (pretax) price drift could reflect changes in expected returns that compensate investors for the implicit taxes. Fama and French (1998) similarly conclude in their abstract that "dividends . . . convey infor-

[28] The most successful proposals in 1997 concerned the rescission of poison pills and the repeal of classified boards.

mation about profitability (expected net cash flows) . . . [which] obscures any tax effects of financing decisions." Further, "a $1 change in annual dividends is associated with about $10 change in value."

Benartzi, Michaely, and Thaler (1997) find that dividend increases are not associated with higher *future* earnings, but with higher *lagged* earnings. Although no rational theory is likely to be fully consistent with a positive dividend-announcement effect without intrinsic information that the firm is better than previously thought (or that it has better future earnings prospects)—which is really an asset-pricing puzzle—the smoothing intrinsic to our theory makes it much more robust to timing than some earlier dividend models. It is conceivable that Benartzi et al. (1997) have difficulties detecting forward-looking performance increases, because firms smooth their dividends/dividend growth in a way that makes their empirical tests relatively powerless. Institutional shareholders in most firms may have continued to "force" management to increase dividends, even in the absence of additional favorable future news.[29]

In contrast to Benartzi et al. (1997), Brook, Charlton, and Hendershott (1998) find that firms do indeed raise their dividends in anticipation of *unusual and large* increases in future earnings, but that they are reluctant to adjust them thereafter (especially downward), perhaps accounting for the Benartzi et al. (1997) nonfinding. The Brook et al. (1998) view is broadly consistent with our own: dividends carry value-based information, but they are also sticky.

V. Conclusion

This paper offers a novel explanation for the puzzle of why some firms prefer to pay dividends rather than repurchase shares. We assumed that institutional investors are more likely to invest in dividend-paying stocks. This could derive, for example, from a tax advantage relative to individual investors, which induces "dividend clientele" effects. We argue that these clientele effects are the very reason for the presence of dividends, because institutions have a relative advantage in monitoring firms or in detecting firm quality. Firms paying dividends attract relatively more institutions and perform better. The theory is consistent with some documented regularities, such as a reluctance of firms to cut dividends, and offers novel empirical implications, such as a prediction that it is the tax difference between institutions and retail investors that determines dividend payments, not the absolute tax payments, as in John and Williams (1985). The theory also offers some implications on the effects of tax code changes, on international evi-

[29] There are also other issues of concern. For example, dividend decreases amount to about 3.3 percent of market cap, and lead to subsequent earnings increases of 0 to 2.0 percent of market cap (usually from a smaller base) two years later. This may not necessarily mean that dividend cuts should be good news. If omitted dividends were reinvested into a three-year interest-bearing annuity (rather than into a perpetuity, as they consider), one should not be surprised to find extra earnings of about 1 to 2 percent per year for two years.

dence, and on the association between dividend payments and coupon payments. (Debt payments could similarly attract other monitoring institutions to a firm.)

Our paper succeeds in offering an explanation for a phenomenon (the presence of dividends, positive announcement reactions, and dividend smoothing) for which there are few other explanations. Yet it cannot explain why firms/managers have not found cheaper ways to signal their inside information or to enact better controls on management than by attracting institutions. Still, even if a number of firms have found alternative ways to accomplish effective oversight, and even if many firms have other reasons for paying dividends (e.g., as suggested by the theories mentioned in the Introduction), the mechanism discussed in this paper remains an intrinsically plausible force to add to the attraction of dividends for better firms/managers, despite their overall tax disadvantages: when a firm pays higher dividends, it attracts disproportionately larger ownership by institutions, and these institutions in turn are more likely to play a larger role in overseeing management than dispersed retail investors. Managers would weigh the positive share price response to the announcement of dividends against the consequences of angering institutional shareholders if they were later forced to cut the dividends in response to poorer performance.

Appendix A. Proofs

Proof of Proposition 1: Wealth for representative investor i is now

$$\widetilde{W}_i = \sum_{j\in[H,L]} \theta_{ij}(\widetilde{V}_j - \tau_i D_j) + \left[W_{0,i} - \sum_{j\in[H,L]} (\theta_{ij} - \bar{\theta}_{ij})p_j\right], \tag{A1}$$

where θ_{ij} is the demand by an investor of type i in stock of type j, D_j is the dividends paid by and p_j is the share price for firms of type j.

In a (symmetric) separating equilibrium, investors correctly conjecture that firms paying dividends $D_j = D^* > 0$ are high quality (i.e., $E[\widetilde{V}_j] = \mu_H$), whereas firms that do not pay dividends are low quality (i.e., $E[\widetilde{V}_j] = \mu_L$). Using this we solve equation (18) which yields

$$\theta_{ij} = \frac{\mu_H - \tau_i D^* - p_j}{\gamma_i \sigma^2} \tag{A2}$$

for the dividend-paying firms and

$$\theta_{ij} = \frac{\mu_L - p_j}{\gamma_i \sigma^2} \tag{A3}$$

for the non-dividend-paying firms. To determine the equilibrium prices, we use the market-clearing condition $\theta_{Ij} + \theta_{Rj} = 1 \forall j$. Substituting these prices into the demand functions (A2) and (A3) yields the desired result. Q.E.D.

Proof of Proposition 2: Let $F(D) \equiv \bar{\tau}D - [1 - \pi(\theta_I(D))](\mu_H - \mu_L)$. D^* satisfies equation (15) with equality iff $F(D^*) = 0$. $F(0) = -[1 - \pi(\theta_I(0))] \times (\mu_H - \mu_L) < 0$ and $F(\bar{D}) = [\pi(\theta_I(\bar{D})) - \pi(\theta_I(0))](\mu_H - \mu_L) > 0$, where $\bar{D} = (1 - \pi(\theta_I(0)))[(\mu_H - \mu_L)/\bar{\tau}]$. Furthermore, $F'(D) = \bar{\tau} + \pi_D(\theta_I(D)) \times (\mu_H - \mu_L) > 0$. Because $\pi(\theta_I(D))$ is assumed differentiable (and continuous), there exists a unique D^* satisfying equation (15) with equality. Finally, because $D^* \leq \bar{D}$, it follows that D^* also satisfies equation (12). Q.E.D.

Proof of Proposition 3: At the optimal choice of D^* we have $F(D^*) \equiv \bar{\tau}D^* - [1 - \pi(\theta_I(D^*))](\mu_H - \mu_L) \equiv 0$. First, note that $F_{D^*} = \bar{\tau} + \pi_D(\theta_I(D^*)) \times (\mu_H - \mu_L) > 0$. Implicit differentiation yields all of our comparative statics results:

(1) $$\frac{\partial D^*}{\partial(\mu_H - \mu_L)} = \frac{(1 - \pi(\theta_I(D^*)))}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} > 0;$$

(2) Holding $(\tau_R - \tau_I)$ fixed, θ_{IH} is fixed, and then

$$\frac{\partial D^*}{\partial \bar{\tau}} = \frac{-D^*}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} < 0;$$

(3) Holding $\bar{\tau}$ fixed,

$$\frac{\partial D^*}{\partial(\tau_R - \tau_I)} = \frac{-\dfrac{\partial \pi}{\partial \theta_I}\dfrac{\partial \theta_I}{\partial(\tau_R - \tau_I)}(\mu_H - \mu_L)}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} < 0;$$

(4) $$\frac{\partial D^*}{\partial \sigma^2} = \frac{-\dfrac{\partial \pi}{\partial \theta_I}\dfrac{\partial \theta_I}{\partial \sigma^2}(\mu_H - \mu_L)}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} > 0;$$

(5) $$\frac{\partial D^*}{\partial \gamma_I} = \frac{-\left[\dfrac{\partial \bar{\tau}}{\partial \gamma_I} D^* + \dfrac{\partial \pi}{\partial \theta_I}\dfrac{\partial \theta_I}{\partial \gamma_I}(\mu_H - \mu_L)\right]}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} \lesseqgtr 0; \text{ and}$$

(6) $$\frac{\partial D^*}{\partial \gamma_R} = \frac{-\left[\dfrac{\partial \bar{\tau}}{\partial \gamma_R} D^* + \dfrac{\partial \pi}{\partial \theta_R}\dfrac{\partial \theta_R}{\partial \gamma_R}(\mu_H - \mu_L)\right]}{\bar{\tau} + \pi_D(\theta_I(D^*))(\mu_H - \mu_L)} \lesseqgtr 0. \quad \text{Q.E.D.}$$

Proof of Proposition 4: Because we have assumed that the payoffs for firm j are independent of the payoffs for all other firms, we can examine the equilibrium for shares in firm j as if they were the only shares the investors could trade in. The first-order condition for the institutional shareholdings in firm j is

$$\mu_j + \frac{1}{\alpha}(M_j^*)^\alpha - \tau_I D_j - p_j - \gamma_I \theta_{Ij} \sigma_j^2 = 0. \tag{A4}$$

Similarly, the first-order condition for the retail investors for shares in firm j is

$$\mu_j + \frac{1}{\alpha}(M_j^*)^\alpha - \tau_R D_j - p_j - \gamma_R \theta_{Rj} \sigma_j^2 = 0. \tag{A5}$$

With the market-clearing condition we have three linear equations to solve the three equilibrium quantities $(\theta_{Ij}, \theta_{Rj}, p_j)$ and we get the desired result. Q.E.D.

Proof of Proposition 5: Differentiate the equilibrium p_j in stage two with respect to D_j and set equal to zero. Q.E.D.

Proof of Proposition 6: All of the results can be derived by partially differentiating D_j^* with respect to the relevant variables and using the fact that $0 < \alpha < 1/2$. Q.E.D.

Proof of Proposition 11: In the first term of the optimal dividend, an increase in the covariance increases ρ (if the covariance is positive) and, since the exponent $\alpha/(1 - 2\alpha)$ is positive, this term increases. The second term is premultiplied by a negative sign so an increase in ρ increases D_1^* and, finally, an increase in the covariance increases the third term if $D_2 > 0$. The second part of the proposition is obvious. Q.E.D.

Proofs of other propositions are in Appendix B, which is available from the author upon request and on the *Journal of Finance* world-wide web site at http://www.afajof.org/.

REFERENCES

Admati, Anat R., Paul Pfleiderer, and Josef Zechner, 1994, Large shareholder activism, risk sharing, and financial market equilibrium, *Journal of Political Economy* 102, 1097–1130.

Allen, Franklin, and Roni Michaely, 1995, Dividend policy, in Robert A. Jarrow, Vojislav Maksimovic, William T. Ziemba, eds.: *North-Holland Handbooks of Operations Research and Management Science: Finance* (Elsevier, New York).

Allen, Franklin, and Anthony M. Santomero, 1998, The theory of financial intermediation, *Journal of Banking and Finance* 21, 1461–1485.

Ayres, Ian, and Peter Cramton, 1994, Relational investing and agency theory, *Cardozo Law Review* 15, 1033–1066.

Bagwell, Laurie Simon, 1991, Shareholder heterogeneity: Evidence and implications, *American Economic Review* 81, 218–221.

Bagwell, Laurie Simon, 1992, Dutch auction repurchases: An analysis of shareholder heterogeneity, *Journal of Finance* 47, 71–106.

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

Barclay, Michael J., and Clifford W. Smith, 1988, Corporate payout policy: Cash dividends versus open-market repurchases, *Journal of Financial Economics* 22, 61–82.

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

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

Bethel, Jennifer E., Julia Porter Liebeskind, and Tim Opler, 1998, Block share purchase and corporate performance, *Journal of Finance* 53, 605–634.

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

Brennan, Michael J., and Anjan V. Thakor, 1990, Shareholder preferences and dividend policy, *Journal of Finance* 45, 993–10.

Brickley, James A., Ronald C. Lease, and Clifford W. Smith, 1988, Ownership structure and voting on antitakeover amendments, *Journal of Financial Economics* 20, 267–292.

Brook, Yaron, William T. Charlton, Jr, and Robert J. Hendershott, 1998, Do firms use dividends to signal large future cash flow increases? *Financial Management* 27, 46–57.

Carleton, Willard T., James M. Nelson, and Michael S. Weisbach, 1998, The influence of institutions on corporate governance through private negotiations: Evidence from TIAA-CREF, *Journal of Finance* 53, 1335–1362.

Cho, In-Koo, and David Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 102, 179–122.

Chowdhry, Bhagwan, and Vikram Nanda, 1994, Repurchase premia as a reason for dividends: A dynamic model of corporate payout policies, *Review of Financial Studies* 7, 321–350.

DeAngelo, Harry, and Linda DeAngelo, 1990, Dividend policy and financial distress: An empirical investigation of troubled NYSE firms, *Journal of Finance* 45, 1415–1431.

Del Guercio, Diane, and Jennifer Hawkins, 1999, The motivation and impact of pension fund activism, *Journal of Financial Economics* 52, 293–340.

Denis, David J., Diane K. Denis, and Atulya Sarin, 1997, Ownership structure and top executive turnover, *Journal of Financial Economics* 45, 193–221.

Denis, David J., and Jan M. Serrano, 1996, Active investors and management turnover following unsuccessful control contests, *Journal of Financial Economics* 40, 239–266.

Dewenter, Kathryn L., and Vincent A. Warther, 1998, Dividends, asymmetric information and agency conflicts: Evidence from a comparison of the dividend policies of Japanese and U.S. firms, *Journal of Finance* 53, 879–904.

Dhaliwal, Dan S., Merle Erickson, and Robert Trezevant, 1998, A test of the theory of tax clienteles for dividend policies, Working paper, University of Arizona, University of Chicago, and University of Southern California.

Elton, Edwin J., and Martin J. Gruber, 1970, Marginal stockholder tax rates and the clientele effect, *Review of Economics and Statistics* 52, 68–74.

Fama, Eugene F., and Kenneth R. French, 1998, Taxes, financing decisions, and firm value, *Journal of Finance* 53, 819–843.

Grossman, Sanford J., and Oliver D. Hart, 1980, Takeover bids, the free-rider problem, and the theory of the corporation, *Bell Journal of Economics* 11, 42–64.

Hausch, Donald B., and James K. Seward, 1993, Signaling with dividends and share repurchases: A choice between deterministic and stochastic cash disbursements, *Review of Financial Studies* 6, 121–154.

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

Keim, Donald, 1985, Dividend yields and stock returns: Implications of abnormal January returns, *Journal of Financial Economics* 14, 473–489.

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

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

Lucas, Deborah J., and Robert L. McDonald, 1998, Shareholder heterogeneity, adverse selection, and payout policy, *Journal of Financial and Quantitative Studies* 33, 233–253.

Maug, Ernst, 1998, Large shareholders as monitors: Is there a tradeoff between liquidity and control? *Journal of Finance* 53, 65–98.

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

Ofer, Aharon R., and Anjan V. Thakor, 1987, A theory of stock price responses to alternative corporate cash disbursement methods: Stock repurchases and dividends, *Journal of Finance* 42, 365–394.

Peck, Sarah W., 1996, The influence of professional investors on the failure of management buyout attempts, *Journal of Financial Economics* 40, 267–294.

Pound, John, 1988, Proxy contests and the efficiency of shareholder oversight, *Journal of Financial Economics* 20, 237–265.

Scholes, Myron S., G. Peter Wilson, and Mark A. Wolfson, 1992, Firms' responses to anticipated reduction in tax rates: The tax reform act of 1986, *Journal of Accounting Research* 30, 161–185.

Shivdasani, Anil, 1993, Board composition, ownership structure, and hostile takeovers, *Journal of Accounting and Economics* 16, 167–198.

Shleifer, Andrei, and Robert Vishny, 1986, Large shareholders and corporate control, *Journal of Political Economy* 94, 461–488.

Smith, Michael P., 1996, Shareholder activism by institutional investors: Evidence from CALPERS, *Journal of Finance* 51, 227–252.

Strickland, Deon, 1997, Determinants of institutional ownership: Implications for dividend clienteles, Working paper, Ohio State University.

Stulz, René M., 1988, Managerial control of voting rights, *Journal of Financial Economics* 20, 25–54.

Zwiebel, Jeffrey, 1995, Block investment and partial benefits of corporate control, *Review of Economic Studies* 62, 161–185.

Zwiebel, Jeffrey, 1996, Dynamic capital structure under managerial entrenchment, *American Economic Review* 86, 1197–1215.