Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Empirical Finance 12 (2005) 77–98
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/econbase

Testing dividend signaling models

Dan Bernhardt[a,*], Alan Douglas[b], Fiona Robertson[c]

[a] Department of Economics, College of Business, University of Illinois, 3330 Wohlers Hall, 1206 South Sixth Street, Champaign, IL 61820-6978, USA
[b] School of Accountancy, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
[c] Department of Economics and Finance, Albers School of Business and Economics, Seattle University, USA

Accepted 2 October 2003
Available online 24 March 2004

Abstract

This paper exploits a key monotonicity property common to dividend signaling models—the greater the rate that dividend income is taxed relative to capital gains income, the greater the value of information revealed by a particular dividend yield—to distinguish the hypothesis that dividends are used as a signaling device from the hypothesis that dividends contain information but are not used as Spencian signals. The monotonicity conditions are tested with robust nonparametric techniques. While the monotonic relationship predicted by signaling theory can be found, a more careful inspection reveals that it does not hold for different levels of the dividend signal, as required by signaling theory. This strongly suggests that existing signaling models cannot explain the dividend policy choices of firms.

Keywords: Dividend signaling models; Monotonicity condition; Income

1. Introduction

Since Ross (1977) and Bhattacharya (1979), financial economists have explored the properties of dividends arising from signaling models. Signaling theories were developed to explain positive abnormal returns following announcements by firms of an increase in dividends.[1] Such excess returns are puzzling in traditional models of perfect information

* Corresponding author.
E-mail addresses: danber@staff.uiuc.edu (D. Bernhardt), adouglas@uwaterloo.ca (A. Douglas), fiona@seattleu.edu (F. Robertson).

[1] See Pettit (1972), Laub (1976), Charest (1978), Aharony and Swary (1980), Asquith and Mullins (1983), and Eades et al. (1984).

0927-5398/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.jempfin.2003.10.002

because dividend income has been taxed less favorably than capital gains. To reconcile why firms pay dividends even though they are more adversely taxed, researchers have constructed signaling models in which firms convey their private information about firm profitability by dispersing costly dividends.

However, while the empirical evidence makes it clear that dividends affect market valuations, it is not at all clear whether management use dividends intentionally as a 'Spencian' signal, or whether the effect arises for some other reason.[2]

This paper exploits a key monotonicity property common to costly dividend signaling models to test signaling explanations of dividend policy. When the cost of the signal increases (i.e., dividend income is more heavily taxed), higher quality firms can differentiate themselves with a smaller amount of the signal. As a result, in equilibrium, any given signal choice (i.e., a particular increase or decrease in dividend yield) leads to a greater revision in the market's estimation of the firm's quality. Consequently, when dividend income is more highly taxed, as long as the increase in value exceeds the increase in taxes, a given increase in the signal should be associated with a greater increase in firm value. Similarly, a given decrease in the signal should be associated with a greater negative excess return when dividends are more highly taxed. Thus, signaling models predict that firm value is more sensitive to a more costly signal: for a given change in dividend yield there should be a monotonic relationship between excess returns and the dividend tax rate. Importantly, this prediction holds whether or not dividend taxes are the feature that facilitates the use of dividends as a signal—the prediction requires only that the marginal cost of the dividend signal be increased by higher relative dividend taxation.

The many changes in the relative tax treatment of dividends and capital gains in the United States provide an ideal setting in which to test for the predictions implied by dividend signaling theories. This paper tests the monotonicity prediction by identifying 18 different tax regimes that can be ordered according to their tax treatment of dividend income relative to capital gains, and investigating whether a given dividend signal (either an unexpected change in dividend yield or an unexpected change in the dividend level) is associated with a greater absolute excess return when dividend income is less favorably taxed.

Bernheim and Wantz (1995) (henceforth BW) also recognize that if dividend income is taxed more severely, then the share price response to a particular dividend signal predicted by the signaling models should be greater. They term the share price response per dollar of dividends the 'bang-for-the-buck'. To test the signaling hypothesis, they impose a linear structure and regress the excess return associated with a positive change in the dividend signal (increase in the dividend level) on the change in dividend and an interaction with the tax regime (as well as a vector of control variables). They find that the interaction term between change in dividend signal and relative tax burden on dividend income has a significantly negative coefficient, which is consistent with the predictions of dividend signaling models.

[2] The empirical relationship between dividends and valuations is documented in the citations of footnote 1. Non-signaling theories that can produce such a relationship include empire-building theories (e.g., Jensen, 1986), repurchase-based expropriation problems (e.g., Brennan and Thakor, 1990)), and behaviorial theories (e.g., Shefrin and Statman, 1984). Bernheim and Wantz (1995) illustrate formally how non-signalling theories can produce the dividend announcement effects described above. Allen and Michaely (1995, in press) provide excellent reviews of the dividend policy literature.

The results of BW contrast with those of Bernhardt and Lee (2001), who find no support for the bang-for-the-buck relationship using quantile regression techniques over the period 1978–1996. Other suggestive evidence, summarized by Allen and Michaely (in press), that formal signaling theory does not underlie dividends is found in the papers by Grullon and Michaely (2001), who find that the market responded more positively to dividend increases following the Tax Reform Act of 1986, which reduced dividend taxation; and Amihud and Murgia (1997) who find that excess returns in the US are similar to those in Germany, where dividends are favorably taxed.

In this paper, we adopt a very different empirical methodology in order to investigate more tightly the formal predictions of dividend signaling theory. We employ robust nonparametric techniques that focus on monotonicity without imposing additional restrictions on the relationship. In particular, we employ a test of rank order correlation to investigate the predicted monotonic relationship between tax regime and bang-for-the-buck. Our nonparametric methodology has two advantages relative to OLS. Most importantly, our methodology permits a test of the dividend signaling hypothesis at all levels of the change in dividend signal.[3] Dividend signaling theory implies that the monotonicity relationship exists at each level of the change in dividend signal. In contrast, OLS tests whether the monotonic relationship holds on average across the different levels of change in dividend signal. Additionally, our nonparametric methodology circumvents concerns about the effects of outliers or poorly behaved error terms on estimates.[4]

It is because we impose no functional form assumptions that we can test the monotonicity prediction at each level of the change in dividend signal: functional form assumptions do not link the tests at different stratifications, so it does not matter whether the impact of the signal choice differs at different levels of the signal. In contrast, if the signaling relationship is misspecified, e.g., if a linear structure is violated for some range of the dividend signal, then the resulting parameter estimates can mislead. Because we can test at different levels of the dividend signal, we reduce the likelihood that some feature of the economic environment, unrelated to dividend signaling, may spuriously generate an average monotonic relationship.

We consider two measures of the change in the dividend signal: (1) change in dividend yield, $((d_t/p_t)-(d_{t-1}/p_{t-1}))$, and (2) change in dividend level scaled by lagged price, $((d_t-d_{t-1})/(p_{t-1}))$, that Bernheim and Wantz employ. Almost all theoretical dividend signaling models, including Bhattacharya (1979) and John and Williams

[3] It is important to distinguish between the dividend signal and the information revealed by difference between the actual and expected dividend signal. In empirical investigations, the dividend (level or yield) remains the signal, but the market response is determined by the difference between the actual and expected signal. That is, if the market expects the firm is of type x, and the firm chooses a dividend consistent with type x, then there is no impact on share price. If instead, the market expects that the firm is of type y, and the firm chooses a dividend consistent with type x, the share price responds. This can give rise to confusion because empirical investigations often refer to the difference between the actual and expected signals as the dividend "signal". Here, we maintain the focus on differences from the expected signal, and because we assume that the expected signal is just last period's dividend signal, this is termed the change in the dividend signal.

[4] OLS has a host of auxiliary assumptions that may be violated (e.g., error terms may be correlated). The CAR distribution has 'fat tails' and is heavily right-skewed, so that the OLS estimates will be driven by the data in the tails of the CAR distribution.

(1985), directly or indirectly indicate that the appropriate signal measure is dividend yield: they imply an optimal dividend relative to share price (i.e., an optimal dividend fraction of ex ante firm valuation). Informal observation (e.g., the tendency for firms to pay the same dividend per share for many periods), however, suggests that firms may be more focused on dividend levels than dividend yields, and it is certainly the case that that changes in dividend yield are largely driven by changes in the stock price. While these observations suggest that formal signaling theory may not underlie dividend choice, their intuitive appeal leads us to consider $((d_t - d_{t-1})/(p_{t-1}))$ as a measure of the dividend signal change.

We first use change in dividend yield as our measure of the change in dividend signal. When we aggregate across stratifications of the dividend signal level (i.e., across different magnitudes of the change in dividend yields), we find that on average the “correct” monotonicity relationship holds between tax regime and bang-for-the-buck. However, signaling theory imposes monotonicity restrictions not only when ‘averaging’ across the changes in dividend signal, where many factors can influence the result, but also at each level of the change in dividend signal.

We then show that the monotonicity property holds only sporadically at a few (change in) dividend signal levels—primarily those with larger reductions in the dividend signal. Similar rejection rates occur both when the data is stratified by size, and when it is stratified by dividend yield. Thus, the major strength of our nonparametric methodology—it allows us to focus on different levels of the signal without concern for functional form—matters. In particular, stratification allows us to incorporate additional information that is pertinent to signaling theory, yet is obscured in aggregate tests.

We then search for evidence in favor of dividend signaling using the more intuitive dividend signal employed by BW—change in dividend level scaled by lagged price. Such a measure captures an environment where the dividend level, rather than yield, is used to signal. The tests based on this measure provide similar results: the monotonic relationship is supported only in pockets of the data, yet the average test statistics indicate that the correct monotonic relationship holds.

In summary, our findings are robust to different definitions of the dividend signal. With both definitions, for most levels of the change in dividend signal, we cannot reject the null hypothesis of independence between tax regime and bang-for-the-buck against the alternative hypothesis of positive correlation between the relative tax treatment of dividends and the magnitude of excess returns associated with dividend announcements. Overall, we believe that our findings represent significant evidence against signaling explanations for corporate dividend policy.[5]

A possible explanation for the contrast between our results and those of BW is that signaling theory predicts a monotonic relationship, but not necessarily a linear one. That is, the nonparametric methodology focuses solely on directional correlations, whereas

[5] Our methodology also permits a test of traditional tax-based dividend models (e.g., the CAPM models of Brennan, 1970; Litzenberger and Ramaswamy, 1979). In these models, investors demand compensation in the form of higher pre-tax returns on high dividend stocks to compensate them for the higher tax cost. The nonparametric tests detect no discernible relationship between the difference in returns on high versus low dividend stocks and tax regime, providing no support for the monotonic (linear) relationship predicted by the tax-based CAPM, or the more general (e.g., arbitrage-based) asset pricing theories (that may predict a monotonic, but not necessarily linear, relationship).

OLS focuses jointly on monotonicity and functional form (i.e., a linear relationship). Indeed, it is well established that the same linear structure does not hold for both dividend increases and decreases (Christie, 1994), so one cannot include both increases and decreases in the same OLS regression. It is for this reason that BW omit dividend decreases from their sample.[6] More generally, imposing a common linear structure may still be inappropriate for the subsample of positive dividend changes.

We conclude our analysis by exploring the underlying relationships in more detail. Further tests reveal that the relationship between excess return and tax regime is marginally stronger than that between bang-for-the-buck and regime. Since bang-for-the-buck is defined as excess return divided by change in dividend yield, this suggests that the excess return relationship may underlie the bang-for-the-buck relationship. Plausibly, this relationship reflects that the results are driven by non-signaling dividend theories.[7]

The paper is organized as follows. The next section discusses the various tax code changes that have occurred in the treatment of capital gains taxes in the US over the period 1960–1996. Section 3 describes the data and details the empirical methodology and nonparametric tests employed. Section 4 presents the test results for the dividend signaling models. Section 5 draws conclusions.

2. Tax treatment of dividends and capital gains

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% 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 deductibility 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% (the same as on all income). Throughout most of the 1960s this rate was 25%. In the mid-1970s, rates rose dramatically for high income earners, so that in 1978 the Congressional Budget Office estimates an effective top

[6] In fact, the bang-for-the-buck feature of signaling models can be derived *only* in models in which increases and decreases occur. Specifically, in dividend signaling models, initial firm value is based on the expected dividend. Some firms credibly separate themselves by choosing a signal (dividend) that exceeds expectations, so that others fall short of the expectations. Thus, our ability to include 'bad news' (dividend decreases) again highlights an important advantage of our nonparametric approach. Indeed, of the four stratifications for which we find a statistically significant monotonic relationship, three feature dividend reductions.

[7] For example, firms may pay dividends to help control the opportunistic spending of free cash flow as in Jensen (1986) or the reasons cited in footnote 2. See Conclusion for a discussion.

Table 1
Major tax changes affecting capital gains and dividend income

Year(s)	Income tax changes	Inclusion rate (%)	Alternate maximum rate	Holding period (months)
1962–1963	Top rate = 90%	50	25%	6
1964	Rates lowered, top rate 77%	50	25%	6
1965–1966–1967	Rates lowered, top rate 70%	50	25%	6
1968	7.5% tax surcharge	50	25%	6
1969	10% tax surcharge	50	25%	6
1970	2.5% tax surcharge	50	29.5%	6
1971	No changes	50	32.5%	6
1972–1976	No major changes	50	None	9
1977–1978	No major changes	50	None	12
1979–1980	No major changes	60	None	12
1981	Rates lowered 5%	60	None	12
1982	Top rates cut to 50% from 69% Other rates lowered 10%	60	None	12
1983	Rates lowered 10%	60	None	12
1984	Rates lowered	60	None	6
1985–1986	No major changes, some bracket adjustments	60	None	6
1987	Lower rates, fewer brackets	100	28%	6
1988	Lower rates (28%)	100	None	6
1989–1990	Rates 28%	100	21%	12
1991–1992	Rates increased to 31%	100	24%	12
1993–1996	Rates increased to 39.6%	100	26%	12

Income tax rates quoted are for married couples.
(a) Income tax data from tables A-3 to A-6 of Pechman (1987), updated to 1996 from Standard Federal Tax Reports, Commerce Clearing House, and from the forms and publications section of the IRS website. Capital gains tax treatment obtained from Standard Federal Tax Reports, Commerce Clearing House (1962–88), and from the forms and publications section of the IRS website.
(b) With US$50,000 cap on alternate maximum rate shield. All capital gains over US$50,000 per individual taxed at 50% of regular marginal rate.

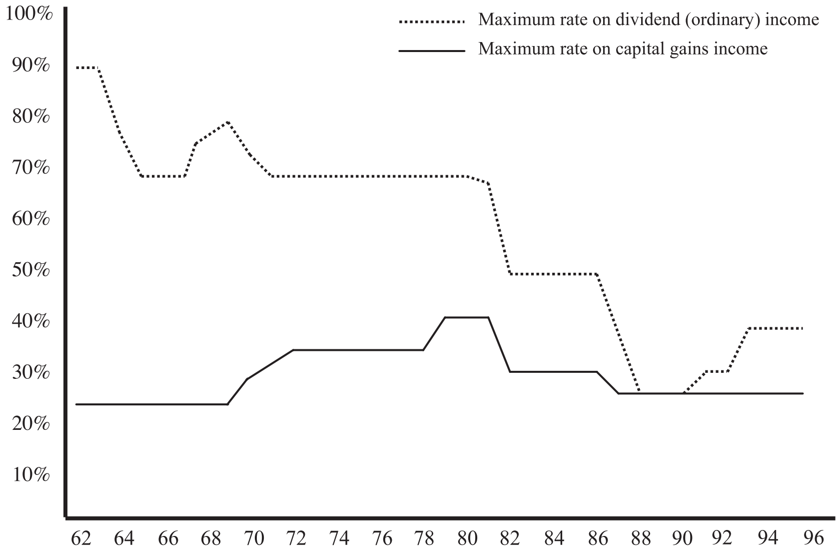
tax rate of 25% on capital gains compared to 22% in the late 1970s and rates of 14% in the early 1980s. Table 1 lists significant changes over this period.[8]

Fig. 1 plots the maximum rates on dividend income and capital gains incomes over the period 1962–1996. This figure provides a graphical illustration of the ordering of tax regimes presented in Table 2. 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.[9] 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 US$215,400. The rankings were adjusted slightly to account for this.[10] Otherwise, tax regimes are clearly

[8] All rates quoted are for married couples.

[9] 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.

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



Note: data are obtained from the sources listed in table 1.

Fig. 1. Maximum rates on capital gains and dividend income 1962–1996.

ranked. It is important to note that in ranking years from "most favorable to dividends" (lowest relative taxation of dividends to capital gains) to "least favorable to dividends" (highest relative taxation of dividends), we have assumed that inflation affects dividend

Table 2
Rankings from lowest relative dividend taxation R_1 to highest relative dividend taxation R_{18}

R_1	1988–1990
R_2	1991–1992
R_3	1987
R_4	1993–1996
R_5	1985–1986
R_6	1984
R_7	1983
R_8	1982
R_9	1972–1978
R_{10}	1979–1980
R_{11}	1981
R_{12}	1965–1967
R_{13}	1971
R_{14}	1970
R_{15}	1969
R_{16}	1968
R_{17}	1964
R_{18}	1962–1963

R_j denotes the jth most favorable difference in tax rates between dividends and capital gains (for example in 1988 this difference is 0, in 1962–1963 for the highest tax bracket, this difference is 65% (90% on dividend income versus 25% on capital gains)). See Fig. 1.

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 of inflation on the differential tax treatment of dividends and capital gains is insignificant.

Our measure of the relative tax burden differs from that used by Bernheim and Wantz. Bernheim and Wantz use a measure taken from James Poterba,

$$\theta_t = \sum_{j=1}^{S} w_{jt} \frac{(1 - m_{jt})}{(1 - z_{jt})(1 - \tau_t^{\mathrm{u}})},$$

where w_{jt} is the equity ownership weight of investor class j, m_{jt} is the marginal dividend tax rate on investors in class j, z_{jt} is the accrual-equivalent capital gains tax rate, τ_t^{u} is the rate of tax on undistributed profits, and S is the number of distinct shareholder classes. While θ_t captures broad movements in the relative tax treatment—if the relative tax burden on dividend increases, then θ_t falls and vice versa-consideration of the likely marginal investor leads us to focus on individuals in the top tax categories. However, since the rank order of the two measures accord for the period of our sample for which Poterba has calculated θ_t (through 1992), and this is what our nonparametric test exploits, differences in measure cannot underlie our different findings.

3. Data and methodology

We obtain data on stock returns, shares outstanding, share price and dividend distributions from the Center for Research in Stock Prices (CRSP) daily returns file. The period of analysis is 1962–1996. For a firm to be included in the sample, it had to be listed on the New York Stock Exchange, make regular quarterly cash dividends, and have a complete set of price, distribution and return information at the declaration date of each dividend.

We use the information available at the dividend declaration dates to construct stock portfolios for each tax regime. In constructing the portfolios, we omit declarations surrounding changes in tax regime, since a change in tax regime can result in a change in dividends (and dividend expectations) in the absence of asymmetric information regarding firm quality. We also omitted "outliers" with daily excess returns exceeding 50% in absolute value, as they are likely data entry errors.[11]

Signaling theory implies a positive (negative) relationship between the relative tax disadvantage of dividends and the excess return associated with a given increase (decrease) in the dividend signal. The excess return associated with a given dividend

[11] In constructing our portfolios, we do not make adjustments for dividend announcements that are close to earnings announcements. This is because we focus not on the magnitude of abnormal returns per se, but on how the abnormal returns are related to the cost of dividends, which should not be systematically related to earnings announcements (i.e., the same reason as in Bernheim and Wantz, 1995, pp. 542–543). In addition, it ensures that our results are directly comparable to Bernheim and Wantz.

signal is termed the bang-for-the-buck (BFB). Thus, the signaling prediction is that, for increases (decreases) in dividend yield, there should be a monotonic increasing (decreasing) relationship between the bang-for-the-buck and the dividend tax rate.

To test this prediction, we must distinguish between the anticipated and unanticipated portions of dividend signal. This requires a stand on what the expected signal should be. We consider two possibilities:

1. The expected signal is the dividend yield from the previous quarter, so that the change in dividend yield, $((d_t/p_t)-(d_{t-1}/p_{t-1}))$, captures the unanticipated portion.
2. The dividend level, rather than dividend yield is the relevant signal, so that the expected signal is the dividend from the previous quarter, in which case the change in dividend, d_t-d_{t-1}, captures the unanticipated portion.

Bernheim and Wantz (1995) effectively take the second approach. They assume that the unanticipated portion is given by $((d_t/p_{t-1})-(d_{t-1}/p_{t-1}))$: since the denominator in both components is the previous quarter's price, they are restricting attention to the subset of firms whose dividends change—their signal is essentially the change in the dividend level.[12]

The generic signaling model does not impose a particular functional relationship. Consequently, we employ a direct nonparametric test of this monotonicity condition, based on the portfolio returns in each tax regime, as follows.

We begin by calculating the change in dividend measure at each declaration date. The change in dividend yield is calculated as $((d_t/p_{t-1})-(d_{t-1}/p_{t-1}))$, where p_t and p_{t-1} are average stock prices from the 5 days immediately preceding the declaration date (in aggregate, our data set contains 196,923 declarations associated with a change in dividend yield). The change in dollar value of dividends signal measure, $((d_t-d_{t-1})/(p_{t-1}))$, is calculated similarly.[13]

For each declaration that represents a change in dividend signal (e.g., a change in yield), the daily stock return associated with the event date is recorded. We subtract the CRSP value-weighted daily market return from the stock return to obtain a market-adjusted return that is our measure of the "excess" daily return. We then either (i) simply average the daily excess returns to obtain the portfolio excess return for each tax regime, or (ii) scale the excess return by the dividend signal to obtain the bang-for-the-buck (excess return per dollar of signal), and then average the scaled excess returns to obtain the portfolio BFB for each tax regime.[14]

[12] The simple representation of market expectations (the value signalled in the previous period) is tenable, both in BW and here, because we focus on differences in announcement effects for different tax regimes.

[13] Using this measure reduces the number of observations due to the well-known stickiness of dividends per share.

[14] For consistency, we employ the same definition of excess returns as Bernheim and Wantz (1995). This simple definition is appropriate in this context because the tests focus on differences in excess returns. To verify, we repeated our analysis using beta-adjusted excess returns (calculated using the market model). This produced the same qualitative results, although the number of observations was reduced due to out-of-sample estimation of betas.

We subsequently stratify the data into smaller portfolios based on the magnitude of the change in dividend signal (in Table 5 we also stratify based on firm size or the firm's dividend yield category, as seen below). Our stratification categories are chosen to ensure that each portfolio is large enough to be well diversified. In particular, we ensure that each portfolio employed in our tests contains at least 50 individual declarations. This makes our analysis robust to both inter-temporal and cross-sectional heterogeneities, allowing us to treat the portfolios as informationally identical (see Brown and Warner, 1980).

3.1. Description of the nonparametric tests

We test the monotonicity prediction for different stratifications of the data with three different nonparametric test statistics. For each stratification category, we examine the hypothesis that the portfolio excess returns (or bang-for-the-buck) and tax regimes are independently distributed using Kendall's tau statistic. This technique treats the data as a sample from a bivariate distribution, which we illustrate here for the case of (unscaled) excess returns. The test statistic is calculated by organizing the data into observation pairs of the form (t_i, ER_i), where t_i represents a particular tax regime and ER_i represents the excess return on a particular portfolio during that tax regime (recall that the portfolios are determined by the particular stratification categories).

Kendall's tau statistic is used to test, in a distribution-free manner, whether the two random variables (t and ER) are correlated. Recall that the 18 different tax regimes in our study are ranked according to the relative treatment of dividends (as in Table 2). The test statistic is calculated by determining the number of "concordant" and "discordant" pairs, denoted by N_c and N_d, respectively. Two observation pairs are concordant if $(t_i - t_j)(\mathrm{ER}_i - \mathrm{ER}_j)>0$, i.e., if both members of one observation pair are larger than the respective members of the other observation pair (e.g., (1, 3), (2, 4)). Otherwise, they are discordant (e.g., (4, 1) and (2, 3)). The test statistic for the bang-for-the-buck tests is calculated analogously, using scaled rather than raw excess returns associated with each declaration, so that two observation pairs are concordant if $(t_i - t_j)(\mathrm{BFB}_i - \mathrm{BFB}_j) > 0$.

Since there are $n = 18$ tax regimes, the maximum possible number of concordant pairs is $N = (n(n-1)/2) = 153$. Kendall's tau is calculated as $\tau = (K/N)$, where $K = N_\mathrm{c} - N_\mathrm{d}$. As such, τ provides an estimate of the correlation between tax regime and the amount of information conveyed by a given signal. For example, if the two variables are perfectly positively correlated, all pairs are concordant and $\tau = 1$. Alternatively, with two independent variables, the expected number of concordant and discordant pairs is equal and the expected value of $\tau = 0$.[15]

The null and alternative hypotheses tested are:

H0. Excess return and tax regime are mutually independent ($\tau = 0$).

H1. Higher absolute values of excess return are associated with higher dividend-tax regimes ($\tau>0$).

[15] For further discussion, see Hollander and Wolfe (1973), p. 185.

We also calculate an aggregate Z test statistic that allows us to aggregate information obtained from each stratification category. Specifically, denote the value of K for stratification category i by K^i. In large samples, z^i=((K^i−E_0(K^i))/σ(K^i)), has an asymptotic N(0,1) distribution under the null hypothesis of independence. Therefore, the aggregate test statistic, $Z = \sum_i z^i / \sqrt{m}$, where m is the number of stratification categories, has an asymptotic N(0,1) distribution under the null.[16] Accordingly, the Z statistic also tests a null of independence against an alternative that higher absolute values of excess return are associated with higher dividend-tax regimes (Z>0) The aggregate Z statistic averages the information from the individual z statistics, determining whether, on average, there is a monotone relationship, much like OLS does.

4. Results

4.1. Tests using the change in dividend level as the signal

Signaling theory implies that if the firm's dividend announcement is greater (less) than expected there should be a positive (negative) excess return, and that these excess returns should be amplified when dividends are taxed more heavily. In particular, signaling theory predicts a concordant relationship between the absolute value of the bang for the buck (excess return per dollar of the signal) and the relative tax disadvantage of dividends.[17]

We first consider change in dividend yield as our measure of the signal change. Table 3 presents the results of our nonparametric tests of the monotonicity prediction, both for the sample as a whole (panel A), and for the portfolios stratified by change in yield (panel B). In each panel, the first row is the number of concordant pairs (out of a possible 153 pairs). The second row presents the Kendall's τ and the third presents the standard normal approximation z test statistics. Below these, we provide the average and median numbers of observations in each portfolio. The aggregate Z statistics are presented in the final rows.

The results in panel A show that, for the sample as a whole (i.e., tests where all levels of the change in signal are grouped together, or where only increases and decreases are considered), the null hypothesis of independence cannot be rejected. These nonparametric results contrast with the OLS results of BW.[18] To investigate further, we examine whether the monotonic relationship holds at different ranges of the signal in panel B, as predicted by signaling theory. That is, we repeat our nonparametric tests after controlling for the magnitude of the change in signal (i.e., the magnitude of the difference between the actual and expected dividend signal). The specific signal stratification categories in panel B (e.g., the portfolio of stocks that increased their yield by between 0.2% and 1% during the first

[16] Our tests for the individual stratifications rely on the individual tau statistics (the same results obtained with the individual z statistics); the individual z statistics are used only to compute the aggregate Z statistics.

[17] We multiply the excess return for decreases by negative one so that concordance is supportive of the theory for both increases and decreases in dividend yield.

[18] The contrasting results are not due to the additional data from 1989 to 1996: our results are qualitatively unchanged if we use BW's sample period (1962–1988). The results in panel A are also qualitatively unchanged if changes in yield close to zero, e.g., $|\Delta DY|$>0.04%, are omitted.

Table 3
Tests for monotonicity between "bang-for-the-buck" and tax regime with stratification by the magnitude of the change in dividend signal

Panel A: Minimal stratification:

	All changes in dividend yield	All decreases in yield	All increases in yield
N_c	83	83	85
τ-statistics	0.08	0.08	0.11
z-statistics	0.49	0.49	0.64
Observations:			
Total in test	196,923	98,958	97,965
In average portfolio	10,940	5498	5443
In median portfolio	5989	3156	3312

Panel B: Additional stratification by change in dividend yield

	$-1\% < \Delta DY < -0.2\%$	$-0.2\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.07\%$	$-0.07\% < \Delta DY < -0.04\%$	$-0.04\% < \Delta DY < -0.015\%$	$-0.015\% < \Delta DY < 0$	$0\% < \Delta DY < 0.015\%$	$0.015\% < \Delta DY < 0.04\%$	$0.04\% < \Delta DY < 0.07\%$	$0.07\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 0.2\%$	$0.2\% < \Delta DY < 1\%$
N_c	108	96	94	69	77	78	79	79	86	93	85	97
τ-statistics	0.41***	0.25*	0.23*	−0.1	0.01	0.02	0.03	0.03	0.12	0.22	0.11	0.27*
z-statistics	2.39***	1.48*	1.33*	−0.57	0.04	0.11	0.19	0.19	0.72	1.25	0.64	1.55*
Observations:												
Total	17,042	18,941	15,719	16,587	17,526	12,297	11,740	16,663	15,869	15,311	19,412	18,375
Average portfolio	947	1052	873	922	974	683	652	926	882	851	1078	1021
Median portfolio	742	667	422	543	547	388	358	533	564	567	718	815
Aggregate Z statistics:												
$\Delta DY < -0.04\%$	2.32**											
$\Delta DY > 0.04\%$									2.08**			
All together	2.69**											

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the one-sided test. Significance levels for Kendall's τ-statistics, $\tau^i = ((N_c^i - N_D^i)/(N_c^i + N_D^i))$ are presented in Hollander and Wolfe (1973), p. 384.
The number of observations in the average portfolio for each τ-test is the total divided by 18 (there are 18 tax regime portfolios); the median is the number of observations in the portfolio for which there are an equal number of portfolios with more and with less observations.

tax regime as in the last column) are chosen to ensure a sufficient number of observations for each portfolio.[19]

Panel B reveals that the signaling prediction holds only sporadically for the different stratification categories. In particular, only 4 of the 12 individual τ and z test statistics reject the null hypothesis of independence between BFB and tax regime (three at the 90%, and one at the 99% confidence level). Also note that three of the four stratifications for which we reject independence are for dividend decreases (which BW do not consider); only one of the six individual test statistics for increased dividend yield is significant. This sporadic support is inconsistent with signaling theory, which implies a monotonic relationship for all levels of the dividend signal.

Nonetheless, note that for the larger change in yield stratifications in panel B, $|\Delta\text{DY}|>0.04\%$, most stratifications have more concordant than discordant pairs. Signaling theory would predict zero excess returns across all tax regimes when there are no changes in yield, and since our measures (e.g., of the expected signal) are noisy, it is plausible that noise would dominate for very small changes in dividend yield. For this reason, we especially focus on observations with larger changes in yield. When we aggregate across stratification categories, the last three rows of panel B reveal that the aggregate Z tests reject the null in favor of the monotonic relationship predicted by the signaling theory at the 95% confidence level.

How then do we interpret these findings? A balanced interpretation would first note that when we aggregate across stratification categories, the Z statistics reveal that on average, there is a positive relationship between bang for the buck and tax regime in the data. This is consistent with dividend signaling theories, and the results of BW. However, this is but weak support for dividend signaling theory. Signaling theory imposes many restrictions on the data; in particular, it implies a monotonic bang for the buck relationship at all levels of the signal. We simply do not find this, especially for increases in the dividend signal. Our 18 distinct tax regimes imply 153 distinct orderings, so that our sample size, while not enormous, is certainly not small. Moreover, each stratification was chosen so that many observations went into our calculation of the excess return associated with each tax regime; in our view this, too, is not a big concern (subject to the caveat that to ensure that we have enough observations to have confidence in our excess return calculations, we still allow non-trivial ranges of the dividend signal within a stratification). In the end, we view the lack of support at different signal levels documented in Table 3 as compelling evidence against signaling theory.

To further investigate the robustness of our findings, we repeat our tests using alternative definitions of the signal and bang-for-the-buck variable, and, to the extent that sample size considerations permit, introduce additional control variables.

4.2. Tests using dividend level as the signal

In Table 3, we employ the change in dividend yield, defined as $((d_t/p_t)-(d_{t-1}/p_{t-1}))$, as the hypothesized signal of new information, and scale excess returns to produce the "bang-for-

[19] As discussed above, large portfolios are maintained to ensure the integrity of the tests (this was particularly an issue for regime 14 (1968)). Our results are not sensitive to the particular stratification categories chosen. Only 1% of quarterly dividend yield changes exceed 1% in magnitude.

the-buck'' variable. As discussed earlier, a plausible alternative is to employ the scaled change in the dividend level, defined as $((d_t - d_{t-1})/(p_{t-1}))$, as the signal. In particular, Bernheim and Wantz employ this measure of the signal, and scale excess returns in this way.

This definition of change in yield essentially focuses the analysis on changes in dividend levels, and begs the question of what is the appropriate signal. While it seems intuitive to consider the dividend itself as the signal, the choice of dividend is typically related to share price because the equilibrium dividend deters mimicry at a cost that is reflected in the pre-signaling share price. For example, in Bhattacharya (1979), the probability of incurring the costs of a dividend shortfall is contingent on the factors that determine the pre-signaling share price; and in John and Williams (1985) the reduction in dilution is contingent on factors that determine the pre-signaling share price.[20]

To examine how the definition of the signal affects the results, we repeat our tests using the change in the dividend level as the signal of new information. The results are presented in Table 4. Employing the new definition of signal sharply reduces the number of observations, because all observations for which dividend level did not change must be dropped, so that fewer stratification categories are possible (while maintaining at least 50 observations in each stratification portfolio). In panel A of Table 4, bang-for-the-buck is calculated by scaling the excess returns associated with these announcements by BW's definition of the change in dividend yield. In panel B, bang-for-the-buck is calculated by scaling the excess returns by the change in dividend yield as in Table 3. (N/A) indicates that we could not construct sufficiently large portfolios for our tests for particular sub-categories due to insufficient observations.

Table 4 reveals the same basic pattern as Table 3. The first set of individual τ tests in both panels indicates that the null hypothesis of independence cannot be rejected when all levels of the signal are grouped together (including all increases and all decreases). The second set of individual τ tests reveal that the signaling prediction holds only for one of the stratification categories (the largest dividend decreases). However, the aggregate Z test reveals that when we aggregate across stratification categories, on average there is a monotone relationship. Again, we view the individual test results as compelling evidence against signaling theory.

4.3. Controlling for firm size and average dividend yield

The empirical literature studying dividends finds that excess returns are related to both firm size and average dividend yield (Keim, 1983; Bajaj and Vijh, 1990). To the extent possible with our methodology,[21] we now allow for the possibility that the

[20] Empirically, there is a tendency for dividends to ''stick'' to last period's dividend per share. This suggests either coarse signaling models (e.g., Kumar (1988)) that produce such stickiness are relevant, or that firms sometimes choose not to signal. It is therefore appropriate both to look at firms for which the dividend level changed (i.e., look at firms that choose to signal) and at firms for which the dividend yield changed sufficiently (if ''stickiness'' reflects coarse signaling).

[21] The nonparametric methodology cannot support as many control variables as OLS, so we cannot incorporate all of the control variables employed by Bernheim and Wantz. However, of the control variables that BW use and we do not, all save a measure of credit-worthiness are insignificant—and credit-worthiness has the wrong sign from a signaling theory perspective—suggesting that the major advantage of the OLS methodology over our nonparametric methodology is not important here; but that the many advantages our methodology remain.

Table 4
Monotonicity between "bang-for-the-buck" and tax regime: results using dividend level as the signal

Panel A: Excess returns scaled by $D_t/P_{t-1}-D_{t-1}/P_{t-1}$

	All changes in dividend yield	All decreases in yield	All increases in yield
N_c	83	91	81
τ-statistics	0.08	0.19	0.06
Observations:			
Total in test	30,964	6770	28,781
In average portfolio	1720	376	1599
In median portfolio	841	254	866

Additional stratification by change in dividend yield

	$-1\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 1\%$	Z statistic
N_c	103	N/A	91	82	2.04**
τ-statistics	0.35**	N/A	0.19	0.07	
Average (median) observation	267 (181)	159 (77)	329 (156)	545 (246)	

Panel B: Excess returns scaled by $D_t/P_t-D_{t-1}/P_{t-1}$

	All changes in dividend yield	All decreases in yield	All increases in yield
N_c	73	70	85
τ-statistics	-0.05	-0.1	0.11
Average (median) observation	2101 (1230)	503 (386)	1599 (866)

Additional stratification by change in dividend yield

	$-1\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 1\%$	Z statistic
N_c	105	N/A	91	84	2.21**
τ-statistics	0.37**	N/A	0.19	0.1	
Average (median) observation	346 (285)	205 (131)	391 (239)	640 (331)	

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the one-sided test. Significance levels for the τ-statistics, $\tau^i=((N_c^i-N_D^i)/N)$, are presented in Hollander and Wolfe (1973), p. 384. The average number of observations is the total in each τ-test divided by 18; the median is the number in the portfolio for which there are an equal number of portfolios with more and with less observations.

signaling relationship differs across firms with differing levels of capitalization or dividend yields.[22]

The tests with this further stratification are presented using dividend yield as the signal, as in Table 3, because the loss of observations when dividend level is used precludes

[22] For example, stratification by size controls for standard size effects as well as systematic differences in public information across firms of different size (e.g., there may be less public information about small firms so that dividends convey relatively more information).

additional stratification. To control for size effects, we stratify the data into three even size categories (large, medium and small, each containing one third of the observations). We then stratify each size category into four change-in-dividend-yield categories, again chosen to ensure sufficient observations in each portfolio. Similarly, to control for effects related to dividend yield (panel C of Table 5), we first stratify the data into three even yield categories—high (1.15% to 3%), medium (0.65% to 1.15%) and low (0% to 0.65%). Again, these categories were chosen to ensure each portfolio contains sufficient observations to maintain the statistical integrity of our tests.[23]

The results are shown in Table 5. Panel A simply repeats the tests of Table 3 with the coarser stratification required to allow for additional stratification by size and yield. The results with additional stratification by firm size and dividend yield are presented in panels B and C, respectively. The results are similar to those above. In panel B, only 4 of the 12 individual τ and z test statistics are significant at conventional levels, and 3 of these 4 are for large decreases. Similarly, in panel C, only 3 of the individual τ and z test statistics are significant, and 2 are for large decreases. Again the aggregate Z test reveals that on average a monotone relationship exists, but this support is driven primarily by a few stratification categories, and neglects the lack of monotonicity at many levels of the signal, which contradicts signaling theory.

4.4. Supplemental analysis

We conclude by exploring whether there is a monotone relationship when we look at excess return and tax regime for a given dividend signal, rather than looking for a monotone relationship between bang-for-the-buck and tax regime, as inTables 3–5. We do so in order to get a better understanding of how the excess return component of the bang-for-the-buck measure is related to the tax regime.

Tables 6 and 7 repeat our tests when the excess returns associated with dividend announcements are not scaled by the change in the dividend signal. Panel A of Table 6 reveals that there is slightly more evidence of a monotone relationship between excess return and tax regime than between bang-for-the-buck and tax regime when all changes in yield (or all increases) are grouped together (compare with panel A of Table 3). Panel B of Table 6 reveals that the relationship between excess return and regime is similar to that between bang-for-the-buck and regime when we stratify by the magnitude of the signal (panel B of Table 3). Panel A of Table 7 indicates that there is a much stronger monotone relationship between excess return and tax regime than there is between bang-for-the-buck and tax regime when the dividend level is taken as the signal, especially for dividend increases (compare with Table 4). Panel B of Table 7 further documents the stronger monotone relationship between excess return and tax regime when we stratify by size and yield.

Since BFB=(ER/ΔDY), Tables 6 and 7 suggest that the bang-for-the-buck relationship might be driven by the excess return relationship: controlling for the magnitude of the

[23] Varying stratification definitions (provided each stratification has sufficient observations) does not substantively alter our findings.

Table 5
Monotonicity between "bang-for-the-buck" and tax regime: additional stratification by firm size and dividend yield with change in yield as the signal

Panel A: Stratification by change in dividend yield only

	−1% < ΔDY < −0.11%	−0.11% < ΔDY < −0.04%	0.04% < ΔDY < 0.11%	0.11% < ΔDY < 1%	Z statistic
N_c	103	81	89	88	2.09**
τ-statistics	0.35**	0.06	0.16	0.15	
Average (median) observation	1999 (1572)	1795 (1125)	1732 (1124)	2099 (1434)	

Panel B: Additional stratification by size

Size		−1% < ΔDY < −0.11%	−0.11% < ΔDY < −0.04%	0.04% < ΔDY < 0.11%	0.11% < ΔDY < 1%	Z statistic
Large	N_c	104	84	75	N/A	1.47*
	τ-statistics	0.4**	0.1	−0.02	N/A	
	Average (median)	578 (424)	701 (452)	599 (379)	543 (462)	
Medium	N_c	107	82	83	83	1.86**
	τ-statistics	0.36***	0.07	0.08	0.08	
	Average (median)	685 (592)	609 (448)	587 (432)	692 (476)	
Small	N_c	96	75	98	91	2.05**
	τ-statistics	0.25*	−0.02	0.28**	0.19	
	Average (median)	736 (452)	484 (311)	546 (339)	865 (450)	
All sizes						2.99***

Panel C: Additional stratification by yield

Yield		−1% < ΔDY < −0.11%	−0.11% < ΔDY < −0.04%	0.04% < ΔDY < 0.11%	0.11% < ΔDY < 1%	Z statistic
High	N_c	N/A	44	107	91	0.55
	τ-statistics	N/A	−0.42	0.4***	0.19	
	Average (median)	787 (507)	453 (281)	536 (320)	1190 (583)	
Medium	N_c	106	71	93	92	2.12**
	τ-statistics	0.39**	−0.07	0.22	0.2	
	Average (median)	776 (565)	649 (591)	660 (492)	744 (531)	
Low	N_c	102	92	78	N/A	1.85**
	τ-statistics	0.33**	0.2	0.02	N/A	
	Average (median)	414 (356)	686 (533)	528 (391)	230 (183)	
All yields						2.65***

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the one-sided test. Significance levels for the τ-statistics, τ^i=((N_c^i−N_D^i)/N), are presented in Hollander and Wolfe (1973), p. 384. The average number of observations is the total in each τ-test divided by 18; the median is the number in the portfolio for which there are an equal number of portfolios with more and with less observations.

costly signal, as suggested by dividend signaling theory, leads to weaker monotonicity findings, rather than stronger monotonicity findings. Simply put, the "bang-for-the-buck" variable associated with the signaling prediction is less correlated with regime than is excess return.

Table 6
Tests for monotonicity between portfolio excess returns and tax regime

Panel A: Minimal stratification

	All changes in dividend yield	All decreases in yield	All increases in yield
N_c	95	79	91
τ-statistics	0.24*	0.03	0.19
z-statistics	1.4*	0.19	1.1
Observations			
In average portfolio	9602	4730	4872
In median portfolio	5059	2694	2666

Panel B: Additional stratification by change in dividend yield

	− 1% < ΔDY < − 0.2%	− 0.2% < ΔDY < − 0.11%	− 0.11% < ΔDY < − 0.07%	− 0.07% < ΔDY < − 0.04%	− 0.04% < ΔDY < − 0.015%	− 0.015% < ΔDY < 0	0% < ΔDY < 0.015%	0.015% < ΔDY < 0.04%	0.04% < ΔDY < 0.07%	0.07% < ΔDY < 0.11%	0.11% < ΔDY < 0.2%	0.2% < ΔDY < 1%
N_c	94	97	96	73	73	80	77	75	102	90	87	92
τ-statistics	0.23*	0.27*	0.25*	− 0.05	− 0.05	0.05	0.01	− 0.02	0.33**	0.18	0.14	0.2
z-statistics	1.33*	1.55*	1.48*	− 0.27	− 0.27	0.27	0.04	− 0.11	1.93**	1.02	0.8	1.17
Average observation	793	894	751	800	853	609	583	827	783	760	966	926
Median observation	534	549	451	467	440	299	292	397	426	412	599	510
Aggregate Z statistics												
ΔDY < − 0.04%	2.05**											
ΔDY > 0.04%									2.46**			
All together	2.58**											

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the one-sided test. Significance levels for Kendall's τ-statistics, $\tau^{i}=((N_c^i-N_D^i)/(N_c^i-N_D^i))$, are presented in Hollander and Wolfe (1973), p. 384.
The average number of observations is the total in each ô-test divided by 18; the median is the number in the portfolio for which there are an equal number of portfolios with more and with less observations.

Table 7
Monotonicity between excess returns and tax regime: alternative signals and additional stratification

Panel A: Results using the change in dividend level as the signal

	$-1\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 1\%$	Z statistic
N_c	89	N/A	97	98	2.38***
τ-statistics	0.16	N/A	0.27*	0.28*	
Total (median) observation	278 (191)	168 (86)	340 (161)	561 (252)	

Panel B: Results using change in dividend yield as the signal

Stratification into four categories for change in dividend yield

	$-1\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 1\%$	Z statistic
N_c	89	83	96*	94*	2.13**
τ-statistics	0.16	0.08	0.24*	0.23*	
Total (median) observation	1686 (1045)	1551 (918)	1543 (864)	1891 (1070)	

Additional stratification by firm size

Size		$-1\% < \Delta DY < -0.1\%$	$-0.1\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.1\%$	$0.1\% < \Delta DY < 1\%$	Z statistic
Large	N_c	100	85	84	N/A	1.75**
	τ-statistics	0.31**	0.11	0.1	N/A	
	Average (median)	480 (320)	611 (315)	536 (290)	443 (282)	
Medium	N_c	88	82	89	89	1.14
	τ-statistics	0.15	0.07	0.16	0.16	
	Average (median)	578 (384)	521 (368)	519 (307)	621 (353)	
Small	N_c	91	73	93	85	1.78**
	τ-statistics	0.19	−0.05	0.22	0.11	
	Average (median)	629 (392)	419 (249)	489 (285)	787 (360)	
All sizes						2.68***

Additional stratification by dividend yield

Yield		$-1\% < \Delta DY < -0.11\%$	$-0.11\% < \Delta DY < -0.04\%$	$0.04\% < \Delta DY < 0.11\%$	$0.11\% < \Delta DY < 1\%$	Z statistic
High	N_c	N/A	59	95	94	1.93**
	τ-statistics	N/A	−0.23	0.24*	0.23*	
	Average (median)	322 (172)	628 (320)	439 (690)	1006 (556)	
Medium	N_c	92	74	100	92	1.97**
	τ-statistics	0.2	−0.03	0.31**	0.2	
	Average (median)	645 (495)	531 (449)	566 (340)	623 (402)	
Low	N_c	104	96	91	N/A	2.69***
	τ-statistics	0.36**	0.25*	0.19	N/A	
	Average (median)	397(341)	652 (405)	530 (331)	211 (156)	
All yields						3.16***

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the one-sided test. Significance levels for Kendall's τ-statistics are presented in Hollander and Wolfe (1973), p. 384.
The average number of observations is the total in each ô-test divided by 18; the median is the number in the portfolio for which there are an equal number of portfolios with more and with less observations.

5. Conclusions

Billions of dollars are paid out each year by firms in the form of cash dividends. The puzzle for researchers is why, given the high associated tax costs of dividends relative to capital gains, are dividends issued? This paper tests whether the underlying explanation for dividends could be signaling based. We focus on a monotonicity prediction that is a feature of dividend signaling models: a positive rank order correlation between the relative tax disadvantage of dividends and the amount of information revealed by a particular dividend signal.

We order 18 distinct tax regimes according to the tax disadvantage of dividend income relative to capital gains. Using distribution-free tests, we cannot reject the hypothesis of independence between tax regime and the excess returns associated with a given change in dividend signal (yield or level). Our findings indicate that the information content in dividends is not positively related to the marginal cost of dividends in the manner implied by dividend signaling theory. In addition, we find that the excess return, rather than the bang-for-the-buck (excess return per dollar of dividends) as predicted by signaling models, is more strongly related to the tax regime. Overall, therefore, our findings suggest that signaling concerns do not explain why dividends are issued. Thus, the fundamental questions—why do firms issue dividends? why does the market respond favorably to dividend increases?—appear to remain unanswered.

Plausibly, our findings reflect non-signaling explanations for dividends. For example, suppose that firms pay dividends to reduce free cash on hand (say to reduce the incentives of managers to misallocate), and that the dividend is based on a marginal benefit-marginal cost calculation that reflects taxes but does NOT reflect signaling. A firm may increase its dividend to control opportunistic spending both if there is an increase in cash on hand relative to investment opportunities (good news), or a decrease in investment opportunities (bad news), and the nature of information revealed can vary with aggregate economic conditions (possibly spuriously correlated with dividend tax regimes) or the firm's stage in its lifecycle (correlated with dividend level). These possibilities may well underlie the spotty evidence that we uncover.

A spurious correlation between tax rates and market conditions may also underlie the stronger relationship between tax regime and excess return. For example, suppose that price is more sensitive to dividends in periods of active takeover markets (e.g., because potential targets use dividends to disburse idle cash as a takeover defense, and higher than expected dividends convey that the firm is "in play"). If takeover markets were more active in periods of higher dividend taxation, as in the late 1980s and early 1990s, a spurious relation is produced. The sensitivity of ER may then reflect the probability of takeover more than the dollar amount of the dividend, and produce a (ER,t) relationship that is stronger than the (BFB,t) relationship.

Future work is required to bear out which, if any, of these explanations is correct. The above discussion suggests that our nonparametric methodology may not be the best way to address these possibilities. Our methodology is best suited to testing theories that (a) predict monotonicity without predicting functional form, (b) require only a limited number of controls, and (c) relate to data that can be ordered into sufficiently many regimes. Signaling theories fit well, provided that a sufficient number of signaling cost regimes can be identified (as in the dividend-tax-based theory). For example, our methodology might also be employed to test signaling theories of capital structure, because the cost of debt reflects the relative taxation of debt and equity, or to test advertising as a signal of quality.

Acknowledgements

The first author is grateful to the SSHRC for financial support. We appreciate the comments and suggestions of Douglas Bernheim, George Deltas, Bart Lipman, Jeff Strnad and the seminar participants at Caltech and Queen's University. We would also like to thank Nick Favron for excellent research assistance. All errors are ours.

References

Aharony, J., Swary, I., 1980. Quarterly dividend and earnings announcements and stockholder returns: an empirical analysis. Journal of Finance 35, 1–12.

Allen, F., Michaely, R., 1995. Dividend policy, Chap. 25. In: Jarrow, R., et al. (Ed.), Handbooks in OR and MS, vol. 9. Elsevier, pp. 793–837.

Allen, F., Michaely, R., 2002. Payout policy. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), North-Holland Handbook of Economics. Elsevier. In press.

Amihud, Y., Murgia, M., 1997. Dividends, taxes, and signaling: evidence from Germany. Journal of Finance 52 (1), 397–408.

Asquith, P., Mullins, D., 1983. The impact of initiating dividend payments on Shareholders' wealth. Journal of Business 56, 77–96.

Avery, R., Elliehausen, G., 1986. Financial characteristic of high income families. Federal Reserve Bulletin 72, 172–182.

Bajaj, M., Vijh, A., 1990. Dividend clienteles and the information content of dividend changes. Journal of Financial Economics 26, 193–219.

Bernhardt, D., Lee, T., 2001. Dividend signalling: fact or fiction? Mimeo, University of Illinois.

Bernheim, B.D., Wantz, A., 1995. A tax-based test of the dividend signaling hypothesis. American Economic Review 85, 532–551.

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

Brennan, M., 1970. Taxes, market valuation, and corporate financial policy. National Tax Journal 23, 417–427.

Brennan, M., Thakor, A., 1990. Shareholder preferences and dividend policy. Journal of Finance, 993–1018.

Brown, S.J., Warner, J.B., 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.

Christie, W., 1994. Are dividend omissions truly the cruelest cut of all? Journal of Financial Economics 29, 459–480.

Eades, K., Hess, P., Kim, H., 1984. On interpreting security returns during the ex-dividend period. Journal of Financial Economics 13, 3–34.

Grullon, G., Michaely, R., 2001. Asymmetric information, agency conflicts and the impact of taxation on the market reaction to dividend changes. Working paper. Cornell University, Ithaca, NY.

Hollander, M., Wolfe, D., 1973. Nonparametric statistical methods. Wiley, New York.

Jensen, M., 1986. Agency costs of free cash flow, corporate finance, and takeovers. American Economic Review 76, 329–423.

John, K., Williams, J., 1985. Dividends, dilution, and taxes: a signaling equilibrium. Journal of Finance 40, 1053–1070.

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., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: theory and empirical evidence. Journal of Financial Economics 7, 163–195.

Pechman, J., 1987. Federal tax policy. Brookings Institution, Washington D.C.

Pettit, R., 1972. Dividend announcements, security performance, and capital market efficiency. Journal of Finance 27, 993–1001.

Ross, S., 1977. The determination of financial structure: the incentive-signaling approach. Bell Journal of Economics 1, 23–40.

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