Queen's Economics Department Working Paper No. 895

Testing Dividend Signalling Models

Dan Bernhardt J. Fiona Robertson Ray Farrow

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-1994

Discussion Paper #895

Testing Dividend Signalling Models

by

Dan Bernhardt
Queen's University

J. Fiona Robertson
Seattle University

Ray Farrow
Seattle University

January 1994

TESTING DIVIDEND SIGNALLING MODELS.

Dan Bernhardt[1], J. Fiona Robertson[2] and Ray Farrow

January, 1994.

[1] Department of Economics
Queen's University
Kingston, Ontario
Canada, K7L 3N6
bernhard@qucdn.queensu.ca

[2] Department of Economics and Finance.
Albers School of Business and Economics
Seattle University
Seattle, Washington.
fiona@seattleu.edu

Abstract

This paper derives a key monotonicity property common to dividend signalling models: the greater the rate that dividend income is taxed relative to capital gains income, the greater the value of information revealed by a given dividend yield, and hence the greater the associated excess return. This monotonicity condition allows us to distinguish the hypothesis that dividends are used as a signalling device from the hypothesis that dividends contain information but are not used as Spencian signals. The monotonicity conditions are tested with robust non-parametric techniques. Although we find strong evidence that dividend announcements contain information, we find no evidence to support dividend signalling. The same results are inconsistent with tax-based CAPM arguments.

Keywords: dividends, signalling, non-parametric.

JEL Classification: G0

We appreciate the comments and suggestions of Bart Lipman and Jeff Strnad. Seminar participants at Caltech and Queen's University provided useful feedback. The first author acknowledges support from the SSHRC. All errors remain our own.

I Introduction.

Since Ross (1977) and Bhattacharya (1979), financial economists have explored the possible signalling properties of dividends and other financial activities. Research into the signalling properties of dividends has been motivated by an attempt to explain the apparent excess returns observed following announcements by firms of favorable dividends. Building on the work of Spence (1973), Bhattacharya (1979) produced an internally consistent model of Modigliani and Miller's (1963) "informational content of dividends hypothesis", demonstrating how dividends could allow insiders to credibly communicate information about the expected future value of the firm to less informed outsiders.[1] Credibility of the signal requires that it not pay low quality firms to mimic the behavior of high quality firms.

The majority of the empirical studies of the "information content of dividends hypothesis" have used an event study methodology to investigate the response of share prices to the announcement of changes in dividend levels.[2] Many of these papers found evidence supportive of the view that there appears to be a stock price response to changes in firms' dividend policies (Aharony and Swary 1980, Asquith and Mullins 1983, Brickley 1983, Charest 1978, Fama, Fischer, Jensen and Roll 1969, Ghosh and Woolridge 1988, Kalay 1980, Kalay and Lowenstein 1986, Laub 1976, Patell and Wolfson 1984, Pettit 1972, 1976). Such results have been viewed as supportive of the view that dividend announcements are interpreted by the market as being informative of firm value. A few early studies focused on examining the predictive content of dividend policy for future share performance, rather than investigating the existence of an announcement effect, and found less evidence to support the view that dividend announcements convey additional information to the market (Ang 1975, Gonedes 1978, Penman 1983, Watts 1973, 1978).

---

[1] Credibility of the signals here refers to the recognition that the signalling aspect of financial policy must be immune to the possibility that insiders could strategically manipulate the signals sent in such a way as to allow them to benefit from temporary mispricing of the firm's shares.

[2] Notable exceptions are Kalay (1982), John and Mishra (1990), and John and Lang (1991), which are discussed below.

Eades (1982) takes a different approach to testing dividend signalling models. Rather than trying to detect the presence of an announcement effect from dividend changes or examine the predictive content of dividend policy, Eades derives and tests specific predictions obtained from comparative static analysis of a formal dividend signalling model.[3] Eades analyzes a version of Bhattacharya's signalling model and derives two testable hypotheses: (1) an inverse relationship between dividend yield and a firm's own variance of returns; and (2) the "relative signalling strength hypothesis" (RSS) which states that "higher risk firms exhibit stronger changes in value relative to their lower risk counterparts for any given change in dividends" (p. 473). Eades finds empirical support for the first of these results but strongly rejects the RSS hypothesis.

This paper's test of signalling theories of dividends is in the spirit of Eades in that we derive and test a comparative static result derived from dividend signalling models. The RSS hypothesis tested by Eades was derived from a particular model of dividend signalling with particular specifications of functional forms and distributional assumptions within that model. In contrast we test a non-parametric comparative static result that holds for a class of dividend signalling models that encompasses virtually every published dividend signalling model, rather than a comparative static derived from a particular example of such a model.

We argue that a testable hypothesis that can be derived within the context of almost any Spencian signalling theory of dividends is that there is a monotonic relationship between the marginal tax rate on dividend income relative to capital gains and the amount of "good news" revealed by any sized dividend yield *level*. The higher is the relative tax rate on dividends, the better is the "type" revealed by any level of dividend yield, and hence, under rational expectations, the greater

---

[3]John and Mishra also derive comparative statics from an explicit signalling model. They allude to empirical evidence from other researchers consistent with their predictions. John and Lang (1991) derive testable implications from a specific signalling model regarding correlations between the announcement of dividend changes, the extent of insider trade and the response of stock prices. They find weak empirical supporting results.

the associated excess return. This result holds *even* in signalling models in which the underlying signalling argument is *not* tax based. It simply reflects the fact that the higher the marginal tax rate on income, the more costly the dividend signal.

This observation suggests a simple and robust method for testing signalling theories of dividends. Over the period 1962-1988 there have been numerous changes in the US Federal tax code governing the taxation of dividend income and capital gains in the US. We identify 16 distinct tax regimes and order them from most favorable to dividend income relative to capital gains income, to least favorable. Signalling theories predict that for a *given* dividend yield *level*, the good news released, and hence the associated excess return, should be least when dividend income receives the most favorable tax treatment, and the excess return should be greatest when dividend income receives the least favorable tax treatment.[4] We refer to this as the monotonicity property of signalling models. It is important to note that we can distinguish between whether information is released as an indirect by-product of dividend announcements (e.g. higher dividends reveal more cash on hand) and whether information is released as a result of Spencian signalling. The former hypothesis predicts that information release should be positive, but independent of the tax regime.

Since a generic dividend signalling model does not predict a particular functional form for the relationship between tax rates and excess returns for a given dividend yield, we test this monotonicity relationship without imposing particular parametric restrictions. In particular, we employ non-parametric tests of rank order correlation (Kendall's tau and Heoffding's distribution free tests) to test for the predicted monotonic relationship between tax regime and excess returns for large portfolios of firms with similar dividend yields. There are several advantages of this approach. The non-parametric tests are robust, allow us to control for the dividend yield levels, and we do

---

[4]Note that since the tax signalling cost is a *static* function of the *level* and *not change* in the dividend yield, theory predicts no particular relationship between excess return and *change* in dividend yield. The theoretical development in Section 2 makes this clear.

not have to worry about mis-specifying the functional form. Further, by aggregating into large representative portfolios our analysis is robust to cross-sectional heterogeneity in firms (e.g. in payoff functions, information).

We find strong evidence that dividend announcements contain payoff relevant information: dividend yield is highly positively correlated with the excess returns earned following the announcements. However, we find *no* evidence of correlation between the relative tax treatment of dividends and the excess returns earned following dividend announcements (controlling for the size of the dividend payout and firm size). We thus find no evidence to support the view that dividends are used to signal firm value. More precisely, the paper provides strong evidence against the joint hypothesis that the marginal investor is taxed and that dividend yield serves as a Spencian signal of firm value. Peterson, Peterson and Ang (1985) examine income tax returns and estimate a marginal effective tax rate on dividend income of 30% for 1979, suggesting that the marginal investor is taxed.

This evidence is also inconsistent with the tax-based CAPM models of Brennan (1979) and Litzenberger and Ramaswamy (1979, 1980, 1983). There, investors demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate for the higher tax cost of dividends relative to capital gains; higher dividend taxes raise the required pre-tax return.[5]

Section 2 derives in a general context the monotonicity property that is the basis for our test. Section 3 examines the theoretic literature on dividend signalling, identifying that the monotonicity property holds in *each* paper. Section 4 discusses the various tax code changes that have occurred in the treatment of capital gains taxes in the U.S. over the period 1960-1988. Section 5 details the non-parametric tests employed and explains the ways in which the analysis is robust. Section 6 describes the data. Section 7 details our findings. Section 8 discusses the results and draws conclusions.

----

[5]A recent study by Christie and Huang (1992) focuses on the tax effect of dividends across tax regimes.

II. A General Signalling Theory.

The essential relationship that we wish to test is a key comparative static that we derive in a generic dividend signalling model. This result holds that the level of dividend payout needed to signal any given level of hidden firm characteristic is lower the higher the marginal tax rate on dividend income relative to capital gains. The simple intuition behind this result is that any signalling model implies selecting an optimal level of dividend payout by equating the marginal cost of the signal level to the marginal benefit of the signal level. The marginal cost of the signal (dividend payout) is a strictly increasing function of the marginal tax rate on dividends and a decreasing function of the marginal tax rate on capital gains. This monotonic relationship holds whether or not the signalling aspect of dividends derives from the tax rate. Since the marginal benefit of the signal is independent of the tax rate, we obtain the predicted inverse relationship between dividend payout levels and dividend tax rates for any quality of firm that signals -- any given amount of good news can be signalled with a lower dividend level if the marginal cost of dividend payouts is higher.

Consider the following general (scalar) dividend signalling model. Let D denote the dividend (signal) level, P denote the market price of a firm's shares, and θ denote some characteristic known only to the informed manager(s) of the firm. θ is the variable to be signalled to the uninformed. The informed's welfare depends on both the current share price and the true value of the firm.[6] Denote this relationship $V(P, \theta)$, where $\partial V/\partial P > 0$. In a signalling equilibrium P depends monotonically on D, P(D) with $P' > 0$, so that we can write $V(P(D), \theta)$.

There is some cost associated with sending the signal (issuing

---

[6]This is a critical characteristic of all dividend signalling models. Absent a dependence on current share price, the insiders would have no incentive to signal firm value to outsiders since they would not benefit from the transmission of information. Absent some concern for future share value (once the truth becomes known to all) the insiders always wish to raise current share price by signalling good news, unconcerned about the future impact of their false signal, so that no credible signal can emerge.

dividends). Part of this cost is the tax cost, τD. In order for a signalling equilibrium to exist, the marginal cost of dividend issuance must depend on firm type. Hence we write the total cost of dividends as $C(D,\theta) + \tau D$ (or more generally as $C(D, \tau, \theta)$, where $\partial C/\partial D > 0$, $dC/\partial\tau > 0$, $\partial^2 C/\partial D\partial\tau > 0$).

The informed choose D to maximize $V(P(D), \theta) - C(D, \tau, \theta)$, taking account of the dependence of P on D. In a signalling equilibrium, P(D) must be "informationally consistent" (Riley, 1979), so that $P(D) = \Pi(\theta, D)$, where $\Pi(.)$ denotes the "true" market value of a firm of type θ paying dividend D, and $\partial\Pi/\partial\theta > 0$ so that higher values of θ correspond to higher quality firms. An optimal signalling equilibrium requires selecting the level D^* that solves:

$$\max_D \quad V(P(D), \theta) - C(D, \tau, \theta), \qquad (1)$$
$$\text{where } P(D) = \Pi(\theta, D).$$

As Riley shows, two necessary and sufficient conditions for the existence of a solution to (1) with $P(D) \neq \overline{P}$, $\forall$ D (i.e. that involves signalling) are:

(1) $\exists\ \hat{D} < \infty$ that maximizes firm value absent any signalling effects.[7]

(2) The "single-crossing property": $$\frac{\partial}{\partial\theta}\left(\frac{\partial C(D,\tau,\theta)/\partial D}{\partial V(P(D),\theta)/\partial P}\right) < 0, \ \forall\theta. \qquad (2)$$

The existence of a positive tax differential on dividends versus capital gains ensures that condition 1 is satisfied since the optimal dividend level absent signalling is then zero. If there are no direct benefits to the signalling activity that depend on θ,[8] then condition (2) reduces to $\partial^2 C/\partial D\partial\theta < 0$; marginal signalling costs must be inversely related to firm quality.

Solving the first order conditions (FOC) for problem (1), we obtain

---

[7]This is needed to prevent all firms from setting D at such a high level that it is impossible to have any information transmitted.

[8]This occurs for instance in Spence's signalling model where education adds no value but simply acts as a dissipative signal.

$$\frac{\partial V(P(D),\ \theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial \theta}\ \frac{d\theta}{dD} + \frac{\partial V(P(D),\ \theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial D} = \frac{\partial C(D,\ \theta,\ \tau)}{\partial D},$$

or, for the special case where $C(D,\ \tau,\ \theta) = C(D,\ \theta) + \tau D$,

$$\frac{\partial V(P(D),\theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial \theta}\ \frac{d\theta}{dD} + \frac{\partial V(P(D),\theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial D} = \frac{\partial C(D,\ \theta)}{\partial D} + \tau.$$

This FOC characterizes a differential equation for the optimal signalling function $D(\theta)$. Infinitely many solutions to this FOC exist. Identification of a particular signalling equilibrium is typically achieved by identifying the most efficient of all the solutions to the differential equation (see e.g. Riley 1979, Cho and Kreps 1987, or Banks and Sobel 1987). In the presence of dissipative signals (such as dividend taxes), this selection procedure identifies the solution in which the lowest quality firm selects a zero dividend. We should emphasize, though, that our empirical analysis is robust to the particular equilibrium selection, provided that the same equilibrium is selected over time. To derive the impact of changes in dividend taxes on the optimal dividend payout level for any quality of firm, we carry out the comparative static:

$$dD^*/d\tau = V_p \Pi_{DD} \theta' + V_p \Pi_{DD} - C_{DD} < 0 \quad \text{from SOC.} \tag{3}$$

An increase in marginal tax rate, τ, leads the level of the dividend, D, to be set at a point where the net marginal benefit is higher, which, by second order conditions, ensures that $dD^*/d\tau < 0$.

Notice that in the absence of some direct benefit of dividends, dividend taxes alone do not satisfy the single-crossing property (2) since $\partial^2 C/\partial\tau\partial\theta = 0$. Thus, if dividends are to act as a signal in a scalar signalling model, some other aspect of dividend costs must generate the needed relationship between marginal dividend costs and firm quality that provides separation. However, *even* if dividend taxes are *not* the feature of the model that generates single-crossing, equation (3) shows that the optimal dividend level still depends on τ.

Engers (1987) shows that in multiple signal models only a quasi-concavity condition on the signalling cost function is required for a signalling equilibrium to exist. In particular, each signal need <u>not</u> satisfy the single-crossing property individually. Engers' condition is satisfied when there are two signals and the cost of one satisfies single crossing and the cost of the other is linear in the signal level. This is important since it means that if the other signal exhibits increasing marginal cost with quality, then dividends can act as a signal even if the only cost of dividends is a (linear) tax cost.

To derive the monotonicity condition in a multiple signalling model, note that with multiple signals the firm chooses the most efficient (cost-minimizing) mix of signals. The cost-minimizing mix of signals is determined where the marginal cost of each signal is equated across all signals used. An increase in the marginal cost of one signal (as would occur with an increase in the tax on dividend payouts) leads to an equilibrium in which the mix of signals is altered so that less signalling is done with the relatively more expensive dividend signal.

It is important to note that even though the information to be revealed by a dividend signal may be dynamic in nature, reflecting, for instance, an increase in the revenues derived from a project, the tax signalling cost is a *static* function of the *level* and *not change* in the dividend yield. That is, the signalling cost reflects levels of the dividend and the tax regime: $C(D, \tau, \theta) = C(D, \theta) + \tau D$. For this reason, any tests of signalling theories of dividends must focus on the level and not change in the dividend yield.

To see this most clearly, consider two firms, A and B, which both have good news to signal in consecutive periods. Firm A has better news than firm B to signal in both periods, but firm A has better news in period 1 than period 2, while firm B has better news in period 2 than period 1. That is, good news is ordered by:

$$R_1^A > R_2^A > R_2^B > R_1^B.$$

The equilibrium levels of the dividend signals reflects this ordering,

$$D_1^A > D_2^A > D_2^B > D_1^B,$$

but were one to look at the change in the dividend yields, one would find that

$$D_2^A - D_1^A < 0 < D_2^B - D_1^B.$$

Since the excess returns reflect the news revealed, they would satisfy

$$ER_2^A > ER_2^B.$$

A researcher looking for a positive relationship between excess return and change in dividend yield as evidence of dividend signalling would incorrectly conclude that firms were *not* using dividends to signal. This observation is of interest because there is evidence that changes (increases) in dividend levels are associated with greater excess returns than are levels of dividend yields. Signalling theories of dividends, however, make no prediction about the relationship between changes in dividend levels and excess returns[9]. Due to the static nature of the signalling cost, the generic dividend signalling model requires that the entire dividend be unanticipated. Since dividend levels are very predictable, this suggests that one should not be surprised that we reject signalling theories of dividends.

III. Dividend Signalling Literature.

Scalar signaling models (Bhattacharya 1979, Kalay 1980, Talmor 1982, Hakanson 1982, Miller and Rock 1985, and Bar Yosef and Hoffman 1986) seek to explain both dividend payouts and their informational content. The models differ in two regards, the form of dividend cost function, $C(D, \tau, \theta)$, and the motivation behind why the informed's objective function is of the form $V(P(D), \theta)$ rather than being concerned solely with $P(D)$ (current value) or solely with future value. Most assume that the cost of dividend issuance[10] arises due to the corporate transaction costs of refinancing cash shortfalls. To motivate the form of objective function they either assume that the form of managerial reward scheme produces a concern for both current and true

---

[9]Differential taxation implies that dividend yield levels always signal information. Still, one could imagine a model in which change in dividend yield levels also conveyed information. Our tests should be robust to such a formulation because we employ large representative portfolios of firms. Intratemporally, aggregation should wash out change effects in our level portfolios. Further, inter-temporally, given dividend level portfolios should have the same cross-sectional distribution of dividend yield changes, so that, holding the tax regime constant, the excess return for a given dividend yield portfolio should be the same across years, allowing us to base our tests solely on levels.

[10]All of these models have the feature that the dissipative signalling cost of dividends is a function of the level of, and not change in, dividend.

value or that managers act in the interests of current firm shareholders whose time horizon is such that they care about both current share prices and future share prices. Dividend payments by firms act as a signal of either current or expected firm value. Managers of high quality firms pay a dividend just high enough to distinguish themselves from low quality firms, which are discouraged from mimicking the behavior of the high quality firm by the greater probability of having to turn to costly external financing to pay the dividend given their lower expected cash flow. The benefit of paying a dividend is the same for all firm types (the positive current stock price return), but the cost is higher for a low quality firm because it is more likely to have to resort to external finance.

Multiple signalling models show that, although dividends are an expensive form of signal, dividends can still play a signalling role in an equilibrium where firms optimally choose among all possible signalling methods. John and Williams (1985), and Ofer and Thakor (1987) address how firms choose between stock repurchases and dividend payouts. The difference in the costs of these payout methods derives from the differential tax cost of dividends versus the dilution cost of repurchases. The optimal signal mix equates marginal costs across payout methods, so that the optimal dividend payout for any quality of firm decreases in the personal tax rate. Ambirush, John and Williams (1987), and Williams (1988) present models that combine the cash disbursement cost of Miller and Rock with John and Williams' relative cost structure for dividends and repurchases to determine the form of cash disbursement used. John and Mishra (1990), and John and Lang (1991) study how dividends, when combined with insider trades, can signal firm quality. Although the analysis does not rely on dividend taxes, an increase in the marginal cost of signalling with dividends (higher tax rates) alters the optimal mix of signals so that more signalling is done with insider trades and less with dividends. Kumar (1988) considers a model where due to differences in risk aversion between the informed manager and uninformed shareholders, there is a conflict of interest between these agents in determining the optimal investment level, and the resulting signalling equilibrium is partially separating. Finally, Bernheim (1991) develops a tax-based theory of dividend signals to determine the optimal mix of dividends and share repurchases when there is the potential for costly bankruptcy. Robertson (1992) documents exhaustively that the monotonicity property holds in each of these models.

In independent research, Bernheim and Wantz (1993) also make the observation concerning the relationship between tax regime and the 'bang-for-the-buck' in excess return associated with a given dividend level. Although the signalling theory just implies a monotonic relationship between dividend yield level and excess return, they regress excess return on changes in dividend yield level, including an interaction with the tax regime, and a vector of control variables. The coefficient on the interaction term is significant (t statistic around 2.5) --- changes in dividend yield level have a bigger effect on excess returns when dividend income is taxed less favorably. This is an interesting finding, but following our theoretical development, it provides neither evidence for nor against existing dividend signalling theories. Even were one to run a regression with dividend levels rather than changes, one would still have to interpret a positive coefficient on the interaction term cautiously. This is because the best linear fit to a non-monotonic relationship may have a significant sign on the interaction coefficient, even though a non-monotonic relationship would be inconsistent with signalling. Consequently, the appropriate test of dividend signalling theory is a test of monotonicity, and there exist a variety of robust non-parametric tests of monotonicity.

IV Tax Treatment of Dividends and Capital Gains, 1962-1988.

The Tax Reform Act of 1986, the U.S. Federal tax code applied the same personal income tax rate to long-term capital gains and ordinary income (including dividend income). This was the first time since 1921 that the income tax code was not discriminated against dividend income relative to capital gains income. While this change may have been the most dramatic change in the relative treatment of capital gains versus ordinary income, numerous changes in the tax code over the last three decades have also affected the relative tax treatment of capital gains and other income.

Prior to the Tax Reform Act, 50 percent or more of capital gains were excludable from taxable income, reducing the effective tax rate on capital gains below that on other forms of income. Over the period

1960-1986 numerous changes in income tax rates, tax brackets, exclusion allowances, changes in maximum alternative tax rates, changes in the definitions of long-term capital gains, and changes in deductability allowances for capital losses, have changed the effective tax disadvantage of dividends relative to capital gains. In 1988, the top tax rate on capital gains was 28 percent (the same as on all income). Throughout most of the 1960s this rate was 25 percent. In the mid-1970s rates rose dramatically for high income earners, so that in 1978 the Congressional Budget Office estimates an effective top tax rate of 25 percent on capital gains compared to 22 percent in the late 1970s and rates of 14 percent in the early 1980s. Table 1 lists the most significant changes over this period[11].

Figure 1 plots the maximum rates on dividend income and capital gains incomes over the period 1962-1988. This figure provides a graphical illustration of the ordering of tax regimes. We choose the maximum rates as our primary focus since as is well known, the primary recipients of capital gains incomes are concentrated in upper income earners[12]. For 1982 the behavior of the very top rate is a deceptive measure of the behavior of the tax treatment of dividends for high income earners because the tax reductions of 1981 *disproportionately* favor those with incomes over $215,400. The rankings were adjusted slightly to account for this.[13] Otherwise, the tax regimes are clearly ranked. It is important to note that in ranking years from "most favorable to dividends" to "least favorable to dividends" we have assumed that inflation affects dividend income and capital gains tax income equally. With a non-indexed Federal Tax code (as was the case through the high inflation period of the 1970s) bracket creep adversely affects the tax treatment of dividends. Since capital gains taxes are levied on nominal capital gains, inflation also adversely affects the tax treatment of capital gains. We implicitly assume that the effect

---

[11]All rates quoted are for married couples.

[12]Avery and Elliehausen [1986] estimate 85% of common stock is held by the top decile of the wealth distribution; 44% by the top one-half percentile.

[13]Alternative rankings of this tax regime, including dropping it from the analysis, did not affect the results.

of inflation on the *differential* tax treatment of dividends and capital gains is insignificant. The ranking of years by tax regime is presented in Table 2.

V. Description of the Non-parametric Tests.

The generic signalling model does not predict a particular functional relationship between firm type and signal strength (here, firm value and tax regime, holding dividend yield constant).[14] The sole prediction is that of a monotonic relationship between type and signal strength, so a direct non-parametric test of this monotonicity condition is appropriate.

Even ignoring functional form concerns, it is no easier to interpret the results from a time series regression. Using a different dummy variable for each tax regime, one is left with determining whether the signs on the dummies are consistent with signalling theory: one must determine whether the signs have the desired monotonic relationship. Thus, even after imposing functional form, one is still left with a test of rank.

An additional concern is that, to implement a test of the monotonicity condition, one must control for the effect of differences in dividend yields across firms since the theory predicts that expected return is an increasing function of dividend yield.[15] Methodologies that do not separate out portfolios according to their dividend yield implicitly impose a linear relationship between dividend yield and return. One could imagine forming portfolios sorted by dividend yield for each announcement date. But then the beta for each portfolio would have to be estimated since the resulting portfolios would be too small to be assumed well diversified. In contrast, the direct tests proposed here can easily control for dividend yield across large portfolios and allow for testing of the model without imposing unnecessary, and

----

[14]For instance, a simple linear regression is, in fact, an inappropriate way to test signalling theory, since it perforce fits a linear relationship on the data when the true relationship may be non-linear.

[15]Implicitly this suggests problems with interpretations of the standard event study which regresses excess return on dividend yield, but fails to control for tax regime. Systematic co-variation in dividend yield and tax regime over time can lead to spurious signalling findings.

potentially false, parametric restrictions.

Equally important, aggregation into large representative portfolios of firms means that our analysis should be robust to cross-sectional heterogeneity in both the dividend cost function, $C_i(D, \tau, \theta)$, and the informed's objective function, $V_i(P(D), \theta)$. Cross-sectional heterogeneity means that the amount of news revealed by a given dividend can vary across (unobserved) firm type. However, intertemporally, given dividend level portfolios should have the same cross-sectional distribution of firm types, so that, holding the tax regime constant, the excess return for a given dividend yield portfolio should be the same across years. The only key property that we require is that the cost and objective *functions* not change over time, so that we can isolate the tax regime effects.

We propose two non-parametric tests that may be used to test the signalling theory of dividends. Kendall's τ calculates an estimate of the correlation between the ranks of tax type and the level of the signalling variable (excess return). Kendall's distribution-free test for independence (K), is based on Kendall's τ. This test examines the hypothesis that X,Y variables of a bivariate population are independent. The test is designed to detect a class of alternatives associated with either positive or negative values of τ. The second test we examine is Heoffding's distribution-free test for independence (D). This test examines the hypothesis that X,Y variables of a bivariate population are independent. The test is designed to detect a much broader class of alternatives than Kendall's K and unlike K it is consistent when tau is zero and the null distribution is false. These tests can be found in Hollander and Wolfe (1973).

5.1 Kendall's Tau and Kendall's Distribution Free Test for Independence.

Kendall's tau is calculated as follows. Assume that the data consist of a bivariate random sample, (X,Y), of size n. Define as concordant two observation pairs if both members of one observation are larger than the respective members of the other observation pair, for example (1, 3), (2, 4). Let N_c denote the number of concordant pairs out of the total $\binom{n}{2} = n(n-1)/2$ possible pairs. Let N_d denote the remaining pairs (the number of discordant pairs, for example (4, 1) and

(2, 3)). Kendall's tau is calculated as $\tau = \frac{N_c - N_d}{n(n-1)/2}$.[16] The hypothesis that we desire to test is the one-tailed test for positive association since the monotonicity hypothesis implies a positive relationship between the relative tax disadvantage of dividends (our X series) and the excess return associated with any given level of dividend yield (our Y series). Thus, the null ($\tau = 0$) and (one-sided) alternative hypotheses to be tested can be written as:

H_0: X and Y are mutually independent:

$$P(X \leq a \text{ and } Y \leq b) = P(X \leq a)\ P(Y \leq b)\ \forall\ a,b.$$

H_1: Larger values of X tend to be paired with larger values of Y.

A test based simply on $N_c - N_d$, Kendall's distribution-free test for independence has wider usage. When H_0 is true, the large sample approximation of the Kendall K statistic, KL, has an asymptotic N(0,1) distribution.

5.2 Heoffding's Distribution-Free Test for Independence.

Heoffding' D statistic has a symmetric construction and tests against both alternatives where the X's and Y's are positively associated <u>and</u> alternatives where the X's and Y's are negatively associated. Unlike τ, D does not distinguish between positive and negative rank order correlation; only two-tailed tests are appropriate. Since we are concerned with examining the data in the context of a one-tailed test for positive rank order correlation we use this test primarily to verify the results based on Kendall's K. Because pooling equilibria result in a prediction of τ equal to zero, we use this test to check for alternatives of dependence when τ is equal to zero (Kendall's K is inconsistent if $\tau = 0$). See Hollander and Wolfe (1973) or Heoffding (1948) for more details. The large sample approximation to the D statistic is given by $nD + (1/36)$; p-values for this distribution are in Hollander and Wolfe.

---

[16] Thus, tau measures whether there are more discordant or concordant pairs. With two perfectly positively correlated variables, all pairs are concordant and $\tau = 1$. With two perfectly negatively correlated series, all pairs are discordant and $\tau = -1$. With two independent series, the expected number of concordant and discordant pairs is identical so that the expected value of τ is 0.

6 Data.

We test for the monotonic relationship between tax rates and excess returns predicted by dividend signalling models using American stock market information provided by the Center for Research in Stock Prices (CRSP). Information on stock returns, firm value, and cash disbursement distribution information, data for price, shares outstanding and all dividend distribution information was obtained from the CRSP monthly master file and the CRSP daily returns file was used to obtain stock return information. The period of analysis begins in July 1962 when daily returns were first collected.

For a firm to be considered part of the sample it had to meet the following criteria:

(1) The firm had to be listed on the New York Stock Exchange as of July 1962 or later.

(2) Firms are only considered over the period when they make regular quarterly cash dividends.[17]

(3) A complete set of price, dividend distribution and return information was available for the declaration date of the dividend.

For each event (declaration of a dividend) occurring at time t, we calculate both firm size and dividend yield.[18] In constructing the dividend yield and firm size variables, it is important to note that since the distribution information is provided by the monthly master file, the price and shares outstanding information are available only for the last trading day in each month. However, dividend declaration dates and the dollar amount of the dividends are available for the actual event date within an event month. To take account of the fact

---

[17]For the majority of firms in the sample their cash dividend disbursement pattern was a quarterly dividend. In addition year-end "extra" dividends for these firms were included.

[18]The dividend yield measure employed is a short-term measure (one-day return) that takes account of dividend size but not dividend timing. This biases it in favor of finding tax-related effects. For our purposes this is preferable to a long-term measure that is biased against finding tax effects (see Rumsey 1988 and Kalay and Michaely 1992 for a discussion of timing effects and measurement of tax effects of dividends).

that for any event occurring in month t, an investor's information set would only contain information known at t-1, firm size and dividend yield are computed using month t-1 price data. Thus, dividend yield is defined as the dollar amount of the cash dividend divided by the price in the month prior to the event month, and firm size is defined as the prior month price times the number of shares outstanding.

In investigating the monotonic relationship between tax regimes and excess return we separately control for any other variable that may impact on excess return. We therefore construct portfolios categorized by dividend yield and firm size. Our prediction is that *for any level of dividend yield* the information released should be more favorable the more disadvantageous the tax treatment of dividends. We therefore clearly need to control for variations in dividend yield. In addition, there are well-documented intertemporal variations in average dividend yields for which we wish to control. The importance of firm size is well documented in studies of asset pricing (see Keim 1985 or Bajaj and Vijh 1990).[19]

For each year, therefore, the total number of events for all firms in the sample were categorized based on the size of the calculated dividend yield. In preliminary work we experimented with various grid sizes.[20] All of our reported results are for the dividend yield groupings increasing in increments of 0.5%. Two considerations led us to choose this dividend yield grouping. First, we find the vast majority of quarterly dividend yields are below 2%. Second, in constructing excess returns for each dividend yield portfolio we follow the methodology of Brown and Warner (1980) and assume that each portfolio is sufficiently large that it is well diversified (so that its beta is constant over time). Equally important, to the extent that there is heterogeneity in the amount of public information about firms within a year, aggregation into large portfolios permits 'laws of large

----

[19]Also, it is possible that the information released may be a function of firm size (e.g. analysts reports may reveal more information about marge firms than small firms, requiring less information signalling for these firms).

[20]Dividend yield divisions in increments of 0.2% and 0.3% were examined. No substantive difference in results was found.

number' arguments that equal dividend yield portfolios are 'informationally' identical over time. Based on these considerations, we create four portfolios based on dividend yield size increments of 0.5%:

Portfolio A:	$0 < dy < 0.5$ %
Portfolio B:	$0.5 \leq dy < 1.0$ %
Portfolio C:	$1.0 \leq dy < 1.5$ %
Portfolio D:	$1.5 \leq dy \leq 2.0$ %

For all firms, the dividend distribution events for each year are then categorized based on these dividend increments. Each firm event is then assigned to one of the portfolios as above. We then calculate for each dividend yield portfolio (A through D) the excess returns for each year 1962 through 1988. For each event documented in each dividend yield portfolio, the associated daily stock return for that event date is recorded. The average daily return for each dividend yield portfolio is then calculated. This average daily return for the event date is converted into an annual return and the annual return on the market portfolio is then subtracted to give us the dividend yield portfolio's annual excess return. The market portfolio is characterized as the value-weighted portfolio of all stocks listed on the NYSE as supplied by CRSP. The non-parametric tests were run using these complete dividend yield portfolio's excess returns.

We also control for possible size effects by sub-dividing dividend yield portfolios A through D on the basis of firm size. This controls both for the standard size effects and for any systematic differences in public information across firms of different sizes (for instance, there may be less public information about small firms so that more information is revealed through dividends). For each year, the portfolios were ranked with respect to firm size and then split into three separate (but equal) size groups, large (L), medium (M) and small (S). For each of these new portfolios, the annual excess return was calculated in the same manner as with the undivided dividend yield portfolios. The non-parametric tests were run on these size-based dividend portfolios using the calculated annual excess returns.

Many studies have found a relationship between dividend yield and excess return and interpreted these results as evidence of tax effects resulting for the differential treatment of dividends and capital gains (e.g. Litzenberger and Ramaswamy 1979). These results are generally derived from estimates drawn from cross-sectional analysis which are aggregated over time. This time aggregation ignores the possibility that the yield-return relationship may vary through time due to changes in the tax regime. Since signalling theory predicts that the strength of the relationship between stock returns and dividend yields varies directly with the extent of the tax disadvantage of dividends, time aggregation across tax regimes can bias the results.

Recall too, that although this investigation is posed as a test of information-signalling models, that tax-based CAPM models provide the same predictions. Hence, a failure to to find the monotonic relationship provides strong evidence against the tax-effect theories which suggest that investors demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate them for the tax cost of dividends.

7. Results

Table 3 summarizes the results for the tests of the theory for the complete portfolios grouped by dividend yield and table 4 presents the corresponding results for the portfolios split by both dividend yield and firm size. Looking first at table 3, the monotonicity prediction that we wish to test implies that there should be concordance between the excess returns and the relative tax disadvantage of dividend income. We can see from the estimate for tau of rank correlation that there are as many discordant as concordant pairs and only for portfolio A may we reject the null hypothesis of independence based on Kendall's distribution free test statistic for independence (K). However, note that while we reject independence for portfolio A, the indication is that there is a *negative* association between tax regime and excess return for this portfolio, and *not* the *positive* correlation that signalling theory predicts! The large sample approximation of Kendall's K statistic, (KL), is also significant at the 5% level for this portfolio. Heoffding's distribution free test for independence (D) shows no significance for any dividend yield group and so according to this test, we cannot reject the null hypothesis of independence for any portfolio.

Table 4 presents results for the size-controlled dividend yield portfolios. Only for portfolio (A-S), the smallest firms paying dividends between 0% and 0.5%, do we reject independence based on Kendall's τ at the 5% level. For portfolios B-S (the smallest firms paying dividends between 0.5% and 1%) we reject independence at the 10% level. The result for A-S is consistent with table 3 where portfolio A was the only portfolio showing significant discordance. Recall that discordance is *not* consistent with the signalling hypothesis which is being tested. Again there are as many negative values as positive values for our estimates of τ. Heoffding's test statistic D, shows no significance for any of the portfolios in the table. Recall that unlike Kendall's K, Heoffding's D statistic is consistent when τ is zero and the null distribution is false. Examining table 4 it is clear that many of our estimates for τ are very close to zero. Based on these results, we cannot reject the null hypothesis of independence.[21]

We then test to see whether, even though dividends do not appear to be used as a signalling instrument, there is information contained in dividend yield. For each of the 27 years in our sample we determine whether higher dividend yield portfolios are associated with greater excess returns. We then aggregate, first across years, and again across tax regimes, to generate our test statistic, exploiting the fact that under the null hypothesis of independence, the variance of the sum of the yearly statistics is the sum of the variances. Aggregating across years we find a $\hat{\tau}$ = 4.18, so that we overwhelmingly reject the null hypothesis of independence at the .01% level, in favor of positive concordance: greater dividend yields are strongly associated with greater excess returns. Similarly, when we look at each of the 16 tax regimes, we obtain a $\hat{\tau}$ = 2.72, so that we again reject the null hypothesis in favor of positive concordance, this time at the .4% level.

In summary, we find very strong evidence that dividend yield levels do contain positive payoff relevant information, but we find no

----

[21]Values of Spellmans' rho for each portfolio were also calculated. These produced identical levels of significance to those presented for Kendall's tau and hence are omitted.

evidence of a positive association between excess return and the relative tax treatment of dividends.[22] That is, we find no evidence that dividends are used to signal information to investors. Heoffding's D does not reject independence of tax regime and excess returns for any of the portfolios considered and Kendall's τ finds evidence of significant concordance for none of the portfolios (while finding discordance for one aggregated and two dis-aggregated portfolios). (Statistically insignificant) discordance is as common as (statistically insignificant) concordance looking across the set of portfolio estimates.

8. Conclusions

Bagwell and Shoven (1989) estimate that $68 billion in cash dividends were paid in 1985. The issue addressing researchers is why, given the apparently high associated tax costs, are dividends issued? One can always postulate that, *ceteris paribus*, certain investors have a preference for dividends over an identical dollar amount of capital gains, but such relative preferences would have to be extreme in light of the tax costs. It is also hard to believe that dividends are really irrelevant given the enormous resources that firms devote to determining their dividend policy.

This paper tests whether the underlying explanation for dividends is signalling based. We show a common prediction of both scalar and vector dividend signalling models is that of a positive rank order correlation between the tax disadvantage of dividends relative to capital gains and the amount of "good" news revealed by any given *level* of dividend yield.

We present tests of this monotonicity prediction using data on dividend yield and excess returns for American stocks for the time period July 1962 to December 1988. Over this period we detail sixteen

----

[22]We also considered the possibility that the effective capital gains tax is zero due to dynamic tax shielding strategies. The results are almost identical to those presented here: portfolios A and A-S still have significant discordance at the 5% level and portfolio C-M has significant concordance at the 10% level. No other portfolios exhibited a significant monotonic relationship between excess returns and dividend tax rates.

distinct tax regimes, ordered according to the relative tax disadvantage of dividends. Constructing portfolios ranked by dividend yield and size we test whether there is any positive association between the relative tax disadvantage of dividends and the excess returns associated with any level of dividend payout. Using distribution-free tests for independence we find no evidence of the predicted positive monotonic relationship.

Our failure to reject the independence of tax regime and excess returns for any given dividend yield implies that we cannot find any evidence in support of the signalling theory of dividends. Because we test only the weakest signalling prediction -- monotonicity, we provide strong evidence that information content in dividends is independent of the marginal cost of using dividends as a signal. This result is inconsistent with any tax-based model of dividend signalling, and more generally with any dividend signalling model *provided* that the marginal investor is taxed. These findings are also inconsistent with tax-based CAPM arguments. Investors do not appear to demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate them for the tax cost of dividends.

It is important to stress that the paper does not find that *no* information is revealed through dividend announcements. Indeed, the evidence of positive excess returns associated with greater dividend announcements is indicative of information release. Rather, the results reveal that information release is an indirect by-product of dividend announcements, instead of a direct signalling goal. By looking at the relationship between excess returns and the marginal tax cost of dividends, we distinguish between classical Spencian signalling explanations and indirect information release. The information content of dividends is uncorrelated with the tax costs of dividends, and hence inconsistent with Spencian signalling.

FIGURE ONE

Maximum rates on capital gains and dividend income 1962-1988

□ = maximum rate on dividend (regular) income.

\+ = maximum rate on capital gains income.

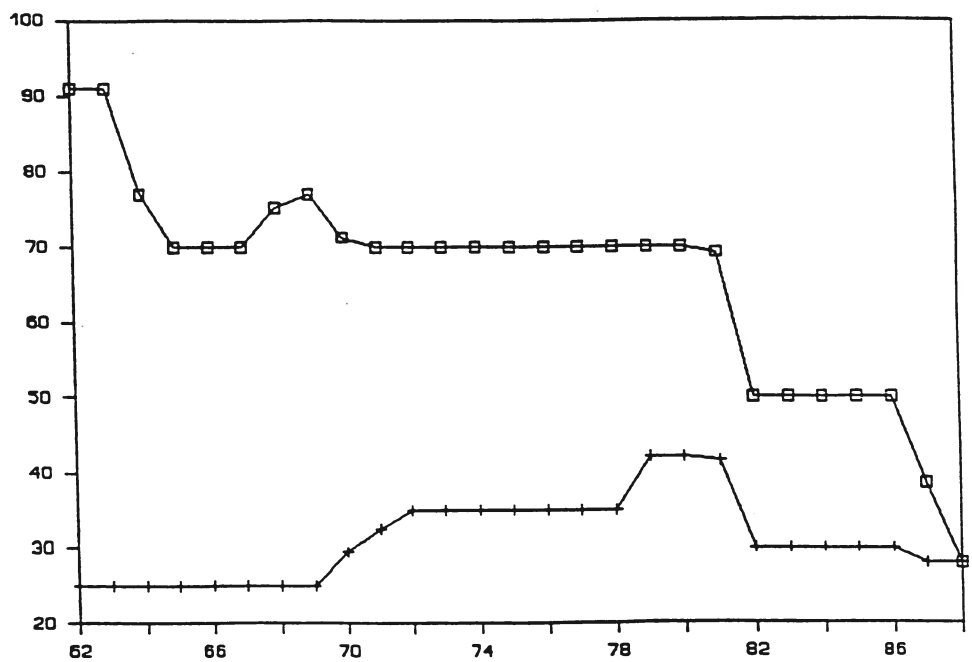


TABLE 1

Major tax changes affecting capital gains and dividend income.[a]
(Income tax rates quoted are for married couples)

Year(s)	Income Tax Changes.	Inclusion Rate	Alternate Maximum Rate	Holding Period *(mths)*
1962/63	Top rate = 90%	50%	25%	6
1964	Rates lowered, top rate 77%.	50%	25%	6
1965/66/67	Rates lowered, top rate 70%	50%	25%	6
1968	Tax surcharge of 7.5%	50%	25%	6
1969	Tax surcharge of 10%.	50%	25%	6
1970	Tax surcharge of 2.5%	50%	29.5%	6
1971	No changes	50%	32.5% [b]	6
1972/76.	No major changes	50%	None	9
1977/78.	No major changes	50%	None	12
1979/80	No major changes	60%	None	12
1981	Rates lowered 5%	60%	None	12
1982	Top rates cut to 50% from 69%. Other rates lowered by 10%.	60%	None	12
1983	Rates lowered 10%	60%	None	12
1984	Rates lowered	60%	None	6
1985/86.	No major changes, some bracket adjustments.	60%	None	6
1987	Lower rates, reduced number of brackets.	100%	28%	6
1988	Lower rates.	100%	None	6

a. Income tax data from tables A-3 to A-6 of Pechman (1987), updated to 1988 from *Standard Federal Tax Reports, Commerce Clearing House Inc, (1988).* Capital gains tax treatment obtained from *Standard Federal Tax Reports, Commerce Clearing House Inc. (1962-1988).*

b With $50,000 cap on alternate maximum rate shield. All capital gains over $50,000 per individual taxed at 50% of regular marginal rate.

TABLE 2

Rankings from most favorable to dividends R_1 to least favorable R_{16}[a].

R_1	1988
R_2	1987
R_3	1985-86
R_4	1984
R_5	1983
R_6	1982
R_7	1972-78
R_8	1979-80
R_9	1981
R_{10}	1965-67
R_{11}	1971
R_{12}	1970
R_{13}	1969
R_{14}	1968
R_{15}	1964
R_{16}	1962-63

a. R_j denotes the jth most favorable difference in tax rates between dividends and capital gains (for example in 1988 this difference is zero, in 1962-1963 for the highest tax bracket, this difference is 65% (90% on dividend income versus 25% on capital gains)). See figure 1.

Table 3

Non-Parametric Test Statistics for the Undivided Sample

X = TAX REGIME RANKING

Y = EXCESS RETURN

	n	$\hat{\tau}$	K	KL	D	nD+(1/36)
A	16	-0.4167 †	-50	-2.25 †	0.004	0.0981
B	16	-0.1833	-22	-0.99	0.001	0.0481
C	16	0.133	16	0.72	-0.001	0.0097
D	16	0.1833	22	0.99	-0.001	0.143

† denotes significance at the α = 0.05 level.

Key for Tables 3 and 4.

A - D represents portfolios with lowest (A) to highest (D) dividend yield.

S, M, L denote firm sizes small, medium, large respectively.

n is the number of tax regimes over the period.

$\hat{\tau}$ is Kendall's tau statistic.

K is Kendall's K statistic.

KL is the large sample approximation to Kendall's K.

D is Heoffding's D statistic.

nD + 1/36 is the large sample approximation to D.

<u>**Table 4**</u>

Non-Parametric Test Statistic for Portfolios
Divided by Firm Size

X = TAX REGIME RANKING
Y = EXCESS RETURN

	n	$\hat{\tau}$	K	KL	D	nD+(1/36)
A-L	16	-0.2167	-26	-1.1706	0.0007	0.0397
A-M	16	-0.0833	-10	-0.4502	0.0008	0.0417
A-S	16	-0.417 †	-50	-2.251 †	0.0034	0.0818
B-L	16	-0.0667	-8	-0.3602	0.0023	0.065
B-M	16	0.0	0	0.0	-0.0005	0.0203
B-S	16	-0.2667 ★	-32	-1.4407 ★	0.0001	0.0278
C-L	16	-0.0167	-2	-0.0900	-0.0006	0.0187
C-M	16	0.2	24	1.0805	-0.0003	0.0233
C-S	16	0.0667	8	0.3602	-0.0010	0.0113
D-L	16	0.1	12	0.5403	0.0002	0.0305
D-M	16	0.1167	20	0.9005	0.0010	0.0430
D-S	16	0.2	24	1.0805	0.0012	0.0463

† denotes significance at the α = 0.05 level.

★ denotes significance at the α = 0.1 level.

Bibliography

Ahrony, J., and I. Swary, (1980) "Quarterly Dividend and Earnings Announcements and Stockholder Returns: An Empirical Analysis." *Journal of Finance* 35: 1-12.

Ambirish, R., K. John, and J. Williams, (1987) "Efficient Signalling with Dividends and Investments." *Journal of Finance* 42: 321-343.

Ang, J.S., (1975) "Dividend Policy: Informational Content or Partial Adjustment?" *Review of Economics and Statistics* 57: 65-70.

Asquith, P., and D. Mullins, (1983) "The Impact of Initiating Dividend Payments on Shareholders' Wealth" *Journal of Business* 56: 77-96.

Auerbach, A.J., (1988) "Capital Gains Taxation in the US: Realizations, Revenue and Rhetoric. *Brookings Papers on Economic Activity 2.*

Avery, R. and G. Elliehausen, (1986)"Financial Characteristic of High Income Families", *Federal Reserve Bulletin* 72: 172-182.

Bagwell, L., and J. Shoven, (1989) "Cash Distributions to Shareholders." *Journal of Economic Perspectives*: 129-140.

Bajaj, M., and A. Vijh, (1990) "Dividend Clienteles and the Information Content of Dividend Changes." *Journal of Financial Economics* 26: 193-219.

Banks, J., and J. Sobel, (1987) "Equilibrium Selection in Signalling Games." *Econometrica*: 647-661.

Bar-Yosef, S., and L. Huffman, (1986) "The Information Content of Dividends." *Journal of Financial and Quantitative Analysis* 21: 47-58.

Bernheim, B.D., (1991) "Tax Policy and the Dividend Puzzle." *RAND Journal of Economics* 22: 455-476.

Bernheim, B.D. and A. Wantz, (1993) "A Tax-Based Test of the Dividend Signaling Hypothesis", *mimeo*, Princeton University.

Bhattacharya, S. (1979) "Imperfect Information, Dividend Policy and the 'Bird in the Hand' Fallacy" *Bell Journal of Economics* 10: 259-270.

Bhattacharya, S. (1980) "Nondissipative Signalling Structures and Dividend Policy." *Quarterly Journal of Economics* 95: 1-24.

Brickley, J., (1983) "Shareholder Wealth, Informational Signalling and the Specially Designated Dividend: An Empirical Study." *Journal of Financial Economics* 12: 187-210.

Brennan, M. (1970) "Taxes, Market Valuation, and Corporate Financial Policy." *National Tax Journal* 23: 417-427.

Brown, S.J., and J.B. Warner, (1980) "Measuring Security Price Performance." *Journal of Financial Economics* 8: 205-258.

Charest, G., (1978) "Dividend Information, Stock Returns and Market Efficiency - II." *Journal of Financial Economics 6: 297-330.*

Cho, I-K., and D. Kreps, (1987) "Signalling Games and Stable Equilibria." *Quarterly Journal of Economics* 102: 179-221.

Christie, W.G., and R. Huang, (1992) "The Changing Functional Relation Between Stock Returns and Dividend Yields." *mimeo* , Vanderbilt

University.

Commerce Clearing House Inc, (1962-88). **Standard Federal Tax Reports.**

Eades, K., (1982) "Empirical Evidence on Dividends as a Signal of Firm Quality." *Journal of Financial and Quantitative Analysis* 17: 471-500.

Engers, M, (1987) "Signaling with Many Signals" *Econometrica* 55 663-674.

Fama, E., Fisher, F., Jensen, M. and R. Roll, (1969)

Ghosh, C., and J.R. Wooldridge, (1988) "An Analysis of Shareholder Reaction to Dividend Cuts and Omissions." *Journal of Financial Research* 9: 281-294.

Gonedes, N.J., (1978) "Corporate Signalling, External Accounting and Capital Market Equilibrium: Evidence on Dividends, Income, and Extraordinary Items." *Journal of Accounting Research* 16: 26-79.

Hakansson, N., (1982) "To Pay or Not to Pay Dividends." *Journal of Finance* 37: 415-428.

Hollander, M., and D. Wolfe, (1973) *Nonparametric Statistical Methods* (Wiley).

John, K., and B. Mishra, (1990) "Information Content of Insider Trading Around Corporate Announcements: The Case of Capital Expenditures." *Journal of Finance* 45: 835-855.

John, K., and L. Lang, (1991) "Insider Trading around Dividend Announcements: Theory and Evidence." *Journal of Finance* 46: 1361-1389.

John, K., and J. Williams, (1985) "Dividends, Dilution and Taxes: A Signalling Equilibrium." *Journal of Finance* 40: 1053-1070.

Kalay, A. (1980) "Signalling, Informational Content and the Reluctance to Cut Dividends." *Journal of Financial and Quantitative Analysis* 15: 855-865.

Kalay, A. and U. Lowenstein, (1986) "the Informational Content of the Timing of Dividend Announcements." *Journal of Financial Economics* 16: 373-388.

Kalay, A., and R. Michaely (1992) "Dividends and Taxes: A Re-examination." *mimeo*, Cornell University.

Keim, D., (1983) "Size Related Anomalies and Stock Return Seasonality." *Journal of Financial Economics* 12: 13-32.

Kumar, P., (1988) "Shareholder-Manager Conflict and the Information Content of Dividends." *Review of Financial Studies* 1: 111-136.

Laub, P.M., (1976) "On the Informational Content of Dividends." *Journal of Business* 49: 73-80.

Litzenberger, R., and K. Ramaswamy, (1979) "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence." *Journal of Financial Economics* 7: 163-195.

Litzenberger, R., and K. Ramaswamy, (1980) "Dividends, Short Selling Restrictions Tax Induced Investor Clienteles and Market Equilibrium." *Journal of Finance* 35: 469-482.

Litzenberger, R., and K. Ramaswamy, (1982) "The Effect of Dividends on

Common Stock Prices: Tax Effects or Information Effects." *Journal of Finance* 37:429-444.

Miller, M. and F. Modigliani, (1961) "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business* 4: 411-433.

Miller, M., and K. Rock, (1985) "Dividend Policy Under Asymmetric Information." *Journal of Finance* 40: 1031-1051.

Miller, M., and M. Scholes, (1978) "Dividends and Taxes" *Journal of Financial Economics* 6: 333-364.

Modigliani, F. and M. Miller, (1958) "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review* 48: 261-297.

Ofer, A., and A. Thakor, (1987) "A Theory of Stock Price Responses to Alternative Corporate Cash Disbursement Methods: Stock Repurchases and Dividends." *Journal of Finance* 42: 365-394.

Patell, J., and H. Wolfson, (1984) "The Intra-day Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements." *Journal of Financial Economics* 13: 223-252.

Pechman, J., (1987) **Federal Tax Policy** (5th Ed.) Brookings Institution.

Penman, S. (1983) "The Predictive Content of Earnings Forecasts and Dividends." *Journal of Finance* 38: 1181-1199.

Peterson, P., D. Peterson and J. Ang (1985) "Direct Evidence on the Marginal Rate of Taxation on Dividend Income" *Journal of Financial Economics* 14: 267-282.

Pettit, R., (1972) "Dividend Announcements, Security Performance, and Capital Market Efficiency." *Journal of Finance* 27: 993-1001.

Pettit, R., (1976) "The Impact of Dividend and Earnings Announcements: A Reconciliation." *Journal of Business* 49: 86-96.

Riley, J, (1979) "Informational Equilibrium." *Econometrica* 47: 331-358.

Robertson, F., (1992) "Three Essays in Financial Economics." *Ph.D. thesis*, Queen's University.

Ross, S., (1977) "The Determination of Financial Structure: The Incentive-Signalling Approach." *Bell Journal of Economics*: 23-40.

Rumsey, J., (1989) "Testing for Tax Effects of Dividend Yields on Pre-Tax Returns." *mimeo*, York University.

Spence, M., (1973) "Job Market Signalling." *Quarterly Journal of Economics* 87: 355-374.

Talmor, E., (1981) "Asymmetric Information, Signalling, and Optimal Corporate Financial Decisions." *Journal of Financial and Quantitative Analysis* 16: 413-435.

Watts, R., (1973) "The Informational Content of Dividends." *Journal of Business* 46: 191-211.

Watts, R., (1978) Systematic 'Abnormal' Returns after Quarterly Earning Announcements." *Journal of Financial Economics* 6: 127-150.

Williams, J., (1988) "Efficient Signalling with Dividends, Investment, and Stock Repurchases." *Journal of Finance* 43: 737-747.

Discussion Paper #895

Testing Dividend Signalling Models

by

Dan Bernhardt
Queen's University

J. Fiona Robertson
Seattle University

Ray Farrow
Seattle University

January 1994

Common Stock Prices: Tax Effects or Information Effects." *Journal of Finance* 37:429-444.

Miller, M. and F. Modigliani, (1961) "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business* 4: 411-433.

Miller, M., and K. Rock, (1985) "Dividend Policy Under Asymmetric Information." *Journal of Finance* 40: 1031-1051.

Miller, M., and M. Scholes, (1978) "Dividends and Taxes" *Journal of Financial Economics* 6: 333-364.

Modigliani, F. and M. Miller, (1958) "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review* 48: 261-297.

Ofer, A., and A. Thakor, (1987) "A Theory of Stock Price Responses to Alternative Corporate Cash Disbursement Methods: Stock Repurchases and Dividends." *Journal of Finance* 42: 365-394.

Patell, J., and H. Wolfson, (1984) "The Intra-day Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements." *Journal of Financial Economics* 13: 223-252.

Pechman, J., (1987) **Federal Tax Policy** (5th Ed.) Brookings Institution.

Penman, S. (1983) "The Predictive Content of Earnings Forecasts and Dividends." *Journal of Finance* 38: 1181-1199.

Peterson, P., D. Peterson and J. Ang (1985) "Direct Evidence on the Marginal Rate of Taxation on Dividend Income" *Journal of Financial Economics* 14: 267-282.

Pettit, R., (1972) "Dividend Announcements, Security Performance, and Capital Market Efficiency." *Journal of Finance* 27: 993-1001.

Pettit, R., (1976) "The Impact of Dividend and Earnings Announcements: A Reconciliation." *Journal of Business* 49: 86-96.

Riley, J, (1979) "Informational Equilibrium." *Econometrica* 47: 331-358.

Robertson, F., (1992) "Three Essays in Financial Economics." *Ph.D. thesis*, Queen's University.

Ross, S., (1977) "The Determination of Financial Structure: The Incentive-Signalling Approach." *Bell Journal of Economics*: 23-40.

Rumsey, J., (1989) "Testing for Tax Effects of Dividend Yields on Pre-Tax Returns." *mimeo*, York University.

Spence, M., (1973) "Job Market Signalling." *Quarterly Journal of Economics* 87: 355-374.

Talmor, E., (1981) "Asymmetric Information, Signalling, and Optimal Corporate Financial Decisions." *Journal of Financial and Quantitative Analysis* 16: 413-435.

Watts, R., (1973) "The Informational Content of Dividends." *Journal of Business* 46: 191-211.

Watts, R., (1978) Systematic 'Abnormal' Returns after Quarterly Earning Announcements." *Journal of Financial Economics* 6: 127-150.

Williams, J., (1988) "Efficient Signalling with Dividends, Investment, and Stock Repurchases." *Journal of Finance* 43: 737-747.

Testing Dividend Signalling Models.

Dan Bernhardt[1], J. Fiona Robertson[2] and Ray Farrow

January, 1994.

[1] Department of Economics
Queen's University
Kingston, Ontario
Canada, K7L 3N6
bernhard@qucdn.queensu.ca

[2] Department of Economics and Finance.
Albers School of Business and Economics
Seattle University
Seattle, Washington.
fiona@seattleu.edu

Abstract

This paper derives a key monotonicity property common to dividend signalling models: the greater the rate that dividend income is taxed relative to capital gains income, the greater the value of information revealed by a given dividend yield, and hence the greater the associated excess return. This monotonicity condition allows us to distinguish the hypothesis that dividends are used as a signalling device from the hypothesis that dividends contain information but are not used as Spencian signals. The monotonicity conditions are tested with robust non-parametric techniques. Although we find strong evidence that dividend announcements contain information, we find no evidence to support dividend signalling. The same results are inconsistent with tax-based CAPM arguments.

Keywords: dividends, signalling, non-parametric.

JEL Classification: G0

We appreciate the comments and suggestions of Bart Lipman and Jeff Strnad. Seminar participants at Caltech and Queen's University provided useful feedback. The first author acknowledges support from the SSHRC. All errors remain our own.

University.

Commerce Clearing House Inc, (1962-88). **Standard Federal Tax Reports.**

Eades, K., (1982) "Empirical Evidence on Dividends as a Signal of Firm Quality." *Journal of Financial and Quantitative Analysis* 17: 471-500.

Engers, M, (1987) "Signaling with Many Signals" *Econometrica* 55 663-674.

Fama, E., Fisher, F., Jensen, M. and R. Roll, (1969)

Ghosh, C., and J.R. Wooldridge, (1988) "An Analysis of Shareholder Reaction to Dividend Cuts and Omissions." *Journal of Financial Research* 9: 281-294.

Gonedes, N.J., (1978) "Corporate Signalling, External Accounting and Capital Market Equilibrium: Evidence on Dividends, Income, and Extraordinary Items." *Journal of Accounting Research* 16: 26-79.

Hakansson, N., (1982) "To Pay or Not to Pay Dividends." *Journal of Finance* 37: 415-428.

Hollander, M., and D. Wolfe, (1973) *Nonparametric Statistical Methods* (Wiley).

John, K., and B. Mishra, (1990) "Information Content of Insider Trading Around Corporate Announcements: The Case of Capital Expenditures." *Journal of Finance* 45: 835-855.

John, K., and L. Lang, (1991) "Insider Trading around Dividend Announcements: Theory and Evidence." *Journal of Finance* 46: 1361-1389.

John, K., and J. Williams, (1985) "Dividends, Dilution and Taxes: A Signalling Equilibrium." *Journal of Finance* 40: 1053-1070.

Kalay, A. (1980) "Signalling, Informational Content and the Reluctance to Cut Dividends." *Journal of Financial and Quantitative Analysis* 15: 855-865.

Kalay, A. and U. Lowenstein, (1986) "the Informational Content of the Timing of Dividend Announcements." *Journal of Financial Economics* 16: 373-388.

Kalay, A., and R. Michaely (1992) "Dividends and Taxes: A Re-examination." *mimeo*, Cornell University.

Keim, D., (1983) "Size Related Anomalies and Stock Return Seasonality." *Journal of Financial Economics* 12: 13-32.

Kumar, P., (1988) "Shareholder-Manager Conflict and the Information Content of Dividends." *Review of Financial Studies* 1: 111-136.

Laub, P.M., (1976) "On the Informational Content of Dividends." *Journal of Business* 49: 73-80.

Litzenberger, R., and K. Ramaswamy, (1979) "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence." *Journal of Financial Economics* 7: 163-195.

Litzenberger, R., and K. Ramaswamy, (1980) "Dividends, Short Selling Restrictions Tax Induced Investor Clienteles and Market Equilibrium." *Journal of Finance* 35: 469-482.

Litzenberger, R., and K. Ramaswamy, (1982) "The Effect of Dividends on

Bibliography

Ahrony, J., and I. Swary, (1980) "Quarterly Dividend and Earnings Announcements and Stockholder Returns: An Empirical Analysis." *Journal of Finance* 35: 1-12.

Ambirish, R., K. John, and J. Williams, (1987) "Efficient Signalling with Dividends and Investments." *Journal of Finance* 42: 321-343.

Ang, J.S., (1975) "Dividend Policy: Informational Content or Partial Adjustment?" *Review of Economics and Statistics* 57: 65-70.

Asquith, P., and D. Mullins, (1983) "The Impact of Initiating Dividend Payments on Shareholders' Wealth" *Journal of Business* 56: 77-96.

Auerbach, A.J., (1988) "Capital Gains Taxation in the US: Realizations, Revenue and Rhetoric. *Brookings Papers on Economic Activity 2.*

Avery, R. and G. Elliehausen, (1986)"Financial Characteristic of High Income Families", *Federal Reserve Bulletin* 72: 172-182.

Bagwell, L., and J. Shoven, (1989) "Cash Distributions to Shareholders." *Journal of Economic Perspectives*: 129-140.

Bajaj, M., and A. Vijh, (1990) "Dividend Clienteles and the Information Content of Dividend Changes." *Journal of Financial Economics* 26: 193-219.

Banks, J., and J. Sobel, (1987) "Equilibrium Selection in Signalling Games." *Econometrica*: 647-661.

Bar-Yosef, S., and L. Huffman, (1986) "The Information Content of Dividends." *Journal of Financial and Quantitative Analysis* 21: 47-58.

Bernheim, B.D., (1991) "Tax Policy and the Dividend Puzzle." *RAND Journal of Economics* 22: 455-476.

Bernheim, B.D. and A. Wantz, (1993) "A Tax-Based Test of the Dividend Signaling Hypothesis", *mimeo*, Princeton University.

Bhattacharya, S. (1979) "Imperfect Information, Dividend Policy and the 'Bird in the Hand' Fallacy" *Bell Journal of Economics* 10: 259-270.

Bhattacharya, S. (1980) "Nondissipative Signalling Structures and Dividend Policy." *Quarterly Journal of Economics* 95: 1-24.

Brickley, J., (1983) "Shareholder Wealth, Informational Signalling and the Specially Designated Dividend: An Empirical Study." *Journal of Financial Economics* 12: 187-210.

Brennan, M. (1970) "Taxes, Market Valuation, and Corporate Financial Policy." *National Tax Journal* 23: 417-427.

Brown, S.J., and J.B. Warner, (1980) "Measuring Security Price Performance." *Journal of Financial Economics* 8: 205-258.

Charest, G., (1978) "Dividend Information, Stock Returns and Market Efficiency - II." *Journal of Financial Economics 6: 297-330.*

Cho, I-K., and D. Kreps, (1987) "Signalling Games and Stable Equilibria." *Quarterly Journal of Economics* 102: 179-221.

Christie, W.G., and R. Huang, (1992) "The Changing Functional Relation Between Stock Returns and Dividend Yields." *mimeo* , Vanderbilt

I Introduction.

Since Ross (1977) and Bhattacharya (1979), financial economists have explored the possible signalling properties of dividends and other financial activities. Research into the signalling properties of dividends has been motivated by an attempt to explain the apparent excess returns observed following announcements by firms of favorable dividends. Building on the work of Spence (1973), Bhattacharya (1979) produced an internally consistent model of Modigliani and Miller's (1963) "informational content of dividends hypothesis", demonstrating how dividends could allow insiders to credibly communicate information about the expected future value of the firm to less informed outsiders.[1] Credibility of the signal requires that it not pay low quality firms to mimic the behavior of high quality firms.

The majority of the empirical studies of the "information content of dividends hypothesis" have used an event study methodology to investigate the response of share prices to the announcement of changes in dividend levels.[2] Many of these papers found evidence supportive of the view that there appears to be a stock price response to changes in firms' dividend policies (Aharony and Swary 1980, Asquith and Mullins 1983, Brickley 1983, Charest 1978, Fama, Fischer, Jensen and Roll 1969, Ghosh and Woolridge 1988, Kalay 1980, Kalay and Lowenstein 1986, Laub 1976, Patell and Wolfson 1984, Pettit 1972, 1976). Such results have been viewed as supportive of the view that dividend announcements are interpreted by the market as being informative of firm value. A few early studies focused on examining the predictive content of dividend policy for future share performance, rather than investigating the existence of an announcement effect, and found less evidence to support the view that dividend announcements convey additional information to the market (Ang 1975, Gonedes 1978, Penman 1983, Watts 1973, 1978).

----

[1] Credibility of the signals here refers to the recognition that the signalling aspect of financial policy must be immune to the possibility that insiders could strategically manipulate the signals sent in such a way as to allow them to benefit from temporary mispricing of the firm's shares.

[2] Notable exceptions are Kalay (1982), John and Mishra (1990), and John and Lang (1991), which are discussed below.

Eades (1982) takes a different approach to testing dividend signalling models. Rather than trying to detect the presence of an announcement effect from dividend changes or examine the predictive content of dividend policy, Eades derives and tests specific predictions obtained from comparative static analysis of a formal dividend signalling model.[3] Eades analyzes a version of Bhattacharya's signalling model and derives two testable hypotheses: (1) an inverse relationship between dividend yield and a firm's own variance of returns; and (2) the "relative signalling strength hypothesis" (RSS) which states that "higher risk firms exhibit stronger changes in value relative to their lower risk counterparts for any given change in dividends" (p. 473). Eades finds empirical support for the first of these results but strongly rejects the RSS hypothesis.

This paper's test of signalling theories of dividends is in the spirit of Eades in that we derive and test a comparative static result derived from dividend signalling models. The RSS hypothesis tested by Eades was derived from a particular model of dividend signalling with particular specifications of functional forms and distributional assumptions within that model. In contrast we test a non-parametric comparative static result that holds for a class of dividend signalling models that encompasses virtually every published dividend signalling model, rather than a comparative static derived from a particular example of such a model.

We argue that a testable hypothesis that can be derived within the context of almost any Spencian signalling theory of dividends is that there is a monotonic relationship between the marginal tax rate on dividend income relative to capital gains and the amount of "good news" revealed by any sized dividend yield *level*. The higher is the relative tax rate on dividends, the better is the "type" revealed by any level of dividend yield, and hence, under rational expectations, the greater

---

[3] John and Mishra also derive comparative statics from an explicit signalling model. They allude to empirical evidence from other researchers consistent with their predictions. John and Lang (1991) derive testable implications from a specific signalling model regarding correlations between the announcement of dividend changes, the extent of insider trade and the response of stock prices. They find weak empirical supporting results.

Table 4

Non-Parametric Test Statistic for Portfolios
Divided by Firm Size

X = TAX REGIME RANKING
Y = EXCESS RETURN

	n	$\hat{\tau}$	K	KL	D	nD+(1/36)
A-L	16	-0.2167	-26	-1.1706	0.0007	0.0397
A-M	16	-0.0833	-10	-0.4502	0.0008	0.0417
A-S	16	-0.417 †	-50	-2.251 †	0.0034	0.0818
B-L	16	-0.0667	-8	-0.3602	0.0023	0.065
B-M	16	0.0	0	0.0	-0.0005	0.0203
B-S	16	-0.2667 *	-32	-1.4407 *	0.0001	0.0278
C-L	16	-0.0167	-2	-0.0900	-0.0006	0.0187
C-M	16	0.2	24	1.0805	-0.0003	0.0233
C-S	16	0.0667	8	0.3602	-0.0010	0.0113
D-L	16	0.1	12	0.5403	0.0002	0.0305
D-M	16	0.1167	20	0.9005	0.0010	0.0430
D-S	16	0.2	24	1.0805	0.0012	0.0463

† denotes significance at the α = 0.05 level.
* denotes significance at the α = 0.1 level.

Table 3

Non-Parametric Test Statistics for the Undivided Sample

X = TAX REGIME RANKING
Y = EXCESS RETURN

	n	$\hat{\tau}$	K	KL	D	nD+(1/36)
A	16	-0.4167 †	-50	-2.25 †	0.004	0.0981
B	16	-0.1833	-22	-0.99	0.001	0.0481
C	16	0.133	16	0.72	-0.001	0.0097
D	16	0.1833	22	0.99	-0.001	0.143

† denotes significance at the $\alpha = 0.05$ level.

Key for Tables 3 and 4.

A - D represents portfolios with lowest (A) to highest (D) dividend yield.
S, M, L denote firm sizes small, medium, large respectively.
n is the number of tax regimes over the period.
$\hat{\tau}$ is Kendall's tau statistic.
K is Kendall's K statistic.
KL is the large sample approximation to Kendall's K.
D is Heoffding's D statistic.
nD + 1/36 is the large sample approximation to D.

the associated excess return. This result holds *even* in signalling models in which the underlying signalling argument is *not* tax based. It simply reflects the fact that the higher the marginal tax rate on income, the more costly the dividend signal.

This observation suggests a simple and robust method for testing signalling theories of dividends. Over the period 1962-1988 there have been numerous changes in the US Federal tax code governing the taxation of dividend income and capital gains in the US. We identify 16 distinct tax regimes and order them from most favorable to dividend income relative to capital gains income, to least favorable. Signalling theories predict that for a *given* dividend yield *level*, the good news released, and hence the associated excess return, should be least when dividend income receives the most favorable tax treatment, and the excess return should be greatest when dividend income receives the least favorable tax treatment.[4] We refer to this as the monotonicity property of signalling models. It is important to note that we can distinguish between whether information is released as an indirect by-product of dividend announcements (e.g. higher dividends reveal more cash on hand) and whether information is released as a result of Spencian signalling. The former hypothesis predicts that information release should be positive, but independent of the tax regime.

Since a generic dividend signalling model does not predict a particular functional form for the relationship between tax rates and excess returns for a given dividend yield, we test this monotonicity relationship without imposing particular parametric restrictions. In particular, we employ non-parametric tests of rank order correlation (Kendall's tau and Heoffding's distribution free tests) to test for the predicted monotonic relationship between tax regime and excess returns for large portfolios of firms with similar dividend yields. There are several advantages of this approach. The non-parametric tests are robust, allow us to control for the dividend yield levels, and we do

---

[4]Note that since the tax signalling cost is a *static* function of the *level* and *not change* in the dividend yield, theory predicts no particular relationship between excess return and *change* in dividend yield. The theoretical development in Section 2 makes this clear.

not have to worry about mis-specifying the functional form. Further, by aggregating into large representative portfolios our analysis is robust to cross-sectional heterogeneity in firms (e.g. in payoff functions, information).

We find strong evidence that dividend announcements contain payoff relevant information: dividend yield is highly positively correlated with the excess returns earned following the announcements. However, we find *no* evidence of correlation between the relative tax treatment of dividends and the excess returns earned following dividend announcements (controlling for the size of the dividend payout and firm size). We thus find no evidence to support the view that dividends are used to signal firm value. More precisely, the paper provides strong evidence against the joint hypothesis that the marginal investor is taxed and that dividend yield serves as a Spencian signal of firm value. Peterson, Peterson and Ang (1985) examine income tax returns and estimate a marginal effective tax rate on dividend income of 30% for 1979, suggesting that the marginal investor is taxed.

This evidence is also inconsistent with the tax-based CAPM models of Brennan (1979) and Litzenberger and Ramaswamy (1979, 1980, 1983). There, investors demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate for the higher tax cost of dividends relative to capital gains; higher dividend taxes raise the required pre-tax return.[5]

Section 2 derives in a general context the monotonicity property that is the basis for our test. Section 3 examines the theoretic literature on dividend signalling, identifying that the monotonicity property holds in *each* paper. Section 4 discusses the various tax code changes that have occurred in the treatment of capital gains taxes in the U.S. over the period 1960-1988. Section 5 details the non-parametric tests employed and explains the ways in which the analysis is robust. Section 6 describes the data. Section 7 details our findings. Section 8 discusses the results and draws conclusions.

----

[5] A recent study by Christie and Huang (1992) focuses on the tax effect of dividends across tax regimes.

TABLE 2

Rankings from most favorable to dividends R_1 to least favorable R_{16}.[a]

R_1	1988
R_2	1987
R_3	1985-86
R_4	1984
R_5	1983
R_6	1982
R_7	1972-78
R_8	1979-80
R_9	1981
R_{10}	1965-67
R_{11}	1971
R_{12}	1970
R_{13}	1969
R_{14}	1968
R_{15}	1964
R_{16}	1962-63

a. R_j denotes the jth most favorable difference in tax rates between dividends and capital gains (for example in 1988 this difference is zero, in 1962-1963 for the highest tax bracket, this difference is 65% (90% on dividend income versus 25% on capital gains)). See figure 1.

TABLE 1

Major tax changes affecting capital gains and dividend income.[a]
(Income tax rates quoted are for married couples)

Year(s)	Income Tax Changes.	Inclusion Rate	Alternate Maximum Rate	Holding Period *(mths)*
1962/63	Top rate = 90%	50%	25%	6
1964	Rates lowered, top rate 77%.	50%	25%	6
1965/66/67	Rates lowered, top rate 70%	50%	25%	6
1968	Tax surcharge of 7.5%	50%	25%	6
1969	Tax surcharge of 10%.	50%	25%	6
1970	Tax surcharge of 2.5%	50%	29.5%	6
1971	No changes	50%	32.5% [b]	6
1972/76.	No major changes	50%	None	9
1977/78.	No major changes	50%	None	12
1979/80	No major changes	60%	None	12
1981	Rates lowered 5%	60%	None	12
1982	Top rates cut to 50% from 69%. Other rates lowered by 10%.	60%	None	12
1983	Rates lowered 10%	60%	None	12
1984	Rates lowered	60%	None	6
1985/86.	No major changes, some bracket adjustments.	60%	None	6
1987	Lower rates, reduced number of brackets.	100%	28%	6
1988	Lower rates.	100%	None	6

a. Income tax data from tables A-3 to A-6 of Pechman (1987), updated to 1988 from *Standard Federal Tax Reports, Commerce Clearing House Inc, (1988).* Capital gains tax treatment obtained from *Standard Federal Tax Reports, Commerce Clearing House Inc. (1962-1988).*

b With $50,000 cap on alternate maximum rate shield. All capital gains over $50,000 per individual taxed at 50% of regular marginal rate.

II. A General Signalling Theory.

The essential relationship that we wish to test is a key comparative static that we derive in a generic dividend signalling model. This result holds that the level of dividend payout needed to signal any given level of hidden firm characteristic is lower the higher the marginal tax rate on dividend income relative to capital gains. The simple intuition behind this result is that any signalling model implies selecting an optimal level of dividend payout by equating the marginal cost of the signal level to the marginal benefit of the signal level. The marginal cost of the signal (dividend payout) is a strictly increasing function of the marginal tax rate on dividends and a decreasing function of the marginal tax rate on capital gains. This monotonic relationship holds whether or not the signalling aspect of dividends derives from the tax rate. Since the marginal benefit of the signal is independent of the tax rate, we obtain the predicted inverse relationship between dividend payout levels and dividend tax rates for any quality of firm that signals -- any given amount of good news can be signalled with a lower dividend level if the marginal cost of dividend payouts is higher.

Consider the following general (scalar) dividend signalling model. Let D denote the dividend (signal) level, P denote the market price of a firm's shares, and θ denote some characteristic known only to the informed manager(s) of the firm. θ is the variable to be signalled to the uninformed. The informed's welfare depends on both the current share price and the true value of the firm.[6] Denote this relationship $V(P, \theta)$, where $\partial V/\partial P > 0$. In a signalling equilibrium P depends monotonically on D, P(D) with $P' > 0$, so that we can write $V(P(D), \theta)$.

There is some cost associated with sending the signal (issuing

---

[6]This is a critical characteristic of all dividend signalling models. Absent a dependence on current share price, the insiders would have no incentive to signal firm value to outsiders since they would not benefit from the transmission of information. Absent some concern for future share value (once the truth becomes known to all) the insiders always wish to raise current share price by signalling good news, unconcerned about the future impact of their false signal, so that no credible signal can emerge.

dividends). Part of this cost is the tax cost, τD. In order for a signalling equilibrium to exist, the marginal cost of dividend issuance must depend on firm type. Hence we write the total cost of dividends as $C(D,\theta) + \tau D$ (or more generally as $C(D, \tau, \theta)$, where $\partial C/\partial D > 0$, $dC/\partial\tau > 0$, $\partial^2 C/\partial D\partial\tau > 0$).

The informed choose D to maximize $V(P(D), \theta) - C(D, \tau, \theta)$, taking account of the dependence of P on D. In a signalling equilibrium, P(D) must be "informationally consistent" (Riley, 1979), so that $P(D) = \Pi(\theta, D)$, where $\Pi(.)$ denotes the "true" market value of a firm of type θ paying dividend D, and $\partial\Pi/\partial\theta > 0$ so that higher values of θ correspond to higher quality firms. An optimal signalling equilibrium requires selecting the level D^* that solves:

$$\max_D \; V(P(D), \theta) - C(D, \tau, \theta), \qquad (1)$$
$$\text{where } P(D) = \Pi(\theta, D).$$

As Riley shows, two necessary and sufficient conditions for the existence of a solution to (1) with $P(D) \neq \overline{P}$, $\forall$ D (i.e. that involves signalling) are:

(1) $\exists\ \hat{D} < \infty$ that maximizes firm value absent any signalling effects.[7]

(2) The "single-crossing property": $\dfrac{\partial}{\partial\theta}\left(\dfrac{\partial C(D,\tau,\theta)/\partial D}{\partial V(P(D),\theta)/\partial P}\right) < 0,\ \forall\theta.$ (2)

The existence of a positive tax differential on dividends versus capital gains ensures that condition 1 is satisfied since the optimal dividend level absent signalling is then zero. If there are no direct benefits to the signalling activity that depend on θ,[8] then condition (2) reduces to $\partial^2 C/\partial D\partial\theta < 0$; marginal signalling costs must be inversely related to firm quality.

Solving the first order conditions (FOC) for problem (1), we obtain

---

[7]This is needed to prevent all firms from setting D at such a high level that it is impossible to have any information transmitted.

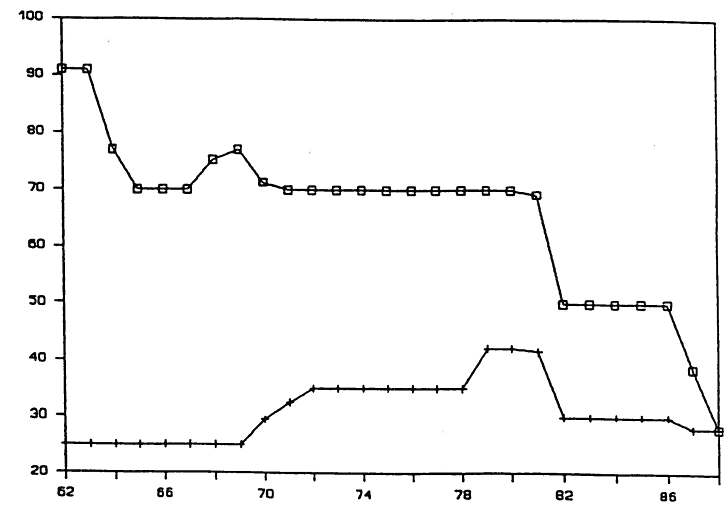
[8]This occurs for instance in Spence's signalling model where education adds no value but simply acts as a dissipative signal.

FIGURE ONE

Maximum rates on capital gains and dividend income 1962-1988

□ = maximum rate on dividend (regular) income.

\+ = maximum rate on capital gains income.



distinct tax regimes, ordered according to the relative tax disadvantage of dividends. Constructing portfolios ranked by dividend yield and size we test whether there is any positive association between the relative tax disadvantage of dividends and the excess returns associated with any level of dividend payout. Using distribution-free tests for independence we find no evidence of the predicted positive monotonic relationship.

Our failure to reject the independence of tax regime and excess returns for any given dividend yield implies that we cannot find any evidence in support of the signalling theory of dividends. Because we test only the weakest signalling prediction -- monotonicity, we provide strong evidence that information content in dividends is independent of the marginal cost of using dividends as a signal. This result is inconsistent with any tax-based model of dividend signalling, and more generally with any dividend signalling model *provided* that the marginal investor is taxed. These findings are also inconsistent with tax-based CAPM arguments. Investors do not appear to demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate them for the tax cost of dividends.

It is important to stress that the paper does not find that *no* information is revealed through dividend announcements. Indeed, the evidence of positive excess returns associated with greater dividend announcements is indicative of information release. Rather, the results reveal that information release is an indirect by-product of dividend announcements, instead of a direct signalling goal. By looking at the relationship between excess returns and the marginal tax cost of dividends, we distinguish between classical Spencian signalling explanations and indirect information release. The information content of dividends is uncorrelated with the tax costs of dividends, and hence inconsistent with Spencian signalling.

$$\frac{\partial V(P(D),\ \theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial \theta}\ \frac{d\theta}{dD} + \frac{\partial V(P(D),\ \theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial D} = \frac{\partial C(D,\ \theta,\ \tau)}{\partial D},$$

or, for the special case where $C(D, \tau, \theta) = C(D, \theta) + \tau D$,

$$\frac{\partial V(P(D),\theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial \theta}\ \frac{d\theta}{dD} + \frac{\partial V(P(D),\theta)}{\partial P}\ \frac{\partial \Pi(\theta,\ D)}{\partial D} = \frac{\partial C(D,\ \theta)}{\partial D} + \tau.$$

This FOC characterizes a differential equation for the optimal signalling function $D(\theta)$. Infinitely many solutions to this FOC exist. Identification of a particular signalling equilibrium is typically achieved by identifying the most efficient of all the solutions to the differential equation (see e.g. Riley 1979, Cho and Kreps 1987, or Banks and Sobel 1987). In the presence of dissipative signals (such as dividend taxes), this selection procedure identifies the solution in which the lowest quality firm selects a zero dividend. We should emphasize, though, that our empirical analysis is robust to the particular equilibrium selection, provided that the same equilibrium is selected over time. To derive the impact of changes in dividend taxes on the optimal dividend payout level for any quality of firm, we carry out the comparative static:

$$dD^{*}/d\tau \ = \ V_p\Pi_{DD}\theta' + V_p\Pi_{DD} - C_{DD} \ < 0 \ \text{ from SOC.} \qquad (3)$$

An increase in marginal tax rate, τ, leads the level of the dividend, D, to be set at a point where the net marginal benefit is higher, which, by second order conditions, ensures that $dD^{*}/d\tau < 0$.

Notice that in the absence of some direct benefit of dividends, dividend taxes alone do not satisfy the single-crossing property (2) since $\partial^2 C/\partial\tau\partial\theta = 0$. Thus, if dividends are to act as a signal in a scalar signalling model, some other aspect of dividend costs must generate the needed relationship between marginal dividend costs and firm quality that provides separation. However, *even* if dividend taxes are *not* the feature of the model that generates single-crossing, equation (3) shows that the optimal dividend level still depends on τ.

Engers (1987) shows that in multiple signal models only a quasi-concavity condition on the signalling cost function is required for a signalling equilibrium to exist. In particular, each signal need <u>not</u> satisfy the single-crossing property individually. Engers' condition is satisfied when there are two signals and the cost of one satisfies single crossing and the cost of the other is linear in the signal level. This is important since it means that if the other signal exhibits increasing marginal cost with quality, then dividends can act as a signal even if the only cost of dividends is a (linear) tax cost.

To derive the monotonicity condition in a multiple signalling model, note that with multiple signals the firm chooses the most efficient (cost-minimizing) mix of signals. The cost-minimizing mix of signals is determined where the marginal cost of each signal is equated across all signals used. An increase in the marginal cost of one signal (as would occur with an increase in the tax on dividend payouts) leads to an equilibrium in which the mix of signals is altered so that less signalling is done with the relatively more expensive dividend signal.

It is important to note that even though the information to be revealed by a dividend signal may be dynamic in nature, reflecting, for instance, an increase in the revenues derived from a project, the tax signalling cost is a *static* function of the *level* and *not change* in the dividend yield. That is, the signalling cost reflects levels of the dividend and the tax regime: $C(D, \tau, \theta) = C(D, \theta) + \tau D$. For this reason, any tests of signalling theories of dividends must focus on the level and not change in the dividend yield.

To see this most clearly, consider two firms, A and B, which both have good news to signal in consecutive periods. Firm A has better news than firm B to signal in both periods, but firm A has better news in period 1 than period 2, while firm B has better news in period 2 than period 1. That is, good news is ordered by:

$$R_1^A > R_2^A > R_2^B > R_1^B.$$

The equilibrium levels of the dividend signals reflects this ordering,

$$D_1^A > D_2^A > D_2^B > D_1^B,$$

but were one to look at the change in the dividend yields, one would find that

$$D_2^A - D_1^A < 0 < D_2^B - D_1^B.$$

Since the excess returns reflect the news revealed, they would satisfy

$$ER_2^A > ER_2^B.$$

evidence of a positive association between excess return and the relative tax treatment of dividends.[22] That is, we find no evidence that dividends are used to signal information to investors. Heoffding's D does not reject independence of tax regime and excess returns for any of the portfolios considered and Kendall's τ finds evidence of significant concordance for none of the portfolios (while finding discordance for one aggregated and two dis-aggregated portfolios). (Statistically insignificant) discordance is as common as (statistically insignificant) concordance looking across the set of portfolio estimates.

8. Conclusions

Bagwell and Shoven (1989) estimate that $68 billion in cash dividends were paid in 1985. The issue addressing researchers is why, given the apparently high associated tax costs, are dividends issued? One can always postulate that, *ceteris paribus*, certain investors have a preference for dividends over an identical dollar amount of capital gains, but such relative preferences would have to be extreme in light of the tax costs. It is also hard to believe that dividends are really irrelevant given the enormous resources that firms devote to determining their dividend policy.

This paper tests whether the underlying explanation for dividends is signalling based. We show a common prediction of both scalar and vector dividend signalling models is that of a positive rank order correlation between the tax disadvantage of dividends relative to capital gains and the amount of "good" news revealed by any given *level* of dividend yield.

We present tests of this monotonicity prediction using data on dividend yield and excess returns for American stocks for the time period July 1962 to December 1988. Over this period we detail sixteen

----

[22] We also considered the possibility that the effective capital gains tax is zero due to dynamic tax shielding strategies. The results are almost identical to those presented here: portfolios A and A-S still have significant discordance at the 5% level and portfolio C-M has significant concordance at the 10% level. No other portfolios exhibited a significant monotonic relationship between excess returns and dividend tax rates.

any portfolio.

Table 4 presents results for the size-controlled dividend yield portfolios. Only for portfolio (A-S), the smallest firms paying dividends between 0% and 0.5%, do we reject independence based on Kendall's τ at the 5% level. For portfolios B-S (the smallest firms paying dividends between 0.5% and 1%) we reject independence at the 10% level. The result for A-S is consistent with table 3 where portfolio A was the only portfolio showing significant discordance. Recall that discordance is *not* consistent with the signalling hypothesis which is being tested. Again there are as many negative values as positive values for our estimates of τ. Heoffding's test statistic D, shows no significance for any of the portfolios in the table. Recall that unlike Kendall's K, Heoffding's D statistic is consistent when τ is zero and the null distribution is false. Examining table 4 it is clear that many of our estimates for τ are very close to zero. Based on these results, we cannot reject the null hypothesis of independence.[21]

We then test to see whether, even though dividends do not appear to be used as a signalling instrument, there is information contained in dividend yield. For each of the 27 years in our sample we determine whether higher dividend yield portfolios are associated with greater excess returns. We then aggregate, first across years, and again across tax regimes, to generate our test statistic, exploiting the fact that under the null hypothesis of independence, the variance of the sum of the yearly statistics is the sum of the variances. Aggregating across years we find a $\hat{\tau} = 4.18$, so that we overwhelmingly reject the null hypothesis of independence at the .01% level, in favor of positive concordance: greater dividend yields are strongly associated with greater excess returns. Similarly, when we look at each of the 16 tax regimes, we obtain a $\hat{\tau} = 2.72$, so that we again reject the null hypothesis in favor of positive concordance, this time at the .4% level.

In summary, we find very strong evidence that dividend yield levels do contain positive payoff relevant information, but we find no

----

[21]Values of Spellmans' rho for each portfolio were also calculated. These produced identical levels of significance to those presented for Kendall's tau and hence are omitted.

A researcher looking for a positive relationship between excess return and change in dividend yield as evidence of dividend signalling would incorrectly conclude that firms were *not* using dividends to signal. This observation is of interest because there is evidence that changes (increases) in dividend levels are associated with greater excess returns than are levels of dividend yields. Signalling theories of dividends, however, make no prediction about the relationship between changes in dividend levels and excess returns[9]. Due to the static nature of the signalling cost, the generic dividend signalling model requires that the entire dividend be unanticipated. Since dividend levels are very predictable, this suggests that one should not be surprised that we reject signalling theories of dividends.

III. Dividend Signalling Literature.

Scalar signaling models (Bhattacharya 1979, Kalay 1980, Talmor 1982, Hakanson 1982, Miller and Rock 1985, and Bar Yosef and Hoffman 1986) seek to explain both dividend payouts and their informational content. The models differ in two regards, the form of dividend cost function, $C(D, \tau, \theta)$, and the motivation behind why the informed's objective function is of the form $V(P(D), \theta)$ rather than being concerned solely with $P(D)$ (current value) or solely with future value. Most assume that the cost of dividend issuance[10] arises due to the corporate transaction costs of refinancing cash shortfalls. To motivate the form of objective function they either assume that the form of managerial reward scheme produces a concern for both current and true

----

[9]Differential taxation implies that dividend yield levels always signal information. Still, one could imagine a model in which change in dividend yield levels also conveyed information. Our tests should be robust to such a formulation because we employ large representative portfolios of firms. Intratemporally, aggregation should wash out change effects in our level portfolios. Further, inter-temporally, given dividend level portfolios should have the same cross-sectional distribution of dividend yield changes, so that, holding the tax regime constant, the excess return for a given dividend yield portfolio should be the same across years, allowing us to base our tests solely on levels.

[10]All of these models have the feature that the dissipative signalling cost of dividends is a function of the level of, and not change in, dividend.

value or that managers act in the interests of current firm shareholders whose time horizon is such that they care about both current share prices and future share prices. Dividend payments by firms act as a signal of either current or expected firm value. Managers of high quality firms pay a dividend just high enough to distinguish themselves from low quality firms, which are discouraged from mimicking the behavior of the high quality firm by the greater probability of having to turn to costly external financing to pay the dividend given their lower expected cash flow. The benefit of paying a dividend is the same for all firm types (the positive current stock price return), but the cost is higher for a low quality firm because it is more likely to have to resort to external finance.

Multiple signalling models show that, although dividends are an expensive form of signal, dividends can still play a signalling role in an equilibrium where firms optimally choose among all possible signalling methods. John and Williams (1985), and Ofer and Thakor (1987) address how firms choose between stock repurchases and dividend payouts. The difference in the costs of these payout methods derives from the differential tax cost of dividends versus the dilution cost of repurchases. The optimal signal mix equates marginal costs across payout methods, so that the optimal dividend payout for any quality of firm decreases in the personal tax rate. Ambirush, John and Williams (1987), and Williams (1988) present models that combine the cash disbursement cost of Miller and Rock with John and Williams' relative cost structure for dividends and repurchases to determine the form of cash disbursement used. John and Mishra (1990), and John and Lang (1991) study how dividends, when combined with insider trades, can signal firm quality. Although the analysis does not rely on dividend taxes, an increase in the marginal cost of signalling with dividends (higher tax rates) alters the optimal mix of signals so that more signalling is done with insider trades and less with dividends. Kumar (1988) considers a model where due to differences in risk aversion between the informed manager and uninformed shareholders, there is a conflict of interest between these agents in determining the optimal investment level, and the resulting signalling equilibrium is partially separating. Finally, Bernheim (1991) develops a tax-based theory of dividend signals to determine the optimal mix of dividends and share

excess return and interpreted these results as evidence of tax effects resulting for the differential treatment of dividends and capital gains (e.g. Litzenberger and Ramaswamy 1979). These results are generally derived from estimates drawn from cross-sectional analysis which are aggregated over time. This time aggregation ignores the possibility that the yield-return relationship may vary through time due to changes in the tax regime. Since signalling theory predicts that the strength of the relationship between stock returns and dividend yields varies directly with the extent of the tax disadvantage of dividends, time aggregation across tax regimes can bias the results.

Recall too, that although this investigation is posed as a test of information-signalling models, that tax-based CAPM models provide the same predictions. Hence, a failure to to find the monotonic relationship provides strong evidence against the tax-effect theories which suggest that investors demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate them for the tax cost of dividends.

7. Results

Table 3 summarizes the results for the tests of the theory for the complete portfolios grouped by dividend yield and table 4 presents the corresponding results for the portfolios split by both dividend yield and firm size. Looking first at table 3, the monotonicity prediction that we wish to test implies that there should be concordance between the excess returns and the relative tax disadvantage of dividend income. We can see from the estimate for tau of rank correlation that there are as many discordant as concordant pairs and only for portfolio A may we reject the null hypothesis of independence based on Kendall's distribution free test statistic for independence (K). However, note that while we reject independence for portfolio A, the indication is that there is a *negative* association between tax regime and excess return for this portfolio, and *not* the *positive* correlation that signalling theory predicts! The large sample approximation of Kendall's K statistic, (KL), is also significant at the 5% level for this portfolio. Heoffding's distribution free test for independence (D) shows no significance for any dividend yield group and so according to this test, we cannot reject the null hypothesis of independence for number' arguments that equal dividend yield portfolios are 'informationally' identical over time. Based on these considerations, we create four portfolios based on dividend yield size increments of 0.5%:

Portfolio A: $0 < dy < 0.5$ %

Portfolio B: $0.5 \leq dy < 1.0$ %

Portfolio C: $1.0 \leq dy < 1.5$ %

Portfolio D: $1.5 \leq dy \leq 2.0$ %

For all firms, the dividend distribution events for each year are then categorized based on these dividend increments. Each firm event is then assigned to one of the portfolios as above. We then calculate for each dividend yield portfolio (A through D) the excess returns for each year 1962 through 1988. For each event documented in each dividend yield portfolio, the associated daily stock return for that event date is recorded. The average daily return for each dividend yield portfolio is then calculated. This average daily return for the event date is converted into an annual return and the annual return on the market portfolio is then subtracted to give us the dividend yield portfolio's annual excess return. The market portfolio is characterized as the value-weighted portfolio of all stocks listed on the NYSE as supplied by CRSP. The non-parametric tests were run using these complete dividend yield portfolio's excess returns.

We also control for possible size effects by sub-dividing dividend yield portfolios A through D on the basis of firm size. This controls both for the standard size effects and for any systematic differences in public information across firms of different sizes (for instance, there may be less public information about small firms so that more information is revealed through dividends). For each year, the portfolios were ranked with respect to firm size and then split into three separate (but equal) size groups, large (L), medium (M) and small (S). For each of these new portfolios, the annual excess return was calculated in the same manner as with the undivided dividend yield portfolios. The non-parametric tests were run on these size-based dividend portfolios using the calculated annual excess returns.

Many studies have found a relationship between dividend yield and

repurchases when there is the potential for costly bankruptcy. Robertson (1992) documents exhaustively that the monotonicity property holds in each of these models.

In independent research, Bernheim and Wantz (1993) also make the observation concerning the relationship between tax regime and the 'bang-for-the-buck' in excess return associated with a given dividend level. Although the signalling theory just implies a monotonic relationship between dividend yield level and excess return, they regress excess return on changes in dividend yield level, including an interaction with the tax regime, and a vector of control variables. The coefficient on the interaction term is significant (t statistic around 2.5) --- changes in dividend yield level have a bigger effect on excess returns when dividend income is taxed less favorably. This is an interesting finding, but following our theoretical development, it provides neither evidence for nor against existing dividend signalling theories. Even were one to run a regression with dividend levels rather than changes, one would still have to interpret a positive coefficient on the interaction term cautiously. This is because the best linear fit to a non-monotonic relationship may have a significant sign on the interaction coefficient, even though a non-monotonic relationship would be inconsistent with signalling. Consequently, the appropriate test of dividend signalling theory is a test of monotonicity, and there exist a variety of robust non-parametric tests of monotonicity.

IV Tax Treatment of Dividends and Capital Gains, 1962-1988.

The Tax Reform Act of 1986, the U.S. Federal tax code applied the same personal income tax rate to long-term capital gains and ordinary income (including dividend income). This was the first time since 1921 that the income tax code was not discriminated against dividend income relative to capital gains income. While this change may have been the most dramatic change in the relative treatment of capital gains versus ordinary income, numerous changes in the tax code over the last three decades have also affected the relative tax treatment of capital gains and other income.

Prior to the Tax Reform Act, 50 percent or more of capital gains were excludable from taxable income, reducing the effective tax rate on capital gains below that on other forms of income. Over the period

1960-1986 numerous changes in income tax rates, tax brackets, exclusion allowances, changes in maximum alternative tax rates, changes in the definitions of long-term capital gains, and changes in deductability allowances for capital losses, have changed the effective tax disadvantage of dividends relative to capital gains. In 1988, the top tax rate on capital gains was 28 percent (the same as on all income). Throughout most of the 1960s this rate was 25 percent. In the mid-1970s rates rose dramatically for high income earners, so that in 1978 the Congressional Budget Office estimates an effective top tax rate of 25 percent on capital gains compared to 22 percent in the late 1970s and rates of 14 percent in the early 1980s. Table 1 lists the most significant changes over this period[11].

Figure 1 plots the maximum rates on dividend income and capital gains incomes over the period 1962-1988. This figure provides a graphical illustration of the ordering of tax regimes. We choose the maximum rates as our primary focus since as is well known, the primary recipients of capital gains incomes are concentrated in upper income earners[12]. For 1982 the behavior of the very top rate is a deceptive measure of the behavior of the tax treatment of dividends for high income earners because the tax reductions of 1981 *disproportionately* favor those with incomes over $215,400. The rankings were adjusted slightly to account for this.[13] Otherwise, the tax regimes are clearly ranked. It is important to note that in ranking years from "most favorable to dividends" to "least favorable to dividends" we have assumed that inflation affects dividend income and capital gains tax income equally. With a non-indexed Federal Tax code (as was the case through the high inflation period of the 1970s) bracket creep adversely affects the tax treatment of dividends. Since capital gains taxes are levied on nominal capital gains, inflation also adversely affects the tax treatment of capital gains. We implicitly assume that the effect

----

[11]All rates quoted are for married couples.

[12]Avery and Elliehausen [1986] estimate 85% of common stock is held by the top decile of the wealth distribution; 44% by the top one-half percentile.

[13]Alternative rankings of this tax regime, including dropping it from the analysis, did not affect the results.

that for any event occurring in month t, an investor's information set would only contain information known at t-1, firm size and dividend yield are computed using month t-1 price data. Thus, dividend yield is defined as the dollar amount of the cash dividend divided by the price in the month prior to the event month, and firm size is defined as the prior month price times the number of shares outstanding.

In investigating the monotonic relationship between tax regimes and excess return we separately control for any other variable that may impact on excess return. We therefore construct portfolios categorized by dividend yield and firm size. Our prediction is that ***for any level of dividend yield*** the information released should be more favorable the more disadvantageous the tax treatment of dividends. We therefore clearly need to control for variations in dividend yield. In addition, there are well-documented intertemporal variations in average dividend yields for which we wish to control. The importance of firm size is well documented in studies of asset pricing (see Keim 1985 or Bajaj and Vijh 1990).[19]

For each year, therefore, the total number of events for all firms in the sample were categorized based on the size of the calculated dividend yield. In preliminary work we experimented with various grid sizes.[20] All of our reported results are for the dividend yield groupings increasing in increments of 0.5%. Two considerations led us to choose this dividend yield grouping. First, we find the vast majority of quarterly dividend yields are below 2%. Second, in constructing excess returns for each dividend yield portfolio we follow the methodology of Brown and Warner (1980) and assume that each portfolio is sufficiently large that it is well diversified (so that its beta is constant over time). Equally important, to the extent that there is heterogeneity in the amount of public information about firms within a year, aggregation into large portfolios permits 'laws of large

----

[19]Also, it is possible that the information released may be a function of firm size (e.g. analysts reports may reveal more information about marge firms than small firms, requiring less information signalling for these firms).

[20]Dividend yield divisions in increments of 0.2% and 0.3% were examined. No substantive difference in results was found.

6 Data.

We test for the monotonic relationship between tax rates and excess returns predicted by dividend signalling models using American stock market information provided by the Center for Research in Stock Prices (CRSP). Information on stock returns, firm value, and cash disbursement distribution information, data for price, shares outstanding and all dividend distribution information was obtained from the CRSP monthly master file and the CRSP daily returns file was used to obtain stock return information. The period of analysis begins in July 1962 when daily returns were first collected.

For a firm to be considered part of the sample it had to meet the following criteria:

(1) The firm had to be listed on the New York Stock Exchange as of July 1962 or later.

(2) Firms are only considered over the period when they make regular quarterly cash dividends.[17]

(3) A complete set of price, dividend distribution and return information was available for the declaration date of the dividend.

For each event (declaration of a dividend) occurring at time t, we calculate both firm size and dividend yield.[18] In constructing the dividend yield and firm size variables, it is important to note that since the distribution information is provided by the monthly master file, the price and shares outstanding information are available only for the last trading day in each month. However, dividend declaration dates and the dollar amount of the dividends are available for the actual event date within an event month. To take account of the fact

----

[17]For the majority of firms in the sample their cash dividend disbursement pattern was a quarterly dividend. In addition year-end "extra" dividends for these firms were included.

[18]The dividend yield measure employed is a short-term measure (one-day return) that takes account of dividend size but not dividend timing. This biases it in favor of finding tax-related effects. For our purposes this is preferable to a long-term measure that is biased against finding tax effects (see Rumsey 1988 and Kalay and Michaely 1992 for a discussion of timing effects and measurement of tax effects of dividends).

of inflation on the *differential* tax treatment of dividends and capital gains is insignificant. The ranking of years by tax regime is presented in Table 2.

V. Description of the Non-parametric Tests.

The generic signalling model does not predict a particular functional relationship between firm type and signal strength (here, firm value and tax regime, holding dividend yield constant).[14] The sole prediction is that of a monotonic relationship between type and signal strength, so a direct non-parametric test of this monotonicity condition is appropriate.

Even ignoring functional form concerns, it is no easier to interpret the results from a time series regression. Using a different dummy variable for each tax regime, one is left with determining whether the signs on the dummies are consistent with signalling theory: one must determine whether the signs have the desired monotonic relationship. Thus, even after imposing functional form, one is still left with a test of rank.

An additional concern is that, to implement a test of the monotonicity condition, one must control for the effect of differences in dividend yields across firms since the theory predicts that expected return is an increasing function of dividend yield.[15] Methodologies that do not separate out portfolios according to their dividend yield implicitly impose a linear relationship between dividend yield and return. One could imagine forming portfolios sorted by dividend yield for each announcement date. But then the beta for each portfolio would have to be estimated since the resulting portfolios would be too small to be assumed well diversified. In contrast, the direct tests proposed here can easily control for dividend yield across large portfolios and allow for testing of the model without imposing unnecessary, and

----

[14]For instance, a simple linear regression is, in fact, an inappropriate way to test signalling theory, since it perforce fits a linear relationship on the data when the true relationship may be non-linear.

[15]Implicitly this suggests problems with interpretations of the standard event study which regresses excess return on dividend yield, but fails to control for tax regime. Systematic co-variation in dividend yield and tax regime over time can lead to spurious signalling findings.

potentially false, parametric restrictions.

Equally important, aggregation into large representative portfolios of firms means that our analysis should be robust to cross-sectional heterogeneity in both the dividend cost function, $C_i(D, \tau, \theta)$, and the informed's objective function, $V_i(P(D), \theta)$. Cross-sectional heterogeneity means that the amount of news revealed by a given dividend can vary across (unobserved) firm type. However, intertemporally, given dividend level portfolios should have the same cross-sectional distribution of firm types, so that, holding the tax regime constant, the excess return for a given dividend yield portfolio should be the same across years. The only key property that we require is that the cost and objective *functions* not change over time, so that we can isolate the tax regime effects.

We propose two non-parametric tests that may be used to test the signalling theory of dividends. Kendall's τ calculates an estimate of the correlation between the ranks of tax type and the level of the signalling variable (excess return). Kendall's distribution-free test for independence (K), is based on Kendall's τ. This test examines the hypothesis that X,Y variables of a bivariate population are independent. The test is designed to detect a class of alternatives associated with either positive or negative values of τ. The second test we examine is Heoffding's distribution-free test for independence (D). This test examines the hypothesis that X,Y variables of a bivariate population are independent. The test is designed to detect a much broader class of alternatives than Kendall's K and unlike K it is consistent when tau is zero and the null distribution is false. These tests can be found in Hollander and Wolfe (1973).

5.1 Kendall's Tau and Kendall's Distribution Free Test for Independence.

Kendall's tau is calculated as follows. Assume that the data consist of a bivariate random sample, (X,Y), of size n. Define as concordant two observation pairs if both members of one observation are larger than the respective members of the other observation pair, for example (1, 3), (2, 4). Let N_c denote the number of concordant pairs out of the total $\binom{n}{2} = n(n-1)/2$ possible pairs. Let N_d denote the remaining pairs (the number of discordant pairs, for example (4, 1) and

(2, 3)). Kendall's tau is calculated as $\tau = \frac{N_c - N_d}{n(n-1)/2}$.[16] The hypothesis that we desire to test is the one-tailed test for positive association since the monotonicity hypothesis implies a positive relationship between the relative tax disadvantage of dividends (our X series) and the excess return associated with any given level of dividend yield (our Y series). Thus, the null ($\tau = 0$) and (one-sided) alternative hypotheses to be tested can be written as:

H_0: X and Y are mutually independent:

$$P(X \le a \text{ and } Y \le b) = P(X \le a)\, P(Y \le b)\ \forall\, a,b.$$

H_1: Larger values of X tend to be paired with larger values of Y.

A test based simply on $N_c - N_d$, Kendall's distribution-free test for independence has wider usage. When H_0 is true, the large sample approximation of the Kendall K statistic, KL, has an asymptotic N(0,1) distribution.

5.2 Heoffding's Distribution-Free Test for Independence.

Heoffding' D statistic has a symmetric construction and tests against both alternatives where the X's and Y's are positively associated and alternatives where the X's and Y's are negatively associated. Unlike τ, D does not distinguish between positive and negative rank order correlation; only two-tailed tests are appropriate. Since we are concerned with examining the data in the context of a one-tailed test for positive rank order correlation we use this test primarily to verify the results based on Kendall's K. Because pooling equilibria result in a prediction of τ equal to zero, we use this test to check for alternatives of dependence when τ is equal to zero (Kendall's K is inconsistent if $\tau = 0$). See Hollander and Wolfe (1973) or Heoffding (1948) for more details. The large sample approximation to the D statistic is given by $nD + (1/36)$; p-values for this distribution are in Hollander and Wolfe.

----

[16] Thus, tau measures whether there are more discordant or concordant pairs. With two perfectly positively correlated variables, all pairs are concordant and $\tau = 1$. With two perfectly negatively correlated series, all pairs are discordant and $\tau = -1$. With two independent series, the expected number of concordant and discordant pairs is identical so that the expected value of τ is 0.