

A Theory of Dividends Based on Tax Clienteles*

Franklin Allen[†]

Antonio Bernardo[‡]

Ivo Welch[§]

May 19, 1998

Abstract

This paper offers a novel explanation for why some firms prefer to pay dividends rather than repurchase shares. It is well-known that institutional investors are relatively less taxed than individual investors, and that this 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.

*This paper will be available from <http://linux.anderson.ucla.edu/research.html>. We thank Shlomo Benartzi, Michael Grinblatt, Avraham Kamara, and Avanidhar Subramanyam for helpful comments. This paper is UCLA-Anderson working paper #13-98 and Wharton working paper . It is abstracted at <http://www.sfs.org/>.

[†]Contact: <mailto:allenf@wharton.upenn.edu>. US mail address: The Wharton School, 2300 Steinberg Hall — Dietrich Hall, Philadelphia, PA 19104-6367, (215) 898-3629.

[‡]Contact: <mailto:antonio.bernardo@anderson.ucla.edu>. US mail address: Anderson Graduate School of Management at UCLA, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, Tel: (310) 825-2508.

[§]Contact: <mailto:ivo.welch@anderson.ucla.edu>. US mail address: Anderson Graduate School of Management at UCLA, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, Tel: (310) 825-2508.

One of the thorniest puzzles in corporate finance is why firms pay dividends. 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 U.S. and many other countries, share repurchases are apparently superior to dividends. Nevertheless, dividends have traditionally been a substantial proportion of earnings while repurchases have not. For example, for the 1973-1983 period, dividends for the largest 1,000 firms in the U.S. averaged 44% of earnings while repurchases averaged only 6% (see Allen and Michaely (1995)). Although repurchases went up significantly in 1984 and have remained high, they were not a substitute for dividends. From 1984-1988, repurchases averaged 38% of earnings while dividends averaged 51%.

A number of theories have been put forward in the literature to explain the predominance of dividends relative to repurchases. John and Williams (1985) present a theory where dividends are a costly signal precisely because they are taxed more heavily than share repurchases. Their starting point is the assumption that the shareholders in a firm have liquidity needs which they must satisfy by selling some of their shares. The firm's managers, who act in the interests of the original shareholders, know the true value of the firm. Outside investors do not. If the firm is undervalued when the shareholders' liquidity needs arise, they would be selling shares at below the true value. However, suppose the firm pays a dividend which is taxed. If outside investors take taxed dividends to be a good signal, then the share price will rise; the shareholders will have to sell less equity to meet their liquidity needs and will maintain a higher proportionate share in the firm.

Why do bad firms not find it worthwhile to imitate good ones in the John and Williams model? When the dividends are paid, it is costly to shareholders because they must pay taxes on them. But there are two benefits. First, a higher price is received for the shares that are sold. Second, and more importantly, a higher proportionate share in the firm is retained. If the firm is actually undervalued, this higher proportionate share is valuable to the shareholder. When the managers' information is bad and the firm is overvalued, the reverse is true. It is this difference that allows separation. Only firms that are actually good will benefit enough from

the higher proportionate share to make it worthwhile bearing the cost of the taxes on the dividends.

Although John and Williams' model has many attractive features, it does not provide a fully satisfactory resolution of the dividend puzzle. In terms of assumptions, they take it as a given that shareholders must meet their liquidity needs by selling their shares. The use of debt by shareholders is ruled out. If this were allowed, it would seem that shareholders' liquidity needs could be met without incurring the cost of signaling. A more important criticism of the John and Williams model is that it is not obvious that its empirical implications are consistent with the stability and smoothing of dividends which has been documented by Lintner (1956), Fama and Babiak (1968) 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.

Building on work by Ofer and Thakor (1987) and Barclay and Smith (1988), Brennan and Thakor (1990) have presented another theory 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. Uninformed buyers will receive only a portion of their order when the stock is undervalued, but will receive the entire amount when it is overvalued. This adverse selection means that they are at a disadvantage in a share repurchase. When money is paid out in the form of dividends, the informed and the uninformed receive a pro rata amount, so there is no adverse selection. 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. When there is a fixed cost of obtaining information, the number of informed will depend on the distribution of shareholdings and the amount paid out. For a given payout, investors with large holdings will have an incentive to become informed. When a small amount is paid out, only the investors with the largest holdings will become informed, and most shareholders will remain uninformed and will prefer dividends. When a larger amount is paid out, more shareholders become informed, so repurchases may 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 the 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. In order to explain the predominance of dividends, another argument that relies on shareholders being homogeneous must be used. For tax rates above the level where adverse selection can explain the preference for dividends, everybody will tender in a repurchase, so it will be pro rata. The U.S. tax code specifies that if repurchases are pro rata, they will be treated the same as dividends, so firms might as well pay dividends. It is critical to this argument that shareholders are the same, so that they all tender. In practice, Bagwell (1991) has provided evidence that there is considerable shareholder heterogeneity, so this part of the explanation for dividends is not very convincing. Another criticism is that if adverse selection were a serious problem, firms could gather the relevant information and publicly announce it.

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 over- or undervalued relative to the current market valuation. Both models provide interesting insights into the advantages and disadvantages of dividends and repur-

chases. 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 papers 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. 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 in order to be taxed at a lower rate than dividends. Again although both papers provide insights into payout policy neither provides a rationale for dividend smoothing.

Our own paper presents a novel explanation of why some firms prefer paying dividends to repurchasing shares which is both consistent with smoothing and which does not depend on investor homogeneity.

We assume there are clienteles of investors who are taxed differently. In the United States, for example, public and corporate pension funds, colleges and universities, and foundations are exempt from taxes. For simplicity, we assume there are just two clienteles: untaxed institutions and taxed individuals. Because of their scale, the (untaxed) institutions have greater incentives to become informed. For a given holding of shares, they have a higher probability of identifying whether a firm is good or bad. Good firms will attempt to take advantage of this feature of institutions by pursuing policies which make it relatively more attractive for them to be held by institutions. (Section 3 discusses alternative mechanisms, such as monitoring, bonding and principal-agent considerations, by which the presence of institutional shareholdings can be associated with higher firm value. These differ-

ent mechanisms are also consistent with our model, but to simplify the exposition, we first discuss the pure signaling mechanism.)

One way of attracting institutions is to pay dividends.¹ Because these are taxed for individuals but untaxed for institutions, the firm will become relatively attractive to institutions. This will result in a higher fraction of ownership by institutions. (And, indeed, many institutions are by charter obliged to purchase *only* dividend-paying stocks.) Consequently, dividends signal firms are good because if they were bad, there would be a higher chance of being detected by the institutions. Bad firms will not find it worth imitating and will pay no dividends to maximize their after-tax payoff. The model thus provides an explanation of why dividends are preferred to repurchases for some (higher quality) firms. It is also consistent with dividend smoothing—reductions in dividends would be disproportionately penalized by the relatively more organized institutions—and investor heterogeneity. Managers would not need to understand the details of the mechanisms proposed in our paper: they would simply weigh the positive share price response to the announcement of dividends against the potential consequences (problems created by institutional shareholders if they were later forced to have to cut the dividends in response to poorer performance). More confident and better managers would be more inclined to pay higher dividends.

Our hypothesis offers a set of implications, some consistent with earlier work, others new which will allow an empiricist to distinguish it from earlier work:

1. Firms with more severe inside information or principal agent problems ex-ante are more likely to use dividends to control them.
2. Firms paying dividends perform relatively better in the future. Firms that paid dividends in the past (relative to repurchases) also tend to perform well.²

¹For example, in the *Bank of America* Rountable (*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..."

²The concurrent association relies on the smoothing implication of our model, which in turn relies on shareholder activism rather than signaling as a reason for the higher firm value when in-

3. Firms paying dividends attract more institutions.³
4. Aggregate dividends paid (relative to aggregate repurchases) relate to the size of the institutional sector. As the institutional sector grows, firms can pay lower dividends and still attract sufficient monitoring services.
5. Aggregate 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.)
6. Dividends are less likely to be reduced (“stickier” or “smoother”) than share repurchases. (This is in addition to the tax laws which prevent regular share repurchases, but which do not seem to be often enforced.) Dividend reductions would result in a costly reallocation of clientele, and institutional shareholders are the exactly the kind of shareholders able to dislodge the management (rather than tacitly accept a dividend reduction). Thus, it is especially those managers whose firms pay high dividends and who have high institutional ownership (monitoring) which would suffer the most dramatic consequences from cutbacks in payouts. Further, when unexpected dividend reductions do occur, we would predict relatively stronger repercussions for the firm and its managers.
7. 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 institutional

stitutional shareholders are present (discussed in Section 3). Benartzi, Michaely, and Thaler (1997) find that dividend increases are not associated with future better performance, but with lagged better performance. 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 prospect)—which is really an asset-pricing puzzle—the smoothing intrinsic to our theory makes it more robust to timing than some earlier dividend models.

³Both tax-exempt and taxable institutions often have charter restrictions which in effect force them to purchase only companies which pay dividends. A tendency by other institutional shareholders (large investors) to prefer dividend paying stocks would further reinforce our hypothesis, although it could weaken the empirical implications discussed in point 5. 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.

investors act as monitors, and in which institutions face a lower tax rate on dividends than retail investors.⁴

8. If institutions with an ability to monitor also tend to hold firms' debt instruments (because they may again face a lower tax rate relative to retail investors who find it more difficult to shelter interest income), then firms can substitute debt payments for dividend payments. This may provide an additional set of implications: those firms paying dividends should also pay interest. However, 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. Also, some funds and banks are allowed to only hold debt, others only to hold equity. Thus, our theory could help 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.
9. 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 share repurchases or other means of payout for dividend payments.

Our paper now proceeds as follows: Section 1 introduces the signaling model's assumptions. Section 2 solves for the separating equilibrium. For the impatient reader, Section A in the appendix provides some numerical illustrations of the model. Section 3 describes different interpretations of the model's key assumption—that large institutional shareholdings would be associated with higher firm value—and solves a principal-agent version of the model. In this version, institutional shareholders add value not by allowing better manager to signal quality, but instead add value by monitoring the subsequent performance of any manager. Section 4 concludes.

⁴Rajan and Zingales (1995), Table V lists tax rates for seven OECD countries. In 1990, France's institutional investors enjoyed the largest differential, U.S. investors the smallest. (But Rajan and Zingales do not list the relative propensity to conduct share repurchases. Our theory suggests that good firms also tend to use bank and other debt to attract other institutional monitors.)

1 The Model

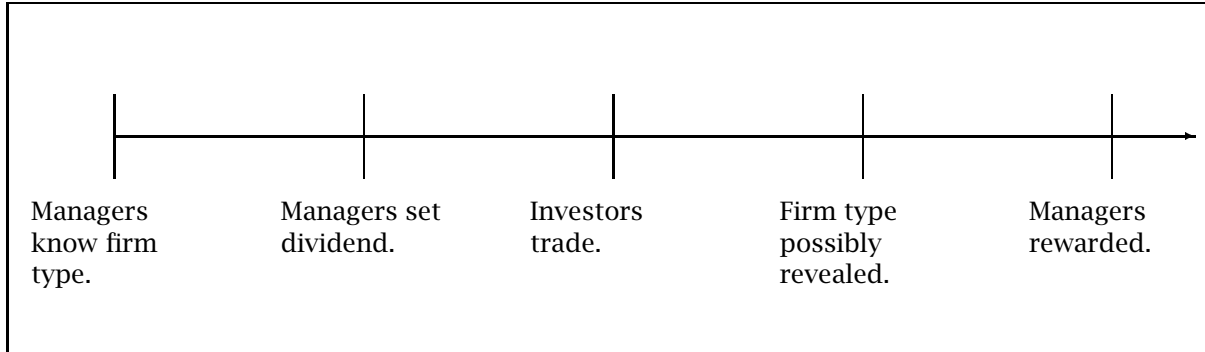
Firm Characteristics: The economy consists of N firms, which can either be of high (H) or low (L) type. The random pre-tax payoff for firms of type $j = H, L$, denoted \tilde{V}_j , are assumed to be normally distributed with mean μ_j and variance σ^2 with $\mu_H > \mu_L$. Consequently, the expected pre-tax payoff 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. 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 managerial compensation increases monotonically with p , the firm's share price. However, if managers of low-quality firms are later revealed to have wasted firm funds by paying dissipative dividends in order to try to obtain a higher share value, she is penalized (e.g., through lower managerial compensation, by being fired, etc.). For simplicity, we assume that managerial utility increases linearly in share price (where m is a proportionality factor), but is normalized to zero when the manager of a *revealed* low-quality firm has paid a dividend.⁵ Managers of the firm know the quality of the firm and are assumed to choose a dividend policy which maximizes their personal expected wealth. We denote the expected payoffs to a manager of true type t_1 who chooses dividends indicating that he is type t_2 as M_{t_1, t_2} . 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 and Investor Types: Without loss of generality, we assume that the capital gains tax rate is zero. Dividends, however, are subject to personal tax: if div-

⁵Because this is an out-of-equilibrium event in our separating equilibrium, this assumption is consistent with our Bayesian equilibrium context. An alternative mechanism that could justify the assumption that managers are penalized is that managers are either low- or high-quality and that institutions want to penalize low-quality managers (for obvious reasons). Then the probability of firing represents the ability of institutions to collect court-verifiable evidence that the manager is in fact low-quality. The greater the dividend the greater is the institutional presence and the more likely it is that such damaging verifiable evidence will be found.

Figure 1: Time Line



dividends D are paid out then the total after-tax payoff (dividends plus capital gains) to shareholders is reduced by $\tau_i \cdot D$, where τ_i is the marginal tax rate on dividend income for investor type i . 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_i and aggregate share endowment $\bar{\theta}_{ij}$ in firm $j = 1, \dots, N$ and we normalize the total share endowment in each of the firms to unity, i.e. $\bar{\theta}_{Ij} + \bar{\theta}_{Rj} = 1$ for all j .

Our model explores two important differences between these investors:

Monitoring/Detection The higher the fraction of the firm owned by institutions, the more likely it is that a manager of a low-quality dividend-paying firm is consequently revealed/discovered to have deceived shareholders. (Section 3 discusses other ways by which the presence of institutions can be associated with better firm performance.)

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 1.

Investor Utility Maximization: The investors problem is to allocate their wealth among the N stocks and a riskless asset which yields a certain gross return which 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. Consumption for representative investor i is given by

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

where θ_{ij} is the demand by an investor of type i in stock of type j , \tilde{V}_j is the dividends paid by and p_j is the shareprice for firms of type j . Investors maximize expected utility of consumption conditional on the information signalled by the firm's dividend policy. Thus, the representative investor solves the following optimization problem:

$$\max_{\theta_{ij}} E \left[-e^{-\gamma_i \tilde{C}_i} | \{D_j\} \right]. \quad (2)$$

Because \tilde{C}_i is normally distributed, equation 2 is equivalent to

$$\max_{\theta_{ij}} -e^{-\gamma_i E(\tilde{C}_i) + \frac{\gamma_i^2}{2} \text{Var}(\tilde{C}_i)}, \quad (3)$$

which has the same solution as

$$\max_{\theta_{ij}} \gamma_i E(\tilde{C}_i) - \frac{\gamma_i^2}{2} \text{Var}(\tilde{C}_i). \quad (4)$$

2 The Separating Equilibrium

We focus attention on a separating equilibrium in which firms with high-quality managers pay a dissipative dividend in order to attract more institutional monitors and signal their higher quality.⁶ Furthermore, we restrict our attention to symmetric equilibria in which all managers of high-quality firms choose the same dividend policy and all managers of low-quality firms choose the same dividend policy.

An important part of our definition of equilibrium will be 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: $\mu^c(D_H^*) = \mu_H$; $\mu^c(D_L^*) = \mu_L$.*
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: $M_{H,H} \geq M_{H,L}$ and $M_{L,L} \geq M_{L,H}$.*

Note that in any reasonable separating equilibrium, low-type investors pay the lowest possible dividend, $D_L^* = 0$, to minimize the tax obligation. Thus, a separating equilibrium requires that investors correctly conjecture that firms paying dividends ($D = D^*$) have expected cash flows of μ_H and firms paying no dividends ($D = D_L^* = 0$) have expected cash flows of μ_L . Given these investor conjectures, investors choose demands optimally and markets clear. Finally, a separating equilibrium

⁶This can be justified with a Cho-Kreps (1987) argument. However, if the dividend tax differential were very small, the costs to signaling with dividends would become very large, and a pooling equilibrium in which no firm pays dividends becomes more plausible. Neither the tax code nor the evidence suggests that this is the case in the major OECD countries.

requires that, given these conjectures by investors, it is optimal for managers of high-quality firms to pay dividends D^* rather than not paying a dividend and it is optimal for managers of low-quality firms to not pay a dividend rather than to mimic managers of high-quality firms.

2.1 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 (i) the after-tax payoffs for a dividend paying firm are normally distributed with mean $\mu_H - \tau_I D^*$ and variance σ^2 with $\tau_I < \tau_R$, by assumption; and (ii) the conjectured after-tax (and pre-tax) payoffs for low-quality firms are normally distributed with mean μ_L and variance σ^2 . The following proposition describes the equilibrium shareholdings and prices in the trading stage given these conjectures:

Proposition 1 *Equilibrium shareholdings and share prices for dividend paying firms are:*

Institutional Shareholdings:

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

Retail Shareholdings:

$$\theta_{RH}^* = \frac{\gamma_I}{\gamma_I + \gamma_R} - \left[\frac{(\tau_R - \tau_I)}{(\gamma_I + \gamma_R)\sigma^2} \right] D^* . \quad (6)$$

Equilibrium Price:

$$p_H = \mu_H - \bar{\tau} D^* - \kappa \sigma^2 , \quad (7)$$

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

Equilibrium shareholdings and share prices for non-dividend paying firms are:

Institutional Shareholdings:

$$\theta_{IL}^* = \frac{\gamma_R}{\gamma_I + \gamma_R}. \quad (8)$$

Retail Shareholdings:

$$\theta_{RL}^* = \frac{\gamma_I}{\gamma_I + \gamma_R}. \quad (9)$$

Equilibrium Price:

$$p_L = \mu_L - \kappa\sigma^2. \quad (10)$$

Proof of Proposition 1 *In a (symmetric) separating equilibrium, investors correctly conjecture that firms paying dividends $D_j = D^* > 0$ are high-quality (i.e. $E[\tilde{V}_j] = \mu_H$) while firms that do not pay dividends are low-quality (i.e. $E[\tilde{V}_j] = \mu_L$). Using this we solve (4) which yields:*

$$\theta_{ij} = \frac{\mu_H - \tau_i D^* - p_j}{\gamma_i \sigma^2} \quad (11)$$

for the dividend paying firms and

$$\theta_{ij} = \frac{\mu_L - p_j}{\gamma_i \sigma^2} \quad (12)$$

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 (11) and (12) yields the desired result.

There are several important features of the equilibrium for dividend paying firms. First, institutional shareholdings consist of an optimal risk sharing component equal to $\frac{\gamma_R}{\gamma_I + \gamma_R}$ and a second component which represents a clientele effect induced by different marginal tax rates. The institutional shareholdings increase linearly in the tax differential. When the tax differential is zero, the equilibrium holdings of the two types are equal to the optimal risk sharing holdings. If institutions have a lower marginal tax rate than individuals they will hold more of the shares of high dividend paying firms. Second, firms paying higher dividends have greater institutional shareholdings when $\tau_I < \tau_R$. Third, the dividend reduces the share price in equilibrium (relative to the full revelation value) because of the tax disadvantage of dividends versus capital gains.

2.2 Optimal Dividend Policies

The competitive equilibrium computed above assumes that dividend policies of the two types are perfectly revealing. For this equilibrium to exist, it must be true that managers of high-quality firms prefer to pay D^* rather than $D = 0$ and it must also be true that it does not pay for managers of low-quality firms to mimic high-quality types by offering the dividend D^* rather than paying $D = 0$. We now examine these issues.

The first stage of the game is analyzed taking as given the competitive equilibrium in the trading stage. In the first stage, managers choose the firm's dividend policy. Although the dividend is dissipative, some managers choose to pay a dividend to signal their quality. For simplicity, managerial compensation is assumed to be a linear function of the share price, mp . If the manager is penalized, she gets a normalized value of zero. Thus, it may be costly for low-quality managers to mimic the high-quality managers. Paying a dividend increases the likelihood that they will be revealed as value-destroying managers, because dividends attract institutions who are assumed to be more effective managerial monitors. Let $\pi(D)$ denote the probability that a manager of a low-quality firm is penalized for deception for a given dividend level D . We assume that $\pi(D)$ is differentiable with $\pi'(D) = \frac{\partial \pi}{\partial \theta_I} \cdot \frac{\partial \theta_I}{\partial D} > 0$ because $\frac{\partial \pi}{\partial \theta_I} > 0$ by assumption and $\frac{\partial \theta_I}{\partial D} = \frac{(\tau_R - \tau_I)}{(y_I + y_R)\sigma^2} > 0$ as we showed above. Furthermore, $\pi(0) = 0$, i.e. if the low-type managers pays no dividend then they have not been deceptive and they will not be penalized

Consider the problem faced by high-quality managers. If they choose $D = 0$ they will not pay a dissipative dividend, but the market will believe the firm is low-quality. Their compensation will be:

$$M_{H,L} = m(\mu_L - \kappa\sigma^2) . \quad (13)$$

If they pay dividends $D = D^*$ their compensation will be

$$M_{H,H} = m(\mu_H - \bar{\tau}D^* - \kappa\sigma^2) . \quad (14)$$

Thus, a necessary condition for a separating equilibrium is:

$$M_{H,H} \geq M_{H,L} \Leftrightarrow D^* \leq \frac{(\mu_H - \mu_L)}{\bar{\tau}}, \quad (15)$$

which simply 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 payoff means, 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, i.e. $D = 0$, they will never be penalized. But the share price will reflect the fact that the market believes the firm is low-quality. In this case, their compensation is

$$M_{L,L} = m(\mu_L - \kappa\sigma^2). \quad (16)$$

If they pay dividends $D = D^*$ managers receive zero compensation with probability $\pi(D^*)$ and mp_H with probability $(1 - \pi(D^*))$. Thus their expected compensation is:

$$M_{L,H} = [1 - \pi(D^*)] [m(\mu_H - \bar{\tau}D^* - \kappa\sigma^2)] + \pi(D^*) [0]. \quad (17)$$

The necessary condition for a separating equilibrium is:

$$M_{L,L} \geq M_{L,H} \Leftrightarrow \pi(D^*) \geq \frac{\mu_H - \mu_L - \bar{\tau}D^*}{\mu_H - \kappa\sigma^2 - \bar{\tau}D^*}. \quad (18)$$

Without loss of generality, we assume that $p_L > 0$, or equivalently $\mu_L > \kappa\sigma^2$. This ensures that the right hand side of (18) is less than one. Furthermore, if (15) is satisfied then the right hand side of (18) is non-negative.

If (15) holds with equality then it is also true that (18) holds (with inequality). In other words, the dividend level $\bar{D} = \frac{\mu_H - \mu_L}{\bar{\tau}} > 0$ will support a separating equilibrium. This choice of dividend, however, is unnecessarily costly. High-quality managers want to choose the lowest D which satisfies both (15) and (18) because it minimizes the dissipative costs of paying dividends. Thus, the optimal level of dividend payouts which supports a separating equilibrium will satisfy (18) with

equality. The following two propositions characterize fully the optimal dividend choice of high-quality managers.

Proposition 2 *There exists a unique D^* satisfying (15) and which satisfies (18) with equality.*

Proof of Proposition 2 *Let $F(D) \equiv \pi(D)(\mu_H - \kappa\sigma^2 - \bar{\tau}D) - (\mu_H - \mu_L - \bar{\tau}D)$. D^* satisfies (18) with equality iff $F(D^*) = 0$. $F(0) = -(\mu_H - \mu_L) < 0$ and $F(\bar{D}) = \pi(\bar{D})(\mu_L - \kappa\sigma^2) > 0$, where $\bar{D} = \frac{(\mu_H - \mu_L)}{\bar{\tau}}$. Furthermore, $F'(D) = \pi'(D)(\mu_H - \kappa\sigma^2 - \bar{\tau}D) + \bar{\tau}(1 - \pi(D)) > 0$. Because $\pi(D)$ is differentiable (and continuous), there exists a unique D^* satisfying (18) with equality. Finally, because $D^* \leq \bar{D}$, it follows that D^* also satisfies (15).*

Proposition 3 *The optimal dividend choice for high-quality firms*

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

Proof of Proposition 3 *At the optimal choice of D^* we have $F(D^*) \equiv \pi(D^*)(\mu_H - \kappa\sigma^2 - \bar{\tau}D^*) - (\mu_H - \mu_L - \bar{\tau}D^*) = 0$. First, note that $F_{D^*} = \pi'(D^*)(\mu_H - \bar{\tau}D^* - \kappa\sigma^2) + \bar{\tau}(1 - \pi(D^*)) = \pi'(D^*)p_H + \bar{\tau}(1 - \pi(D^*)) > 0$, because $\pi' > 0$ and $p_H > 0$ by (15). Implicit differentiation yields all of our comparative statics results:*

1. *Holding $(\tau_R - \tau_I)$ fixed, θ_{IH} is fixed, and then $\frac{\partial D^*}{\partial \bar{\tau}} = \frac{-D^*(1 - \pi(D^*))}{\pi'(D^*)p_H + \bar{\tau}(1 - \pi(D^*))} < 0$;*
2. *Holding $\bar{\tau}$ fixed, $\frac{\partial D^*}{\partial (\tau_R - \tau_I)} = \frac{-\frac{\partial \pi}{\partial \theta_I} \frac{\partial \theta_I}{\partial (\tau_R - \tau_I)} p_H}{\pi'(D^*)p_H + \bar{\tau}(1 - \pi(D^*))} < 0$;*

$$\begin{aligned}
3. \quad \frac{\partial D^*}{\partial \sigma^2} &= \frac{-\frac{\partial \pi}{\partial \theta_I} \frac{\partial \theta_I}{\partial \sigma^2} + \pi \kappa}{\pi'(D^*) p_H + \bar{\tau}(1-\pi(D^*))} > 0; \\
4. \quad \frac{\partial D^*}{\partial \mu_H} &= \frac{(1-\pi)}{\pi'(D^*) p_H + \bar{\tau}(1-\pi(D^*))} > 0; \\
5. \quad \frac{\partial D^*}{\partial \gamma_I} &= \frac{-[\frac{\partial \pi}{\partial \theta_I} \frac{\partial \theta_I}{\partial \gamma_I} p_H - \pi \frac{\partial \kappa}{\partial \gamma_I} \sigma^2 + (1-\pi) \frac{\partial \bar{\tau}}{\partial \gamma_I} D^*]}{\pi'(D^*) p_H + \bar{\tau}(1-\pi(D^*))} \leq 0; \text{ and} \\
6. \quad \frac{\partial D^*}{\partial \gamma_R} &= \frac{-[\frac{\partial \pi}{\partial \theta_I} \frac{\partial \theta_I}{\partial \gamma_R} p_H - \pi \frac{\partial \kappa}{\partial \gamma_R} \sigma^2 + (1-\pi) \frac{\partial \bar{\tau}}{\partial \gamma_R} D^*]}{\pi'(D^*) p_H + \bar{\tau}(1-\pi(D^*))} \leq 0.
\end{aligned}$$

The intuition for these results is as follows. If the weighted average tax rate increases, paying the same dividend is more costly in terms of its effect on equilibrium prices and the benefits of mimicking a high-quality firm becomes smaller. Thus, high-quality firms can pay a smaller dividend and still prevent low-quality firms from mimicking high-quality firms. An increase in $(\tau_R - \tau_I)$ induces stronger clientele effects and the same level of dividends will attract more institutional shareholders, in equilibrium. Thus, it will be more costly for low-quality managers to pay this dividend and a smaller dividend will be sufficient to prevent low-types from mimicking high-types. (Note, however, that footnote 6 argues that the separating equilibrium may break down as the tax differential approaches zero.) If the variance of the stock's payoffs increases the equilibrium price of all stocks falls and low types have less to lose if they are determined to be low types, thus, a higher dividend is necessary to induce separation. When the expected cash flows of the high-type increases it becomes more attractive for low-quality firms to mimic the high-quality firms because it leads to a greater price increase thus a higher dividend is required to induce separation.

An increase in γ_I has two countervailing effects on the optimal dividend. On one hand, an increase in γ_I dampens the clientele effect in the sense that institutions become less aggressive in exploiting the tax differential. Thus, it becomes relatively cheap for low types to mimic high types because they will not be as aggressively monitored. On the other hand, an increase in γ_I increases the preference-weighted average tax rate which makes it relatively expensive to signal quality by paying a dividend. Consequently, the effect of a change in γ_I on the optimal dividend level is ambiguous. Conversely, an increase in γ_R strengthens the clientele effect but

lowers the preference-weighted average tax rate which again leads to an ambiguous effect on the optimal dividend.

Proposition 3 also informs us about the magnitude of the price effect of a dividend announcement. Suppose that prior to trading, all investors believed that a firm is high-quality with probability q and low-quality with $1 - q$. At this time, the equilibrium price of all firms would be identical and equal to an appropriate weighted average of the price of low-quality and high-quality firms. A necessary condition for a separating equilibrium is that firms that pay dividends will have a higher price than firms that don't pay dividends. Consequently, the share price of dividend-paying firms will rise upon the announcement of the dividend. The extent of the price rise is related to the magnitude of the dividend necessary to induce separation because the dividend is dissipative. The following proposition characterizes the price effects of a dividend announcement.

Proposition 4 *Suppose that prior to trading, the share price of each firm reflects an average of the price of a high-quality and low-quality firm. Then the magnitude of the price effect of a dividend announcement*

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

Proof of Proposition 4 *The magnitude of the price effect of a dividend announcement is directly related to the difference between the price of a high-quality firm and a low-quality firm which, in equilibrium, is given by $\Delta \equiv p_H - p_L = \mu_H - \mu_L - \bar{\tau}D^*$. Our comparative statics results follow from Proposition 3:*

1. holding $(\tau_R - \tau_I)$ fixed we have $\frac{\partial \Delta}{\partial \bar{\tau}} = -D^* - \bar{\tau} \frac{\partial D^*}{\partial \bar{\tau}} = \frac{-\pi'(D^*)D^*p_H}{\pi'(D^*)p_H + \bar{\tau}(1-\pi(D^*))} < 0$;
2. holding $\bar{\tau}$ fixed we have $\frac{\partial \Delta}{\partial (\tau_R - \tau_I)} = -\bar{\tau} \frac{\partial D^*}{\partial (\tau_R - \tau_I)} > 0$;
3. $\frac{\partial \Delta}{\partial \sigma^2} = -\bar{\tau} \frac{\partial D^*}{\partial \sigma^2} < 0$;
4. $\frac{\partial \Delta}{\partial \mu_H} = 1 - \bar{\tau} \frac{\partial D^*}{\partial \mu_H} = \frac{\pi'(D^*)p_H}{\pi'(D^*)p_H + \bar{\tau}(1-\pi(D^*))} > 0$;
5. $\frac{\partial \Delta}{\partial y_I} = -\frac{\partial \bar{\tau}}{\partial y_I} D^* - \bar{\tau} \frac{\partial D^*}{\partial y_I} \leq 0$;
6. $\frac{\partial \Delta}{\partial y_R} = -\frac{\partial \bar{\tau}}{\partial y_R} D^* - \bar{\tau} \frac{\partial D^*}{\partial y_R} \leq 0$.

The intuition for these results is as follows. When $\bar{\tau}$ is higher, a given dividend dissipates more firm value. However, a smaller dividend is needed to signal quality. In our model, the second effect dominates and the stock price effect of a dividend announcement is smaller when $\bar{\tau}$ is greater. Holding $\bar{\tau}$ fixed, however, an increase in $\tau_R - \tau_I$ leads to a greater price reaction, because the dissipative effect of the dividend depends only on $\bar{\tau}$ while an increase in $\tau_R - \tau_I$ leads to a smaller dividend to induce separation. Thus, our theory predicts that positive announcement effects depend on the difference in tax treatments and not just the average tax level as suggested by John and Williams (1985).

The appendix provides a numerical illustration of the model.

3 Alternative Mechanisms To Associate Share Value With Institutional Ownership

We have argued that firms pay dividends because dividends attract institutional investors, whose presence is associated with an increased likelihood of better future earnings. The specific structure we used to get this association was a “signaling model,” in which institutions’ information gathering could unveil management quality that worse managers would like to avoid. Managers with good inside information signaled quality by paying dividends.

However, the insight that dividends serve as a mechanism to attract institutions holds more generally than just within an inside information signaling model. It is consistent with other justifications for why firms with more institutional shareholders are associated with better performance such as agency (monitoring), influence (an ability to displace managers detected to be bad), or simply wholesale dumping of shares.

The remainder of this section sketches a simple agency framework, in which the in-equilibrium behavior is similar to the one we derive in our signaling model—firms that pay dividends attract institutions, which is associated with a higher firm value.

3.1 An Agency Framework

In our agency framework, institutions have the ability to increase the value of the firm by monitoring the behaviour of its managers. The ability of institutions to monitor substitutes for the previously assumed ability to differentiate between existing high and low-quality firms, consequently, there are no “good” or “bad” firms and we can omit the subscripts L and H from the parameters μ and p . The model consists of three stages. In the first stage, managers choose a dividend policy to maximize share price. In the second stage, trade between institutional and retail investors occurs. The institutional (retail) shareholdings will depend on the firm’s dividend policy because of the differential tax treatment of dividends to the two types of investors in the same way as our previous signaling model. In the third stage, insti-

tutional investors choose an optimal amount of monitoring, M . We assume that only institutions will monitor because retail investors suffer from the free-rider problem: the retail clientele consists of many investors each of whom have a small stake in the firm and thus have no incentive to expend personal resources to provide monitoring services that benefit all shareholders. The institutions acquire information about the firm and make “suggestions” for improvements to management. These suggestions lead to changes in operating policy which increase the value of the firm. Alternatively, the presence of blockholders may in itself motivate managers to work harder. We assume that M units of monitoring increases firm value by $\ln(M)$ dollars and monitoring costs the institutions c dollars per unit.⁷ All decisions in our model are sequentially rational. Managers are fully aware of the implications of their dividend choice on institutional shareholdings, monitoring, and share prices. Furthermore, retail and institutional shareholders are fully aware of subsequent monitoring choices and incorporate this into their valuations of the firm’s shares. Finally, institutions choose an optimal level of monitoring, given their shareholdings.

We solve the model by backward induction. Given a dividend level D , institutions maximize

$$\max_{\theta_I, M} E \left(-e^{-\gamma_I \tilde{W}_I} \right), \quad (19)$$

where

$$\tilde{W}_I = \theta_I(\tilde{V} + \ln(M) - \tau_I D) + (W_{0,I} - p\theta_I - cM). \quad (20)$$

The two first-order conditions are thus

$$\mu + \ln(M) - \tau_I D - p - \theta_I \gamma_I \sigma^2 = 0 \quad (21)$$

$$M = \frac{\theta_I}{c} \quad (22)$$

⁷To get a meaningful equilibrium we must assume that the cost of monitoring c is small enough to ensure that the optimal M exceeds unity otherwise $\ln(M)$ will be less than zero implying that very small amounts of monitoring decrease firm value. It is straightforward to derive sufficient conditions on exogenous parameters that ensure that the optimal M exceeds unity.

Notice that the optimal level of monitoring is increasing in the institutional shareholdings. This is the mechanism by which dividend paying firms enhance firm value: dividends attract institutions who provide value-enhancing monitoring services. Substituting the known monitoring-given-shareholdings equation (22) into (21) yields the institutional investors' demand function for shares in the second stage of the model:

$$\theta_I = \frac{\mu + \ln(\theta_I/c) - \tau_I D - p}{\gamma_I \sigma^2} \quad (23)$$

Retail investors can predict the level of institutional monitoring in the third stage and solve:

$$\max_{\theta_R} E \left(-e^{-\gamma_R \tilde{W}_R} \right), \quad (24)$$

where

$$\tilde{W}_R = \theta_R (\tilde{V} + \log(\theta_I/c) - \tau_R D) + (W_{0,R} - p\theta_R), \quad (25)$$

which in turn yields the retail investors' demand function for shares in the second stage:

$$\theta_R = \frac{\mu + \log(\theta_I/c) - \tau_I D - p}{\gamma_I \sigma^2}. \quad (26)$$

Finally, we close the model with the market-clearing condition:

$$\theta_I + \theta_R = 1. \quad (27)$$

The equilibrium (θ_I, θ_R, p) cannot be solved in closed form. However, we can solve explicitly for the optimal dividend policy, D^* which is given in the following proposition.

Proposition 5 *In equilibrium, firms pay dividends of*

$$D^* = \frac{\gamma_I + \gamma_R}{\gamma_R \tau_I + \gamma_I \tau_R} - \frac{\gamma_R \sigma^2}{\tau_R - \tau_I} \quad (28)$$

provided the expression on the right hand side is greater than 0. Otherwise, $D^ = 0$.*

Proof of Proposition 5 *Totally differentiate the equilibrium conditions (23), (26), and (27) with respect to D :*

$$\begin{pmatrix} \theta_I^{-1} - \gamma_I \sigma^2 & 0 & -1 \\ \theta_I & -\gamma_R \sigma^2 & -1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial \theta_I}{\partial D} \\ \frac{\partial \theta_R}{\partial D} \\ \frac{\partial p}{\partial D} \end{pmatrix} = \begin{pmatrix} \tau_I \\ \tau_R \\ 0 \end{pmatrix}, \quad (29)$$

thus

$$\begin{pmatrix} \frac{\partial \theta_I}{\partial D} \\ \frac{\partial \theta_R}{\partial D} \\ \frac{\partial p}{\partial D} \end{pmatrix} = \frac{1}{(\gamma_I + \gamma_R) \sigma^2} \begin{pmatrix} -1 & 1 & \gamma_R \sigma^2 \\ 1 & -1 & \gamma_I \sigma^2 \\ -\theta_I^{-1} - \gamma_R \sigma^2 & \theta_I^{-1} - \gamma_I \sigma^2 & (\theta_I^{-1} - \gamma_I \sigma^2) \gamma_R \sigma^2 \end{pmatrix} \begin{pmatrix} \tau_I \\ \tau_R \\ 0 \end{pmatrix}. \quad (30)$$

The share-price maximizing dividend choice solves:

$$\frac{\partial p}{\partial D} = \frac{\theta_I^{-1}(\tau_R - \tau_I) - \sigma^2(\gamma_R \tau_I + \gamma_I \tau_R)}{(\gamma_I + \gamma_R) \sigma^2} = 0. \quad (31)$$

Thus, at the manager's optimum D^* it must be true that

$$\theta_I |_{D^*} = \frac{\tau_R - \tau_I}{\sigma^2(\gamma_R \tau_I + \gamma_I \tau_R)}. \quad (32)$$

We can also derive a second condition which relates the institutional shareholding θ_I to any level of dividends D by combining equations (23) and (26) into

$$\tau_R D + \gamma_R \theta_R \sigma^2 = \tau_I D + \gamma_I \theta_I \sigma^2, \quad (33)$$

and using the condition $\theta_R = (1 - \theta_I)$ to get

$$\theta_I = \frac{\gamma_R \sigma^2 + (\tau_R - \tau_I) D}{(\gamma_I + \gamma_R) \sigma^2}. \quad (34)$$

Combining (32) and (34) yields the desired result. Finally, it is straightforward to check that the second-order conditions are satisfied.

Even though dividends are tax disadvantaged relative to capital gains, dividends increase institutional shareholdings in the firm which increases the endogenous level of monitoring which, in turn, increases the value of the firm. The optimal

level of dividends equates the marginal benefit of monitoring with the marginal dissipative tax cost.

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 above model in which the firm choose its dividend policy each period and institutional monitoring increases firm value only in that period. 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 will 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. Whenever dividends are chosen to attract institutions, e.g. in order to motivate managers, there will be dividend smoothing. The smoothing result is thus likely to hold in a much wider range of circumstances than the specific model used. In a more realistic dynamic setting, there is another effect that reinforces our smoothing result. 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., in an 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 multi-period dynamics reinforces the single-period clientele smoothing choice that arises from the assumed constant parameters.

The ability to detect and monitor may be different from an ability to actually dislodge managers, but the two are correlated. Institutions may be able to open a dialogue with the board of directors or back up outside raiders if the firm performs poorly. Shleifer and Vishny (1986), for example, find that hostile takeover attempts are more likely to occur if there are large, unaffiliated blockholders.

Smith (1996) examines 51 firms targeted by CalPERS from 1987–93. CalPERS, the “California Pension Employee Retirement System,” manages about US-\$125 billion in assets as of early 1998 and is widely regarded as a leader in shareholder activism. CalPERS typically held large stakes in the aforementioned 51 firms, which had experienced poor stock price performance. 72% of these firms eventually adopted proposed CalPERS changes. 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.⁸ Ayres and Cramton (1994) argue that long-term institutional investors have particularly strong influence on firms. Because pension funds tend to be disproportionately long-term investors, they are also more likely to be able to play the role attributed to them in our paper. Pension funds also need not exert pressure on firms directly—according to *Pension and Investments* (1995) about 60% of their assets are allocated to external managers, who in turn are replaced if they perform poorly.

We end with a note that it would be difficult to empirically decompose the activities of institutional shareholders into “signaling detection” and facilitation of management “monitoring.” And even without direct detection or monitoring, institutions may still be able to penalize poor performance. Institutions might sell shares in bulk when they discover the firm is underperforming. Because the firm that pays dividends will then be worse off than if it had just admitted it was bad initially (by not paying dividends), separation may still occur.

⁸The assumption of activism by more organized and larger shareholders is also present in other work. For example, Admati, Pfleiderer and Zechner (1992) show that large shareholders tend to monitor more, even though smaller shareholders successfully freeride. Titman and Trueman (1986) show that it can be ex-post consistent for an institution to investigate, even though the very act of soliciting this information identifies the firm to be of higher quality.

4 Conclusion

This paper has offered 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 offered some implications on the effects of tax code changes, on international evidence, 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 a 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 have to cut the dividends in response to poorer performance.

References

- Admati, Anat R., Paul Pfleiderer, and Josef Zechner, 1994, "Large shareholder activism, risk sharing, and financial market equilibrium." *Journal of Political Economy*, 102 (6), 1097-1130.
- Allen, Franklin and Roni Michaely (1995), "Dividend Policy," Chapter 25 in *North-Holland Handbooks of Operations Research and Management Science: Finance* (Volume 9) edited by R. A. Jarrow, V. Maksimovic and W. T. Ziemba, New York: Elsevier, 1995, 793-837.
- Ayres, Ian, and Peter Cramton, 1994, "Relational Investing and Agency Theory." *Cardozo Law Review* 15, 1033-1066.
- Benartzi, Shlomo, Roni Michaely, and Richard Thaler, 1997, "Do Changes in Dividends Signal the Future or the Past?" *The Journal of Finance*, 52 (3), 1007-1034
- Bagwell, Laurie Simon, 1991, "Shareholder Heterogeneity: Evidence and Implications," *American Economic Review*, 81 (2), 218-221.
- Barclay, Michael J. and Clifford W. Smith, Jr., 1988, "Corporate Payout Policy: Cash Dividends versus Open-Market Repurchases," *Journal of Financial Economics*, 22 (1), 61-82.
- Brennan, Michael J. and Anjan V. Thakor, 1990, "Shareholder Preferences and Dividend Policy," *Journal of Finance*, 45 (4), 993-1019.
- Cho, In-Koo and David Kreps, 1987, "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102 (2), 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.
- Fama, Eugene F. and Harvey Babiak, 1968, "Dividend Policy: An Empirical Analysis," *Journal of the American Statistical Association*, 63 (324), 1132-1161.
- 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 (1), 121-154.
- John, Kose and Joseph Williams, 1985, "Dividends, Dilution, and Taxes: A Signaling Equilibrium," *Journal of Finance*, 40 (4), 1053-1070.
- Lintner, John, 1956, "Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes," *American Economic Review*, 46 (2), 97-113.
- Lucas, Deborah J. and Robert L. McDonald, 1998, "Shareholder Heterogeneity, Adverse Selection, and Payout Policy," *Journal of Financial and Quantitative Analysis*, forthcoming.
- 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 (2), 365-394.
- Rajan, Raghuram G., and Luigi Zingales, 1995, "What do we know about Capital Structure? Some Evidence from International Data," *Journal of Finance*, 50 (5), 1421-1460.
- 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* 51-1, 227-252.

Titman, Sheridan, and Brett Trueman, 1986, "Information Quality and the Valuation of New Issues." *Journal of Accounting and Economics*, 8 (2), 159-172.

Zwiebel, Jeffrey, 1996, "Dynamic Capital Structure under Managerial Entrenchment," *American Economic Review*, 86 (5), 1197-1215.

A Appendix: A Numerical Illustration of the Model

For the sake of this numerical example, assume

- The probability of detection is just the proportion of institutional investors, i.e., if institutions hold 100%, type is always discovered, if institutions hold 0%, type is never discovered.

$$\pi(D) = \theta_I(D)$$

- The pre-tax (pre-dividend) expected value of a high-quality firm is \$10, the pre-tax expected value of a low-quality firm is \$5.

$$\mu_H = 10, \mu_L = 5$$

The variance of realized value of both types of firm is \$1.

$$\sigma^2 = 1$$

- Institutions pay no tax, retail investors pay 50% tax on dividends.

$$\tau_I = 0, \tau_R = 50\%$$

We now consider three cases. In the first, risk aversion of both institutional and retail investors (aggregate) is the same. In the second, institutions are more risk-averse. In the third, retail investors are more risk averse.

A.1 Identical Risk Aversion

Both the institutional and the retail investors have the same risk aversion

$$\gamma_I = 2, \gamma_R = 2.$$

Solution:

- High-quality firms pay \$0.404 dividends and are worth \$8.899 (pre-dividend-taxes) today.
- Low-quality firms pay no dividends and are worth \$4 today.
- Institutions hold 55%, retail investors hold 45% of the high-quality firm.

- Ergo, the probability of “natural” quality revelation/detection for dividend-paying firms is 55%.

To be a separating equilibrium, we need to show that high-quality firms prefer paying dividends than be confused for low-quality firms; that low-quality firms are as well off with a price of \$4 today as they would be if they were to imitate high-quality firms; and that investors are willing to pay the aforementioned amounts for both firms.

- The high-quality firm’s constraint is never binding (not shown here). Equation (7) shows that such firms receive a price equivalent to the weighted post-dividend mean payoff, minus the standard exponential-utility penalty for risk,

$$p_H = \$10 - 25\% \cdot \$0.404 - 1 \cdot \$1 = \$8.899$$

This price will be justified below. High-quality managers have price-proportional preferences, so $M_{H,H} = m \cdot \$8.899$ (equation (14)).

- If the low-quality firm were to “cheat,” it would attract 55% institutional investors, making its probability of revelation 55%. By definition of equation (15), a discovered deceptive manager is penalized and receives zero utility. Thus, 55% of the time, such a deceptive manager would receive utility zero; 45% of the time, he would receive the utility from a price of \$8.899.⁹ Because $55\% \cdot \$0 + 45\% \cdot \$8.899 = \$4$, low-quality managers are as happy with just revealing their identity as they would be falsely imitating high-quality firms.

- Now, why does the market pay \$8.899 for the high-quality firm and attract 55% institutions? First note that the market can infer firm quality in equilibrium. So the problem boils down to fair pricing and appropriate allocations, given the firms’ true identities. Given the assumed exponential utility functions and normally distributed randomness on payoffs, this is a standard problem, and solved in Proposition 1.

The claim is that the equilibrium price p_H of the high-quality firm is given in equation (7) as

$$p_H = \$10 - 25\% \cdot \$0.404 - 1\$1 = \$8.899 ;$$

the price of the low-quality firm is given in equation (10) as

$$p_L = \$5 - 1\$1 = \$4 ;$$

and that, at these prices, institutional investors want to hold $\theta_{IH} = 55\%$ of the dividend-paying stock (equation (5)) and $\theta_{IL} = 50\%$ of the non-dividend-paying stock (equation (8)). Ergo, at these prices, retail investors must want to hold $\theta_{RH} = 45\%$ of the dividend-paying stock (equation (9)), $\theta_{RL} = 50\%$ of the non-dividend-paying stock (equation (9)).

If non-taxable institutional investors do as we say, their consumption is given in equation (1) as

$$\tilde{C}_I \equiv 55\% \cdot (\tilde{V}_H - 0\% \cdot \$0.404) + 50\% \cdot (\tilde{V}_L) + \left[W_I - (55\% - \bar{\theta}_H) \cdot \$8.899 - (50\% - \bar{\theta}_L) \cdot \$4 \right] ,$$

where the following variables will be irrelevant (drop out in the maximization): W is their original total endowment, $\bar{\theta}_H$ is their endowment in shares of firm H and $\bar{\theta}_L$ is their endowment in shares

⁹There is a proportionality factor, m , in translating market prices into managerial utility, but it does not matter.

of firm L . (Note: the first two terms are the random payoffs from holding 55% and 50% of the high- and low-quality firms respectively. $(55\% - \bar{\theta}_H)$ is the percent of shares in high-quality firms that institutions do not yet own, i.e., that they still have to purchase at a price of \$8.899.) Rearranging, for the institutions,

$$\tilde{C}_I \equiv 55\%(\tilde{V}_H - \$8.899) + 50\%(\tilde{V}_L - \$4) + \begin{matrix} \text{non-random endowment terms} \\ \text{independent of } \theta\text{'s} \end{matrix}$$

Taking expectations and variances,

$$\begin{aligned} E(\tilde{C}_I) &\equiv 55\% \cdot (\$10 - \$8.899) + 50\% \cdot (\$5 - \$4) + E\left(\begin{matrix} \text{non-random endowment terms} \\ \text{independent of } \theta\text{'s} \end{matrix}\right) \\ &= 55\% \cdot \$1.101 + 50\% \cdot \$1 + \left(\begin{matrix} \text{non-random endowment terms} \\ \text{independent of } \theta\text{'s} \end{matrix}\right) \\ \sigma^2(\tilde{C}_I) &\equiv 55\%^2 \cdot \$1 + 50\%^2 \cdot \$1 = 55\%^2 + 50\%^2 = 55.25\% \end{aligned}$$

(We assume zero covariance across stocks.) It follows that with an exponential risk-aversion coefficient of “2” and returning to use θ as variable name for “purchased share quantity,” the institutions’ maximization is given in equation (3) as

$$\propto \max_{\theta_H, \theta_L} -e^{-2(\theta_H \cdot \$1.101 + \theta_L \cdot \$1) + \frac{2^2}{2}(\theta_H^2 + \theta_L^2)}$$

It is straightforward to check that institutional investors indeed choose

$$\theta_H = 55\%, \theta_L = 50\% .$$

Going through the same exercise for dividend-paying retail investors shows that they have only one extra term, a subtraction of total dividend taxes of \$0.202 on dividends of \$0.404. Thus,

$$\begin{aligned} E(\tilde{C}_R) &\equiv \theta_H \cdot (\$10 - \$8.899 - \$0.202) + \theta_L \cdot (\$5 - \$4) + E\left(\begin{matrix} \text{non-random endowment terms} \\ \text{independent of } \theta\text{'s} \end{matrix}\right) \\ \sigma^2(\tilde{C}_R) &\equiv \theta_H^2 \cdot \$1 + \theta_L^2 \cdot \$1 \end{aligned}$$

and the retail investors’ utility optimization yields

$$\theta_H = 45\%, \theta_L = 50\% .$$

Clearly, the prices of $p_H = \$8.899$ and $p_L = \$4$, which we assumed in solving for institutional and retail demand (solved generally in equations (7) and (10)), clear the markets for both stocks, because institutional and retail demand add up to 100%.

A.2 Institutions have higher risk aversion

Now assume that institutional investors have a risk aversion parameter of 10, retail investors a risk aversion parameter of 2:

$$y_I = 10, y_R = 2$$

Solution:

$$D^* = 6.144; P_H = 5.774; P_L = 3.333;$$

$$\pi = 0.423; M_{L,L} = 3.333; M_{L,H} = 3.333; M_{H,H} = 5.774; M_{H,L} = 3.333$$

Again, notice that it does not pay for either type to mimic the other type. Even though the optimal dividend is higher in this example, the probability of detection falls, because institutional risk-aversion, γ_I , rises, implying that institutions will hold less of the firm in equilibrium.

A.3 Retail Investors have higher risk aversion

Now assume that institutional investors have a risk aversion parameter of 2, retail investors a risk aversion parameter of 2.5:

$$\gamma_I = 2, \gamma_R = 2.5$$

Solution:

$$D^* = 0.057; P_H = 8.876; P_L = 3.889;$$

$$\pi = 0.562; M_{L,L} = 3.889; M_{L,H} = 3.889; M_{H,H} = 8.876; M_{H,L} = 3.889$$

Please note that when $\gamma_I \ll \gamma_R$, $\pi = \theta_I$ could go above 1. In other words, we cannot allow parameter γ_R to be extremely high, because retail investors' tiny benefit from capturing risk-sharing surplus could be overwhelmed by their desire to avoid taxes. In this case, we would fall outside the separating equilibrium, which prevents us from working out solutions. (We therefore just restrict our attention to parameter values in which the separating equilibrium is feasible.)

Table 1: List of Symbols

Abbreviation	Page	Description
Exogenous Variables: Firms and Managers		
N	8	Number of firm types. (Two in our paper.)
$j \in \{H, L\}$	8	Firm type (quality): <i>High</i> or <i>Low</i> .
$\tilde{V}_j \sim N(\mu_j, \sigma^2)$	8	Before-tax firm value, distributed with mean μ_j and variance σ^2 .
$\pi(\cdot)$	14	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., $\pi \equiv \pi(\theta_I(D)) \equiv \pi(D)$.
$M_{t1,t2}$	8	Expected utility to manager of type $t1$, claiming to be a manager of type $t2$.
m	8	Multiplies share price into managerial utility, if manager is not penalized.
Exogenous Variables: Investors		
$t \in I, R$	9	Shareholder Type, where I denotes institutional shareholders, R denotes retail shareholders.
τ_t	9	Tax rate on dividends faced by investors of class t . Assumed: $\tau_I < \tau_R$.
$\bar{\tau}_t$	12	Preference weighted average tax rate, $(\tau_I \gamma_R + \tau_R \gamma_I) / (\gamma_R + \gamma_I)$.
W_t	9	Aggregate wealth of investors of class t
γ_t	9	Risk aversion of investors by class t
κ	12	Inverse of the sum of risk tolerances, $[\frac{1}{\gamma_I} + \frac{1}{\gamma_R}]^{-1}$.
$\bar{\theta}_{t,j}$	9	(Aggregate) endowments of holdings in firm j invested by investors of class t .
Endogenous Variables		
$\theta_{t,j}$	10	(Aggregate) percent of holdings in firm j invested by investors of class t . (Note: θ 's add to 1 across all investors for each firm, but not across all firms for each investor.)
p_j	10	Share price for firms of type j .
D_j	10	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 $D = 0$. Page 2 defines $\bar{D} = \frac{(\mu_H - \mu_L)}{\tau}$, an upper limit of D^* .
Agency Model Variables (Section 3.3.1)		
c	20	Cost Per Unit of Monitoring. Assumed to be "small."
M	20	Amount of Monitoring (Endogenous)