JOHN G. POWELL

JING SHI*

TOM SMITH

ROBERT E. WHALEY

SPURIOUS REGRESSION, SPURIOUS CORRELATION, AND DIVIDEND YIELD RETURN PREDICTABILITY

Abstract

Dividend yield return predictability has been nominated as one of the "new facts in finance", thus justifying widespread use of the dividend yield as a standard control variable in financial time series analysis. This paper provides strong evidence that return predictability of the dividend yield is a spurious result, even when the dividend yield definition is broadened to include share repurchases. The paper demonstrates how a spurious regression problem that is due to dividend persistence is compounded by a spurious correlation problem when the dependent and independent variables in return predictability regressions are ratios composed of common component variables, and it utilizes an analytical formula as well as a simulation procedure to take account of these problems. The paper's results indicate that that the spurious correlation problem (in interaction with spurious regression) is the major contributing factor to incorrect inference in dividend yield return predictability studies, thus implying that extreme care should be taken when using ratios as predictor or explanatory variables in time series regression. The paper finds that standard dividend behaviour explanatory models are also affected by the spurious regression problem, and introduces a reformulated Lintner first difference dividend behaviour model that is not subject to spurious regression.

This version: October, 2005

*Corresponding author. School of Finance and Applied Statistics, The Australian National University, Canberra ACT 0200, Australia Telephone: +61 2 61254864, Fax: +61 2 61250087, Email: Jing.Shi@anu.edu.au. Powell is from Department of Finance Banking and Property, Massey University, Palmerston North, New Zealand; Smith and Shi are from the School of Finance and Applied Statistics, The Australian National University, Canberra, Australia; Whaley is from Fuqua School of Business, Duke University, USA. We have benefited from and are grateful for comments by seminar participants at Massey University, Jiangxi University of Finance and Economics, Guangxi University, the 13th Annual Conference on Pacific Basin Finance Economics and Accounting, and the 2005 Accounting and Finance Association of Australia and New Zealand.

SPURIOUS REGRESSION, SPURIOUS CORRELATION, AND DIVIDEND YIELD RETURN PREDICTABILITY

Evidence of dividend yield return predictability has been presented so widely and consistently that return predictability is sometimes considered to be one of the "new facts in finance" (Cochrane, 1999), thus justifying widespread use of the dividend yield as a standard business cycle control variable in financial time series analysis. This paper provides strong evidence that return predictability of the dividend yield is a spurious result, regardless of whether the definition of the dividend yield variable used to predict returns is updated by broadening it to include cash distributed to investors via share repurchases. The paper demonstrates how a spurious regression problem that is due to regression variable persistence is compounded by a spurious correlation problem because the dependent and independent variables in return predictability regression equations are ratios constructed from common component variables. Standard dividend behaviour explanatory models are also shown to be strongly affected by spurious regression that is compounded by the regression model dependent and independent variables being ratios with common component variables. A simulation procedure is utilized to take account of the spurious regression and spurious correlation problems, and the paper derives a reformulated dividend explanatory model to indicate how this econometric problem can be avoided.

A first hint that return predictability of the dividend yield is spurious can be obtained by examining how the dividend yield predicts returns relative to other persistent variables. A standard approach to test for dividend yield return predictability is to regress the time $t+1$ stock index rate of return (r_{t+1}) against the time t dividend yield (D_t / P_t) using the regression equation

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_1\left[\frac{D_t}{P_t}\right] + e_{t+1}, \tag{1}$$

where D_t is the level of real annual dividends during the twelve months preceding time t and P_t is the real stock index level at time t. The adjusted R^2 statistic (hereafter, $\bar{R}^2$) for regression equation (1) for the annual CRSP Value-weighted NYSE Index during the time period 1927 to 2004 is close to 2%, for instance, and the results are usually interpreted in relation to a correlation between the current level of dividends and subsequent returns (see, e.g., Campbell, Lo and MacKinlay, 1997).[1] It can be noted, however, that if variation in dividends plays an important role in the return predictability relationship then substituting a constant level of dividends (c) for the dividend term (D_t) in the numerator of the dividend yield variable (D_t / P_t) in regression equation (1) should reduce or eliminate return predictability in the following altered regression equation

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_2\left[\frac{c}{P_t}\right] + e_{t+1}, \tag{2}$$

where c the is the unconditional average of the dividend level during the sample period. Surprizingly, the $\bar{R}^2$ for equation (2) actually rises to above 3% for the value-weighted NYSE Index. This unexpected result is a first indication that a spurious effect might be occurring and is a very strong hint of a spurious regression problem.

To explore this possibility, the persistence properties of the independent variables in regression equations (1) and (2) are examined. They are found to be highly persistent. This persistence combines with return autocorrelation to create spurious return

[1] Lewellen (2004) points out that dividend yield return predictability could also be due to temporary mispricing, an interpretation that will be discussed in Section 3.

predictability because the regression error term inherits autocorrelation from the dependent variable, thus causing standard error estimates to be downward biased and a significant overall relationship to appear when none actually exists (see Ferson, Sarkissian and Simin, 2003a; Foster, Smith and Whaley, 1997).[2] Simulations indicate that the spurious effect of regressing returns against a persistent dividend yield variable is strongly reinforced by a spurious correlation that is due to returns on the left hand side of equation (1) and the dividend yield on the right hand side being ratios that are constructed from the same underlying variables (the share index level and the dividend level; see also Pearson, 1897; Kronmal, 1993; Kim, 1999).

The simulation procedure is adapted to account for the interaction between the spurious correlation and spurious regression problems by simulating uncorrelated share index (P_t) and dividend (D_t) series that have the same auto-correlation properties as the actual data. The simulated share index and dividend series are then used to construct return series ($r_{t+1} \equiv (P_{t+1} + D_{t+1} - P_t)/P_t$) and dividend yield series (D_t / P_t) that are potentially spuriously correlated (rather than simulating returns and dividend yields directly), and the constructed return series are regressed against the constructed dividend series to provide a simulated cut-off $\bar{R}^2$ benchmark. The influence of spurious correlation on dividend yield return predictability regression results is also demonstrated using an analytical formula derived in Kim (1999) that provides a spurious correlation benchmark when correlating ratios with a common divisor. The paper reveals that spurious correlation due to common regression variable components therefore combines with the

[2] Campbell, Lo and MacKinlay (1997) demonstrate that stock index returns are autocorrelated, for instance, and Ferson, Sarkissian and Simin (2003a) point out that even if returns are not highly persistent, underlying expected returns are persistent, so there is a risk of spurious regression when testing return predictability.

persistence properties of the dividend yield to account for the apparent dividend yield return predictability, rather than (as widely argued) any property of the dividend level that is related to risk, return or mispricing. Evidence that dividend yield return predictability is spurious is also shown to be just as strong even when the definition of dividends is broadened to include distributions to investors via share repurchases, a component of shareholder cash distributions that has become increasingly important in recent decades (see also Boudoukh *et al*, 2004; Grullon and Michaely, 2002).

Given the strong persistence properties of dividends, it is not surprising to find that standard dividend behaviour explanatory models such as Lintner (1956) are also affected by the spurious regression problem, since the dependent and independent variables in these models are all very persistent (see also Ferson, Sarkissian and Simin, 2003a). To correct this problem, the paper demonstrates how the Lintner behavioural model of dividends can be reformulated and extended using Marsh and Merton (1987) so that the model is not subject to spurious regression.

The structure of the paper is as follows. Section I contains a review of the dividend yield return predictability and dividend behaviour literature. The spurious regression literature is then reviewed to indicate the potential problems that might be present with the return predictability and dividend behaviour literature. Section II shows that something "funny" is going on in the return predictability literature by revealing that other highly persistent explanatory variables can be used to improve upon standard dividend yield return predictability results, and then uses simulation procedures to show that all the results are spurious. The contribution of spurious correlation to the dividend yield return predictability spurious regression problem is also outlined. Section III demonstrates how standard dividend behaviour regression equations are also subject to spurious regression. Section IV corrects this problem by reformulating the Lintner (1956) dividend model. Section V provides a brief conclusion to the paper.

I. LITERATURE REVIEW

A. Dividend Yield Return Predictability

The literature on stock return predictability is extensive, with the dividend yield arguably being the best-known of the many variables that are found to have forecasting power for stock returns. The dividend yield is usually measured as the ratio of the previous year's dividend payments to the current share price index level. This process removes the influence of strong seasonal effects in dividend payments, but it also artificially raises autocorrelation in monthly and quarterly overlapping dividend yield series.

Many studies find that dividend yields predict the cross-sectional and times series variation of stock returns (see, e.g., Fama and French, 1988; Harvey, 1989; Ferson and Harvey, 1991; Whitelaw, 1994; Pesaran and Timmermann, 1995; Pontiff and Schall, 1998; Bossaerts and Hillion, 1999; Cremers, 2002; Boudoukh *et al*, 2004; Lewellen, 2004; and Touros, Valkanov, and Yan, 2004).[3] Table I summarizes a number of the important return predictability studies that focus exclusively or almost exclusively on the dividend yield. Dividend yield return predictability had been documented so frequently that it was recently nominated as one of the "new facts in finance" (Cochrane, 1999),

[3] Goyal and Wech (2003, p. 639) state "This empirical regularity - that dividend ratios seem to predict equity return – ranks amongst the most important findings of academic finance, and it shows no signs of subsiding. For instance, a citation search lists more than 200 published articles citing the Fama and French (1988) article alone."

with the dividend yield now tending to be included whenever business cycle control variables are employed in financial time series analysis.[4]

[Table I about here]

A number of recent studies have also responded to challenges to findings of dividend yield return predictability. Lewellen (2004) finds evidence in favour of return predictability using the natural logarithm of the dividend yield once a small sample bias correction is improved to more precisely account for a coefficient estimate bias that is induced by a strong correlation between the dividend yield slope coefficient and the dividend yield's autocorrelation. Campbell and Yogo (2005) find that the Lewellen (2004) bias-corrected test has poor power relative to Bonferroni inequality probability confidence interval tests when persistence does not equal unity, but never-the-less find evidence in support of dividend yield return predictability using a Uniformly Most Powerful (UMP) test. The Campbell and Yogo (2003) UMP test subtracts from the dividend yield dependent variable the innovations in the yield that are correlated with returns in order to reduce the independent variable's noise and consequently to increase the power of the return predictability regression coefficient significance test. Boudoukh *et al* (2004) demonstrate that a broadly defined dividend yield variable that includes share repurchases still predicts returns even in recent decades when share repurchases have become important, whereas the significance of the traditional dividend yield for return predictability has declined markedly in recent decades (see also Robertson and Wright, 2003).

[4] A search of the Journal of Finance and the Journal of Financial Economics indicates that most recent financial time series studies that employ business cycle control variables include the dividend yield as one of the control variables.

Only a few studies indicate that there is not a strong statistical relationship between dividend yields and stock returns, including Black and Scholes (1974) and Goetzmann and Jorion (1993). Bossaerts and Hillion (1999) use a number of statistical model selection criteria to examine the predictability of stock returns using dividend yields and find in-sample predictability but no out-of-sample forecasting power. Goyal and Welch (2003) find the predictive power of the dividend yield is present in pre- but not post-1990 data (see also Lettau and Ludvigson, 2005). Valkanov (2003) tests the dividend yield's ability to forecast returns based upon restrictions implied by the log-linearized dynamic Gordon growth model, and finds no evidence of return predictability at 1 to 6 year return horizons. Stambaugh (1999) finds that return predictability disappears when the bias induced by a correlation between the regression error and innovations in the autoccorelated dividend yield regressor is accounted for using Bayesian autocorellation priors. Nelson and Kim (1993) use simulations to show that this bias has an important influence on statistical inference in samples that include pre- World War II returns. Ferson, Sarkissian and Simin (2003a) indicate that the dividend yield's predictive power for monthly returns is questionable when account is taken of the spurious regression problem combined with data mining.

Dividend yields have also been used to predict long horizon returns (e.g. Fama and French, 1988; Campbell, Lo and MacKinlay, 1997). Fama and French (1988) estimate regressions of returns on the lagged dividend yield using post-war NYSE index data for return horizons from one month to four years. They observe that the dividend yield explains a significant proportion of multiple year returns, and the explanatory power of the dividend yield increases with the return horizon. Hodrick (1992) uses three alternative methods of conducting inference and measurement for long-horizon forecasting and finds that changes in dividend yields forecast significantly persistent changes in expected stock returns.

Theoretical explanations for the predictive power of the dividend yield have been developed using the dividend discount model (see, e.g., Campbell and Shiller, 1988a and 1988b; Fama and French, 1988; Donaldson and Kamstra, 1996; and Campbell, Lo and MacKinlay, 1997; Lamont, 1998; Bansal and Yaron, 2004). Campbell and Shiller (1988a, 1988b) and Fama and French (1988) argue that the dividend discount model implies that a high current dividend level relative to the share price index level predicts some combination of either higher expected future returns or lower future dividends. A high current dividend yield therefore forecasts higher future returns if expected future dividends are held constant (e.g., under the perhaps strong assumption of "all else being equal"), thus providing a theoretical basis for dividend yield return predictability.

Although the dividend yield return predictability literature is vast and theoretical explanations for predictability are also well-known, this literature has not been directly connected to the equally well-known literature on behavioural explanations of dividends (see, e.g., Lintner, 1956; Marsh and Merton, 1987). A review of the dividend behaviour literature provides insights as to why dividend yield return predictability regression models are likely to be strongly affected by dividend persistence.

B. Dividend Behaviour

The starting point for a discussion of behavioural models of dividends is the Lintner (1956) speed of adjustment model (see, eg., Marsh and Merton, 1987). Lintner

(1956) argues that corporate managers feel that they have a duty to pay out a proportion of earnings to shareholders, but are reluctant to increase dividends too quickly in reaction to an increase in earnings in case the earnings increase turns out to be temporary and the dividend increase subsequently has to be reversed. The Lintner (1956) model of dividends therefore implies that the current change in dividends is equal to a target dividend payout minus last period's dividend (D_{t-1}) times a speed of adjustment factor, plus a constant. The target payout is equal to the current level of trailing annual earnings (E_t) times a long run payout ratio target. The current dividend level (D_t) does not adjust instantaneously to the target payout level, thus avoiding the situation where an increase in earnings is only temporary and would have to be reversed in future years. Lintner's dividend model (equation (2) in Lintner, 1956) therefore states that the time t dividend is given by

$$D_t = q_0 + q_1 E_t + q_2 D_{t-1} + e_t \,, \qquad (3)$$

where t is in years, e_t is the error term, $q_1/(1-q_2)$ is the long-run target payout ratio, and $(1-q_2)$ is a speed-of adjustment factor whose value will be closer to zero the more slowly that firms adjust their dividends to their long-run target level. Lintner (1956) reports an R^2 in excess of 90% when testing this model (regression equation (3)) using aggregate dividends and earnings data. Grullon and Michaely (2002) use the Lintner (1956) model to generate dividend forecast errors for companies in order to test whether share repurchases are substitutes for dividends, thus updating the Lintner (1956) results for individual companies, but they do not test the Lintner (1956) model using aggregate dividends (see also Fama, 1974).[5]

It can be noted that current time t earnings have not yet been observed when the current dividend level is decided upon and declared, since companies report their earnings after the end of the quarter whereas dividends are decided upon and announced prior to their payment each quarter. This problem can be eliminated in the Lintner model by lagging earnings by either a quarter or a year so that the time t dividend level choice is modelled only in relation to information that is observable at time t:

$$D_t = q_0 + q_1 E_{t-1} + q_2 D_{t-1} + e_t . \tag{4}$$

Underlying the Lintner model is the idea that dividends would not be adjusted to changes in earnings that are only temporary, so Marsh and Merton (1987) introduce the concept of permanent earnings into dividend behaviour models and argue that dividends will only be determined in relation to permanent earnings. They further argue that, in an efficient market, the current share price index level is equal to the present value of all future permanent earnings. The rate of change of share prices can then be shown to be equal to the rate of change of expected permanent earnings when it is assumed that the long-run discount rate is constant, thus implying a log linear relationship between dividend and price changes that leads to regression equation (11) in Marsh and Merton (1987):

$$\log\left(\frac{D_{t+1}}{D_t}\right) + \frac{D_t}{P_{t-1}} = y_0 + y_1 \log\left(\frac{P_t + D_t}{P_{t-1}}\right) + y_2 \log\left(\frac{D_t}{P_{t-1}}\right) + e_{t+1} . \tag{5}$$

[5] The average adjusted R^2 they obtain of 45.7% is not directly comparable with the Lintner (1956) results because Grullon and Michaely (2002) estimate the Lintner model using changes in dividends (equation (1) in Lintner, 1956), not the dividend level regression model that Lintner (1956) tests (see equation (3) above), but the parameter estimates obtained are similar to the Lintner (1956) results.

The Marsh and Merton (1987) model explains roughly half as much of the variation in aggregate dividends as does Lintner's equation (3) above (see Lintner, 1956; Marsh and Merton, 1987; as well as the Results section).

An alternative explanation of aggregate dividend behaviour is provided by Shiller's (1983) trend-autoregressive model.[6] The model implies that deviations in aggregate dividends follow a trend-autoregressive process whereby half of the deviation in dividends from trend disappears within three years, with the underlying trend being explained by ongoing (not anticipated) earnings growth. Changes in current dividends are therefore determined by a time trend as well as past deviations from trend in Shiller's model, rather than anticipated earnings growth, but empirical tests indicate that the trend-autoregressive model provides (at most) half as much explanatory power as the Marsh and Merton (1987) model (see Marsh and Merton, 1987).

C. Spurious Regression

Behavioural models of dividends share in common the use of lagged dividend terms to explain current dividend choices, a feature which implies that dividends are highly persistent. The recent literature on spurious regression points out that the use of persistent independent variables can lead to spurious regression results when the dependent variable is also at least partially persistent, since error terms in the regression equation inherit autocorrelation from the persistent dependent variable (Ferson, Sarkissian, and Simin, 2003a). This autocorrelation in the error term leads to biased standard error estimates and can therefore indicate a significant overall relationship when

[6] For an extensive theoretical and empirical review of dividend policy and dividend behaviour at the individual firm level, see Allen, Bernardo and Welch (2000).

none exists, especially when data sets are mined for potentially significant explanatory variables (Ferson, Sarkissian, and Simin, 2003a; Foster, Smith and Whaley, 1997). Behavioural models of dividends are therefore likely candidates for spurious regression due to their strong persistence properties. It is important to note that the spurious regression problem does not disappear as the sample size is enlarged, unlike the persistent regressor coefficient estimate bias outlined by Stambaugh (1999) and Nelson and Kim (1993), so it is a potential problem for dividend behaviour studies regardless of the study sample that is employed (Ferson, Sarkissian, and Simin, 2003a).

Return predictability regressions that use the dividend yield as a predictor variable are also likely candidates for spurious regression because the numerator of the independent variable, the dividend level, is highly persistent and will consequently make the dividend yield variable highly persistent as well. It is therefore likely that persistence might be responsible for the apparent return predictability of the dividend yield, rather than any property of dividends related to risk, return or mispricing, and in this situation other similarly persistent variables might also possess apparent return predictability, as will be illustrated below.

II. DIVIDEND YIELD RETURN PREDICTABILITY

A starting point for examining whether dividend yield return predictability is a spurious result is to reproduce standard dividend yield return predictability results and compare them to the return predictability results that can be obtained using other persistent explanatory variables. The total return and capital return stock indices used in the study are the annual CRSP Value-weighted index and the annual Standard and Poor

500 Index for the time period 1926 to 2004. All stock index data are obtained from CRSP and all series are converted from nominal to real values by dividing by the Consumer Price Index (CPI) obtained from Robert Shiller's web page. The dividend yield on the index at time t equals the level of real dividends during the twelve months preceding time t (D_t) divided by the real stock index level at time t (P_t).[7] Panel A of Table 2 reproduces standard dividend yield return predictability results by regressing annual time $t+1$ rates of return against a constant as well as the annual time t dividend yield (see, e.g., Fama and French, 1988):

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_1 \left[\frac{D_t}{P_t} \right] + e_{t+1}. \tag{1}$$

The Fama and French (1988) results and the results of other studies are replicated in Panel A of Table II, with the regression $\overline{R}^2$ equalling 1.95% for the CRSP Value-weighted Index and 2.05% for the S&P 500 Index.[8]

[Table II about here]

The return predictability results documented in Panel A of Table II are often explained in terms of a theoretical relationship between current dividend yields and subsequent returns that is derived from the dividend discount model (see, e.g., Campbell and Shiller, 1988a and 1988b; Fama and French, 1988). It can be noted that this theoretical justification for return predictability relies upon a high current dividend level relative to the share index level predicting higher future returns, so substituting a constant

---

[7] Numerous dividend yield return predictability studies also use monthly or quarterly returns with overlapping dividend yield observations. The use of overlapping observations introduces excess dividend yield autocorrelation. This well-known overlapping observation problem is avoided with the use of annual data.

[8] Regressions are estimated using Ordinary Least Squares and t-statistics are estimated using Newey-West (1987) standard errors.

level of dividends (c) for the dividend term (D_t) in the numerator of the dividend yield variable (D_t / P_t) in regression equation (1) should reduce or eliminate return predictability in the following altered regression equation:

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_2 \left[\frac{c}{P_t} \right] + e_{t+1}, \tag{2}$$

where c is the unconditional average of the dividend level during the sample period. The results for regression equation (2) are outlined in Panel B of Table II. Surprizingly, the $\bar{R}^2$ actually rises to above 3.2% for the S&P 500 Index and 3.3% for the Value-weighted CRSP Index. The results for this dividend yield from a constant dividend variable (c / P_t) are therefore so strong that they provide an indication that something "funny" is going on in previous dividend yield return predictability studies since, interestingly, the dividend yield from a constant dividend variable (c / P_t) has higher explanatory power than does the standard dividend yield variable (D_t / P_t). This indicates that dividends are unlikely to be providing the explanatory power in return predictability regressions.

The Table II results instead suggest that the observed empirical relationship between subsequent returns and the dividend yield might be spurious. To investigate this possibility, the persistence properties of the dividend, dividend yield, price index, dividend yield from a constant dividend, and return series are investigated in Table III. Table III confirms that the real dividend series are highly persistent, as are the dividend yield series. The dividend yield from a constant dividend series are even more persistent than the dividend yield series, thus indicating that persistence might have an important influence on return predictability, since predictability increases when more persistent variables are used as predictor variables in Table II. Table III indicates that returns are sufficiently autocorrelated to suggest that regressing subsequent rates of return against highly persistent dividend yield explanatory variables could lead to spurious regression.

[Table III about here]

To follow up this possibility, a simulation procedure is utilized that provides the cut-off $\bar{R}^2$ that would be obtained from a regression for which the dependent and independent variables are uncorrelated but have the same autocorrelation properties as the actual data, thus taking account of the potential for spurious regression (see also Foster, Smith and Whaley, 1997; Ferson, Sarkissian and Simin, 2003a). The simulation procedure is closely related to the procedure used in Nelson and Kim (1993), with the important difference being that dependent variable auto-correlation is directly incorporated, thus moving the analysis from the small sample persistent regressor coefficient estimate bias examined by Nelson and Kim (1993) to the spurious regression problem examined by Ferson, Sarkissian and Simin (2003a).

To obtain the cut-off $\bar{R}^2$ using simulation, the moments and the serial correlation properties of the regression variables are first estimated for each data series, as described in the Appendix. Uncorrelated dependent and independent variables with the same serial correlation properties and sample moments are then simulated for a time period equal to the sample length (1927 to 2004), and a regression is run on these simulated series. The process is repeated 10,000 times, and the $\bar{R}^2$**s** are recorded for each regression and ranked from lowest to highest. The 95th percentile $\bar{R}^2$ is then reported as the 5% cut-off $\bar{R}^2$ and is compared to the actual $\bar{R}^2$ obtained using the original data to assess the overall significance of the estimated regression relationship (see Foster, Smith and Whaley, 1997; Ferson, Sarkissian and Simin, 2003a).

A second set of cut-off $\bar{R}^2$ statistics is also obtained by simulation to take account of the potential effect of data mining since, as was mentioned earlier, the dividend yield is only one of many variables that have been explored for evidence of return predictability (see Cremers, 2002, as well as Lo and MacKinlay (1990); Foster, Smith, and Whaley (1997); Sullivan, Timmermann and White (1999) and (2001); Ferson, Sarkissian and Simin, 2003a). Data mining for explanatory variables reinforces the spurious regression problem, since highly persistent explanatory variables are more likely to display apparent significance. The second set of cut-off $\bar{R}^2$ statistics use Bonferonni correction intervals to take account of the number of series that are examined in the search for potentially significant relations, with a conservative adjustment factor of five being used, thus representing the assumption that at least five explanatory variables have been used in the search for stock market index return predictability (Foster, Smith and Whaley, 1997). This modification is equivalent in an operational sense to requiring a one percent level of overall significance rather than a five percent level due to the number of explanatory variables that have been searched for return predictability.

The eighth and ninth columns of Table II report the cut-off $\bar{R}^2$ and the cut-off $\bar{R}^2$ with data mining that are obtained using the simulation procedure so that it can be compared with the $\bar{R}^2$ reported in the seventh column of Table II. For instance, the cut-off $\bar{R}^2$ (cut-off $\bar{R}^2$ with data mining) reported for the Value-weighted CRSP Index in Panel A of Table II indicate that a $\bar{R}^2$ of 3.95% (7.4%) would be expected to be obtained by regressing subsequent returns against uncorrelated independent variables that are equally as persistent as the dividend yield variable (D_t / P_t), whereas a $\bar{R}^2$ of only 1.95% is actually obtained. The $\bar{R}^2$ of 1.95% does not exceed the critical cut-off $\bar{R}^2$ levels, thus implying that the Value-weighted dividend yield return predictability regression $\bar{R}^2$ is insignificant (see Foster, Smith and Whaley, 1997; Ferson, Sarkissian and Simin, 2003a). The cut-off $\bar{R}^2$ levels reported in Table II therefore indicate that all of the dividend yield return predictability regression results are due to spurious regression since the reported $\bar{R}^2$s never exceed the cut-off $\bar{R}^2$s (see the eighth column and ninth columns of Table II).

The simulation procedure treats the variables in regression equation (1),

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_1 \left[\frac{D_t}{P_t} \right] + e_{t+1}, \tag{1}$$

as if they are independent. This independence assumption is not appropriate since returns on the left hand side of equation (1) and the dividend yield on the right hand side both come from the same underlying variables (the share index level P and the dividend level D). The simulation procedure is therefore modified to recognize the dependency of both the return and dividend yield variables on the share index and dividend levels (see the Appendix for details of the modified simulation procedure). Rather than simulating the return series r_{t+1} as the dependent variable using the properties of the return series and simulating an uncorrelated dividend yield series (D_t / P_t) using the properties of the dividend yield series to calculate the cut-off $\bar{R}^2$, uncorrelated dividend (D_t) and share index (P_t) series are instead simulated using the estimated properties of these series. Dividend yield (D_t / P_t) and return ($r_{t+1} \equiv (P_{t+1} + D_{t+1} - P_t) / P_t$) observations are then calculated using the simulated dividend and share index values, a regression is run using the constructed simulated variables, and the modified cut-off $\bar{R}^2$ is then reported using the 95th percentile $\bar{R}^2$ obtained from the modified simulation regression procedure.

The modified simulation procedure leads to a considerably higher simulated cut-off $\bar{R}^2$ (see the modified cut-off $\bar{R}^2$ reported in the final column of Table II). The modified cut-off $\bar{R}^2$ for the Value-weighted CRSP Index dividend yield regression rises sharply to 19.77%, for instance, and the modified cut-off $\bar{R}^2$ for the S&P 500 Index rises to 19.51% (see the final column of Panel A of Table II). The modified cut-off $\bar{R}^2$s reported in Table II therefore make it very clear-cut that dividend yield return predictability is a spurious result since the actual $\bar{R}^2$s obtained for the dividend yield regressions which range from 1.95% to 2.05% are an order of magnitude smaller than the modified cut-off $\bar{R}^2$ levels.

The increase in the modified cut-off $\bar{R}^2$s in Table II highlights the influence of the dividend yield and rates of return variables sharing a common denominator (the share price index level P_t), an effect that can create a spurious correlation even when all variables that make up the numerators and the common denominator in a regression are independent (Pearson, 1897; Kronmal, 1993). This spurious correlation problem arises when all variables except for the constant in a "true" regression equation are divided by a common variable, often in an (incorrect) attempt by the researcher to control for a "common confounding influence". Kronmal (1993, page 381), summarizing Friedlander (1980), outlines how this leads to a biased least squares estimate of the independent coefficient when the "true" regression constant is non-zero (see also Pearson, 1897; Tanner, 1949; Neyman, 1952; and Friedlander, 1980). More importantly, it can also lead to a significant estimate of the overall relationship between the dependent and independent regression variables even when all of the component variables making up the numerators and the denominator of the dependent and independent regression variables are uncorrelated. Intuitively, the common denominator on both sides of the regression model can introduce a correlation between the independent and the dependent variable if the regression constant is not divided by the same “control” variable. Kronmal (1993) documents how the spurious correlation problem leads to incorrect inferences in empirical studies that examine the cross-sectional relationship between weight size and body size as well as the question of whether females lose lung capacity at a faster rate than males as they age.

Pearson (1897) first noted the spurious correlation problem and derived an approximate correlation formula that provides a reference point spurious correlation benchmark when correlating ratios with a common divisor. When random variables *X*, *Y*, and *Z* are uncorrelated then Pearson’s approximate formula for the correlation $r_{(X/Z,\,Y/Z)}$ between ratios with a common divisor *Z* is

$$r_{(X/Z,\,Y/Z)} = \frac{V_Z^2}{\sqrt{(V_X^2+V_Z^2)(V_Y^2+V_Z^2)}}, \qquad (6)$$

where *V* indicates a variable’s coefficient of variation (standard deviation divided by expected value). Pearson’s approximate formula indicates that the ratio correlation will be positive unless *Z* is a constant; if the variables *X*, *Y* and *Z* are independently and identically distributed then the approximate formula correlation value will be .5, even though *X*, *Y* and *Z* are independent!

Pearson’s (1897) approximate formula derivation ignores higher order terms, so Kim (1999) derives an exact formula for the correlation $r_{(X/Z,\,Y/Z)}$ between ratios of uncorrelated variables with a common divisor *Z*:

$$r_{(X/Z,\ Y/Z)} = \frac{V_{1/Z}^2}{\sqrt{[V_X^2(1+V_{1/Z}^2)+V_{1/Z}^2][V_Y^2(1+V_{1/Z}^2)+V_{1/Z}^2]}} \ . \qquad (7)$$

Kim's exact formula is more precise than Pearson's approximate formula when the coefficient of variation of the common divisor (Z) differs significantly from the coefficient of variation of the inverse of the common divisor ($1/Z$). Kim (1999) provides theoretical examples to demonstrate that the correlation between ratios composed of three independent variables can be arbitrarily close to one, and also documents how the spurious correlation problem can affect inferences concerning the empirical relationship between birth rates and death rates in 97 countries in the year 1990. Application of equation (7) indicates that a correlation of .091 would be expected by chance between birth rates and death rates even if births, deaths and population are uncorrelated, thus providing a benchmark to assess whether birth rates are contemporaneously related to death rates.

The Kim (1999) spurious correlation formula can explain why the modified cut-off $\bar{R}^2$s which take account of both spurious correlation and spurious regression are so much higher than the cut-off $\bar{R}^2$s in Table II. Table IV reports implied modified cut-off $\bar{R}^2$ for return predictability regression model (1) that are obtained using analytical correlation formula (7).[9] The Table IV results reveal that the implied modified cut-off $\bar{R}^2$s obtained using analytical formula (7) are 24.25% for the CRSP Value-weighted index and 17.56% for the S&P 500 index, thus providing a benchmark for the influence

[9] Correlation formula (7) can be converted to an implied modified cut-off $\bar{R}^2$ for return predictability regression (1) using the equation, $1-(1-R^2)(T-1)/(T-K)$, where R^2 is estimated using equation (7) and is the square of the correlation between dividend yields (D_t/P_t) and subsequent returns ($r_{t+1} \equiv (P_{t+1}+D_{t+1}-P_t)/P_t$), T is the number of observations, and K is the number of independent variables in the regression model.

of spurious correlation (in isolation) on dividend yield return predictability regression results. The implied modified cut-off $\bar{R}^2$s that are obtained using the analytical formula closely match the modified cut-off $\bar{R}^2$s obtained by simulation of 19.77% for the CRSP Value-weighted index and 19.51% for the S&P 500 index (see Table II), thus indicating that the modified cut-off $\bar{R}^2$s reported in Table II are precisely what should be expected! The modified simulation procedure takes account of the interaction between spurious correlation and spurious regression whereas the analytical formula accounts for spurious correlation only, thus explaining the difference between the analytical and simulation results, but the two sets of results are remarkably close.[10] Both the simulation procedure and analytical formula cut-off $\bar{R}^2$s greatly exceed the implied $\bar{R}^2$s of 1.95% for the CRSP Value-weighted index and 2.05% for the S&P 500 index that are estimated using the actual data, thus strongly reinforcing the findings that dividend yield return predictability is a spurious result (see Tables II and IV). Table IV also implies that concerns about spurious regression in dividend yield return predictability studies can largely be attributed to the spurious correlation problem.

[Table IV about here]

This paper is the first to document the importance of the spurious correlation problem in a financial time series regression setting with auto-correlated regression variables. The difference between the modified cut-off $\bar{R}^2$ and the cut-off $\bar{R}^2$ identifies the influence of the spurious correlation problem on the overall regression relationship. For instance, the cut-off $\bar{R}^2$ in Panel A of Table II indicates that in the absence of a

[10] The Kim (1999) analytical formula provides a benchmark for spurious correlation in isolation by assuming that variables *X*, *Y*, and *Z* are uncorrelated, whereas they are not uncorrelated in the modified simulation procedure, thus explaining why the two sets of results can differ.

spurious correlation problem (and data mining) then we would expect by chance an $\bar{R}^2$ of 3.95% for the CRSP Value-weighted return predictability regression. The spurious correlation problem is therefore instrumental in helping to contribute almost 16% to the modified cut-off $\bar{R}^2$ of 19.77% for the Value-weighted return predictability regression. It can therefore be argued, based upon the results of Table II, that the spurious correlation problem (in interaction with spurious regression) can be viewed as the major contributing factor to incorrect inference in dividend yield return predictability studies.

The Table II spurious regression return predictability results can also be related to the persistent regressor coefficient bias literature (see, e.g., Stambaugh, 1999; Lewellen, 2004). Table II focuses on the significance of the overall relationship between dividend yields and subsequent returns to determine whether the observed relationship is spuriously affected by dependent variable autocorrelation, and does not focus on whether the dividend yield coefficient estimates are biased, since the paper does not focus on the coefficients *per se*.[11] Still, it can be useful to examine how the results relate to recommendations made in the return predictability literature to overcome the spurious regression and coefficient bias problems.

To make this comparison, it can first be noted that if high dividend yields really do predict higher subsequent returns, then an increase in the dividend yield should also forecast an increase in subsequent returns. Ferson, Sarkissian and Simin (2003b) suggest that testing the relationship between returns and changes in predictor variables away from their trailing moving average can overcome the spurious regression problem, since detrending predictor variables will create a less persistent independent variable.

[11] Ferson, Sarkissian and Simin (2003b) point out that the return predictability coefficient bias literature does not address the interaction between spurious regression and data mining.

Subsequent returns can therefore be regressed against changes in the dividend yield from its trailing moving average, with a twelve month moving average lag length being utilized (as recommended by Ferson, Sarkissian and Simin, 2003b). Lewellen (2004) implements a somewhat related regression model to test for return predictability, since he uses the natural logarithm of the dividend yield as the regression independent variable; the natural logarithm of the dividend yield can be interpreted as a change in the dividend yield away from a dividend yield of one. Campbell and Yogo (2003) also recommend detrending dividend yields when testing for return predictability.[12] Subsequent real returns are therefore regressed against the detrended dividend yield using the regression model

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_1 X_t + e_{t+1}, \tag{8}$$

where r_{t+1} denotes the annual real index return at time $t+1$, X_t denotes the stochastic detrended dividend yield calculated as

$$X_t = \frac{D_t}{P_t} - \frac{1}{t}\sum_{j=1,\ldots,t} \frac{D_{t-j}}{P_{t-j}}, \tag{9}$$

D_t is the level of real annual dividends during the twelve months preceding time t, and P_t is the real stock index level at time t. While different numbers of lags could be used in the detrending, a 12-month lag is used, as recommended by Ferson, Sarkissian and Simin (2003b).

Results for return predictability regression model (8) are reported in Table V. Interestingly, the results imply that an increase in the dividend yield actually foreshadows lower, not higher, returns (although the overall relationship is clearly insignificant).

[12] Campbell and Yogo (2003) recommend subtracting from the dividend yield independent variable the innovations in the yield that are correlated with returns to obtain a less noisy independent variable, thus eliminating some of the noise and increasing the power of the test.

Regression model (8) has not been tested before, but the Table V results can be compared to Lewellen's (2004) finding that the return predictability regression coefficient of the natural logarithm of the dividend yield is actually negative within sub-periods for the monthly CRSP equally-weighted and value-weighted Indices when estimated using the Stambaugh (1999) bias-adjustment.[13] The Table V results again reinforce the Table II findings that the observed relationship between dividend yields and subsequent returns is spurious, and they also make it extremely unlikely that predictor coefficient bias adjustments would lead to a significant overall relationship, since bias adjustments do not normally reverse the sign of the coefficients. The Table V results, combined with the Table II results, further suggest that a spurious correlation effect influences the estimated relationship between dividend yields and subsequent returns, since the reversal of the dividend yield regression coefficient sign between Table II and Table V is consistent with spurious correlation effects outlined in Kronmal (1993).[14] The Table V results also cast severe doubt on the temporary mispricing explanation of return predictability, since the Table V results imply that an increase in mispricing would predict higher, not lower, subsequent returns.

[Table V about here]

A final challenge to the spurious regression conclusion obtained from Table II is the argument that that dividend yield return predictability can be shown to be a non-spurious result if the definition of dividends is expanded to include share repurchases

---

[13] Lewellen (2004) reports very low adjusted R^2s (less than .014) when regressing subsequent monthly returns against the natural logarithm of the dividend yield, and tends to find a significant dividend yield coefficient estimate only when imposing the assumption of unitary dividend yield serial correlation, an assumption that is clearly violated in Table III for annual data.

[14] Kronmal (1993) reports how storks can be responsible for either a decrease or an increase in births, depending on how the dependent variable births is regressed against the independent variable storks in alternative incorrectly specified regression models that are subject to spurious correlation!

(Boudoukh *et al*, 2004; Robertson and Wright, 2003). This is especially important in recent decades, since the decline in the traditional dividend yield and in dividend yield return predictability has occurred at the same time as share repurchases have been substituted for traditional dividends.[15] Table VI therefore reproduces the return predictability regression coefficients and R^2s reported in Table 2 of Boudoukh *et al* (2004) using the broadly defined "total payout" dividend yield variable that includes share repurchases. In addition, Table VI also reports the modified cut-off $\bar{R}^2$ that is obtained from the modified cut-off simulation procedure using the parameters of the Boudoukh *et al* (2004) total payout dividend and share price index series. Once again, the modified cut-off $\bar{R}^2$s reported in Table VI make it very clear-cut that dividend yield return predictability is a spurious result, no matter what dividend yield series is used, since the actual $\bar{R}^2$s reported in Table VI of 5.96% for the full sample and 10.32% for the shorter 1926 to 1985 sample are considerably smaller than the modified cut-off $\bar{R}^2$ levels.

[Table VI about here]

The results of Tables II, IV, V, and VI, taken together, imply that the spurious effect of regressing returns against a highly persistent explanatory variable such as the dividend yield is strongly reinforced by a spurious correlation effect. To further explore the origins of this spurious return predictability effect, models of dividend behaviour are also tested for spurious regression.

[15] Note that, unlike the traditional dividend yield variable used in Panel A of Table II, the dividend yield from a constant dividend return predictability variable used in Panel B of Table II is not affected by this recent substitution away from dividends.

III. DIVIDEND BEHAVIOUR RESULTS

The persistence properties of dividend yields that contribute to spurious return predictability imply that dividend behaviour models are also likely to be subject to spurious regression, since the dependent and independent variables in the Lintner (1956) and Marsh and Merton (1987) dividend behaviour models are all very persistent (see also Ferson, Sarkissian and Simin, 2003a). These dividend behaviour models are examined in Table VII. The Table VII results indicate that the Marsh and Merton (1987) model is not subject to spurious regression, whereas spurious regression appears to play an extremely important role in alternative versions of the Lintner (1956) model which use highly persistent lagged dividend and earnings terms to explain subsequent dividend levels or dividend changes.

Panel A of Table VII reveals that current log returns play a very important role in explaining subsequent dividend changes in the Marsh and Merton (1987) model (recall equation (5)):

$$\log\left(\frac{D_{t+1}}{D_t}\right)+\frac{D_t}{P_{t-1}}=y_0+y_1\log\left(\frac{P_t+D_t}{P_{t-1}}\right)+y_2\log\left(\frac{D_t}{P_{t-1}}\right)+e_{t+1}\ .$$

The $\bar{R}^2$s of 33.95% for the CRSP Value-weighted index and 29.84% for the S&P 500 index greatly exceed the modified cut-off $\bar{R}^2$s of 9.23% and 12.04% (respectively) in Panel A of Table VII, thus implying that the results are not due to spurious regression.

[Table VII about here]

The Lintner (1956) model of the time *t* dividend level choice (recall equation (4)),

$$D_t = q_0 + q_1 E_{t-1} + q_2 D_{t-1} + e_t,$$

is tested using the S&P 500 index only due to earnings data availability (see also Arnott and Asness, 2003).[16] Panel B of Table VII reveals that the estimated independent variable coefficients for regression model (4) are very similar to those found in Lintner (1956), even though the data set is extended by four decades. The results imply that lagged earnings and (especially) lagged dividends explain almost all of the variation in the current dividend level. The $\bar{R}^2$ exceeds 90%, as in Lintner (1956), and slightly exceeds the modified cut-off $\bar{R}^2$, thus implying that the original Lintner model appears to somewhat survive the spurious regression problem. The very high modified cut-off $\bar{R}^2$ reported for the Lintner model in Panel B of Table VII suggests that the spurious regression problem has an influence on the Lintner model regression, even if it is not solely responsible for the results, thus indicating the potential need for a reformulation of the original Lintner model.

A step towards a reformulation of the Lintner regression model is already provided in Lintner (1956), where the Lintner dividend behaviour theoretical model is originally presented in terms of changes in dividends, not the dividend level as in regression model (4) above (see equation (1) in Lintner, 1956, and see also Grullon and Michaely, 2002; Fama, 1974). Restating the Lintner dividend change model (Lintner (1956) equation (1)) in terms of information that is observable at time t leads to the following regression model:

$$D_t - D_{t-1} = l_0 + l_1 E_{t-1} + l_2 D_{t-1} + e_t \qquad (10)$$

[16] The Standard and Poor 500 earnings data set is obtained from Robert Shiller's web site. This paper follows the recent practice of using reported earnings to test the Lintner model due to data availability, rather than operating or cash earnings.

Results for the Lintner (1956) dividend change regression model (10) are presented in Panel C of Table VII and are consistent with the results for the Lintner (1956) dividend level regression model (4) presented in Panel B. Having changes in dividends rather than the dividend level as the dependent variable in the Lintner regression model greatly reduces the $\bar{R}^2$ (from 92.9% down to only 3.13%) as well as the modified cut-off $\bar{R}^2$ (from 84.48% to 14.8%) in Panel C, so the $\bar{R}^2$ is exceeded by the modified cut-off $\bar{R}^2$, thus indicating the Panel C results are spurious. The following section takes further steps to reformulate the Lintner dividend behaviour model entirely in terms of first differences on both sides of the regression equation so that it can remain as a simple and intuitive alternative to (or complement for) the Marsh and Merton (1987) dividend behaviour model without being subject to spurious regression.

IV. AN ALTERNATIVE DIVIDEND BEHAVIOUR MODEL SPECIFICATION

A source of persistence in the Lintner (1956) dividend level model that is likely to have a very important effect on the model's time series regression properties is a common time trend in both the dependent and independent regression variables (recall equation (4)):

$$D_t = q_0 + q_1 E_{t-1} + q_2 D_{t-1} + e_t ,$$

A standard solution to this problem is a reformulation entirely in terms of first differences (see also equation (1) in Lintner, 1956). To derive the Lintner model in terms of first differences, first note that equation (4) also implies:

$$D_{t-1} = q_0 + q_1 E_{t-2} + q_2 D_{t-2} + e_{t-1} . \qquad (11)$$

Equation (4) minus equation (11) leads to a reformulated first-difference Lintner model

$$D_t - D_{t-1} = q_1(E_{t-1} - E_{t-2}) + q_2(D_{t-1} - D_{t-2}) + (e_t - e_{t-1})\,. \tag{12}$$

A second potential improvement to the Lintner model follows from the Lintner (1956) argument that dividends would not be adjusted to changes in earnings that are only temporary. This led Marsh and Merton (1987) to introduce the concept of permanent earnings and to argue that dividends are determined in relation to permanent earnings. They further argue that, in an efficient market, the current share price index level is equal to the present value of all future permanent earnings. By assuming that the long-run discount rate is relatively constant, then "permanent earnings" divided by the long-run discount rate is proportionate to the current share index level, so by adjusting dividends through time in reaction to permanent earnings then managers are adjusting dividends to the share index level. The Marsh and Merton permanent earnings innovation can therefore be directly introduced into the Lintner model by substituting the price index level P in for earnings E in the reformulated Lintner first-difference regression model (12):

$$D_t - D_{t-1} = q_1(P_{t-1} - P_{t-2}) + q_2(D_{t-1} - D_{t-2}) + (e_t - e_{t-1})\,. \tag{13}$$

Results for the reformulated Lintner first-difference regression model (13) are reported in Table VIII.[17, 18] Regression model (13) is estimated with an intercept to test the implication from equation (13) that the intercept should equal zero.[19] A notable result

---

[17] Note that the reformulated Lintner first difference regression model (13) nests the constant growth dividend model *P = kE/(r-g)*, where *k* is the (constant) dividend payout ratio from earnings, *r* is the constant discount rate, and *g* is the constant growth rate.

[18] The error term in equation (9), $e_t - e_{t-1}$, is accommodated using a Heteroskedasticity and autocorrelation Consistent Covariance estimation procedure within Generalized Method of Moments.

[19] The Table VIII results are insensitive to the exclusion of the intercept term in regression model (13); results not reported.

from the table are much lower $\bar{R}^2$s for the Lintner first-difference model (12.42% for the CRSP Value-weighted index and 8.69% for the S&P 500 index) which now exceed the greatly reduced modified cut-off $\bar{R}^2$s of 7.12% and 6.91% respectively, thus implying a lessening of the influence of spurious regression in the re-specified model. The results imply that changes in the aggregate level of dividends are explained by lagged share price index innovations, and are consistent with the predictions of Marsh and Merton (1987).[20] The model prediction that the intercept equals zero is also not rejected. The re-specified Lintner dividend first-difference regression model (13) therefore appears to provide a correctly specified dividend behaviour model, thus allowing the Lintner model to be retained as a simple and intuitive explanation of dividend behaviour.

[Table VIII about here]

V. CONCLUSION

It is now known that spurious regression is a very serious problem when highly persistent variables are used in a time series regression model to predict or explain dependent variables that are at least partially persistent. The spurious regression problem is compounded by a spurious correlation problem when the dependent and independent variables in a regression equation are ratios constructed from common component variables. Both of these problems are present in dividend yield return predictability regression models, thus explaining why standard dividend return predictability regression results are spurious, even when the dividend yield definition is broadened to include distributions to investors via share repurchases. Return predictability of the dividend

[20] The overall relationship is significant, but Table VIII also indicates that the individual coefficients are marginally significant (at best) due to heteroskedasticity.

yield is shown to be due to the persistence properties of the dividend yield as well as a spurious correlation that results from regression variables being constructed from common component variables, rather than (as widely argued) any property of the dividend level that is related to risk, return or mispricing. The paper's results therefore strongly reinforce the Kronmal (1993) message that extreme care should be taken when using ratios in regression analysis, especially in time series regressions when highly persistent ratios are used as predictor or explanatory variables.

The standard Lintner dividend behaviour model is also strongly affected by the spurious regression problem since the dependent and independent variables in the model are all very persistent. A reformulation of the Lintner (1956) dividend model in terms of first differences provides a dividend behaviour explanatory model that is far less subject to spurious regression, and also directly incorporates the Marsh and Merton (1987) permanent earnings explanation of dividend behaviour. Results for the reformulated Lintner first difference dividend model imply that changes in the aggregate level of dividends are explained by lagged share price innovations**, as** predicted by Marsh and Merton (1987).

APPENDIX

A1 Cut-off $\bar{R}^2$ Simulation Procedure

The cut-off $\bar{R}^2$ simulation procedure provides the cut-off $\bar{R}^2$ that is obtained by regressing simulated dependent and independent variables that are uncorrelated but have the same autocorrelation properties as the actual dependent and independent variable data series. The simulated independent variable series X_t is generated as

$$X_t = a_X + r_X X_{t-1} + e_t \quad \text{for } t = 2, 3, \ldots, n\,, \tag{A1}$$

where a_x is the intercept and r_x is the first order autocorrelation coefficient. The unconditional mean of the independent variable series (m_X) is

$$m_X = \frac{a_X}{1 - r_X}$$

and the variance s_X^2 is

$$s_X^2 = \frac{s_e^2}{1 - r_X^2}\,.$$

The parameters that calibrate the simulation, m_X, s_X^2 and r_x, are taken from the actual data. The simulation is started at the unconditional mean m_X and the error term is generated from a normal with mean zero and variance $s_X^2(1 - r_x^2)$.

Similarly, the uncorrelated dependent variable series is generated as

$$r_t = a_r + r_r r_{t-1} + e_t \quad \text{for } t = 2, 3, \ldots, n \tag{A2}$$

with

$$m_r = \frac{a_r}{1 - r_r}$$

and

$$s_r^2 = \frac{s_e^2}{1 - r_r^2},$$

where the parameters m_r, s_r^2 and r_r are taken from the actual data.[21] The dependent variable simulation is started at the unconditional mean m_r and the error term for process A(2) is generated independently of process A(1) from a normal with mean zero and variance $s_r^2(1 - r_r^2)$.

The simulated dependent and independent variable series are stored, and an ordinary least squares regression is run. The process is repeated 10,000 times. The $\bar{R}^2$s are recorded for each regression and ranked from lowest to highest. The 95th percentile $\bar{R}^2$ is then reported as the 5% cut-off $\bar{R}^2$.

A2 Modified Cut-off $\bar{R}^2$ Simulation Procedure

To obtain the modified cut-off $\bar{R}^2$, the cut-off $\bar{R}^2$ simulation procedure is modified to recognize the dependency of both returns (r_{t+1}) and the dividend yield variable ($X_t \equiv D_t / P_t$) on the same underlying component variables (the share index level P and the dividend level D) in regression model

$$r_{t+1} \equiv \frac{P_{t+1} + D_{t+1} - P_t}{P_t} = b_0 + b_1 \left[\frac{D_t}{P_t} \right] + e_{t+1} . \tag{1}$$

The modified cut-off $\bar{R}^2$ simulation procedure simulates the dividend series (D_t) and price index series (P_t) independently prior to constructing the dependent and

[21] Return persistence in equation (A2) is directly due to return autocorrelation whereas return persistence in equation (2) of the Ferson, Sarkissian, and Simin (2003a) simulation model is due to an unobserved persistent latent variable.

independent variable return and dividend yield series. The simulated dividend series D_t is generated as

$$\ln D_t = m_D + r_D \ln D_{t+1} + e_t \quad \text{for } t = 2, 3, \ldots, n\,, \tag{A3}$$

and is then exponentiated to obtain the data in levels, where the parameters m_D, s_e^2 and r_D are taken from the actual dividend data. The simulation is started at the unconditional mean of the dividend series m_D and the error term is generated from a normal with mean zero and variance $s_D^2(1-r_D^2)$.

Similarly, the uncorrelated price index series is generated as

$$\ln P_t = m_P + r_P \ln P_{t-1} + e_t \quad \text{for } t = 2, 3, \ldots, n\,, \tag{A4}$$

and is then exponentiated to obtain the data in levels, where the parameters m_P, s_P^2 and r_P are taken from the actual price index series. The simulation is started at the unconditional mean m_P and the error term for process A(4) is generated independently of process A(3) from a normal with mean zero and variance $s_P^2(1-r_P^2)$.

A dividend yield series ($X_t \equiv D_t / P_t$) is subsequently created from the simulated dividend series D_t and simulated price index series P_t:

$$X_t = \left[\frac{D_t}{P_t}\right] \quad \text{for } t = 1, \ldots, n\,.$$

A return series ($r_{t+1} \equiv (P_{t+1} + D_{t+1} - P_t)/P_t$) is created from the same simulated dividend series D_t and simulated price index series P_t:

$$r_{t+1} = \frac{P_{t+1} + D_{t+1} - P_t}{P_t} \quad \text{for } t = 1, 2, \ldots, n\text{-}1\,.$$

The dividend yield series (X_t) and the return series (r_{t+1}) are then stored. An ordinary least squares regression is run with the stored dividend yield series (X_t) as the independent variable, and the stored return series (r_{t+1}) as the dependent variable.[22] The process is repeated 10,000 times, with the adjusted R^2s being recorded for each regression and ranked from lowest to highest. The 95th percentile $\bar{R}^2$ is then reported as the 5% modified cut-off $\bar{R}^2$.

[22] In Table VI, the simulated regression uses logarithmic real returns as the dependent variable and the natural logarithm of the dividend yield as the independent variable, as in Boudoukh *et al* (2004).

REFERENCES

Allen, Frank, Antonio E. Bernardo, and Ivo Welch, 2000, A Theory Of Dividends Based On Tax Clienteles, Journal of Finance 55, 2499-2536.

Arnott, Robert D., and Clifford S. Asness, 2003, Surprise! Higher Dividends = Higher Earnings Growth, Financial Analysts Journal 59, 70-87.

Balvers, Ronald J., Thomas F. Cosimano, and Bill McDonald, 1990, Predicting stock returns in an efficient market, Journal of Finance 45, 1109-1128.

Bansal, Ravi and Amir Yaron, 2004, Risks For The Long Run: A Potential Resolution Of Asset Pricing Puzzles, Journal of Finance 59, 1481-1509.

Black, Fischer, and Myron Scholes, 1974, The Effects Of Dividend Yield And Dividend Policy On Common Stock Prices And Returns, Journal of Financial Economics 1, 1-22.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?, Review of Financial Studies 12, 405-428.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2004, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, Working Paper, University of Pennsylvania.

Campbell, John Y., and Robert J. Shiller, 1988a, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies 1, 195-228.

Campbell, John Y., and Robert J. Shiller, 1988b, Stock Prices, Earnings and Expected Dividends, Journal of Finance 43, 661-676.

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

Campbell, John Y., and Motohiro Yogo, 2003, Efficient Tests of Stock Return Predictability, forthcoming at Journal of Financial Economics.

Cochrane, John H., 1999, New Facts In Finance, Economic Perspectives 23, 36-58.

Cremers, K.J. Martijn, 2002, Stock Return Predictability: A Bayesian Model Selection Perspective, Review of Financial Studies 15, 1223-1249.

Donaldson, R. Glen and Mark Kamstra, 1996, A New Dividend Forecasting Procedure That Rejects Bubbles In Asset Prices: The Case Of 1929's Stock Crash, Review of Financial Studies 9, 333-383.

Fama, Eugene F., 1974, The Empirical Relationships between the Dividends and Investment Decisions of Firms, American Economic Review 64, 304-318.

Fama, Eugene F., and Kenneth R. French, 1988, Dividend Yields and Expected Stock Returns, Journal of Financial Economics 22, 3-25.

Ferson, Wayne E., and Campbell R. Harvey, 1991, The Valuation of Economic Risk Premium, Journal of Political Economy 99, 385-415.

Ferson, Wayne E., Sarkissian, Sergei, and Timothy T. Simin, 2003a, Spurious Regressions in Financial Economics?, Journal of Finance 58, 1393-1413.

Ferson, Wayne E., Sarkissian, Sergei, and Timothy T. Simin, 2003b, Is Stock Return Predictability Spurious?, Journal of Investment Management 1, 1-10.

Friedlander, L.J., 1980, *A Study of the Correlation between Ratio Variables*, PhD Dissertation, University of Washington.

Foster, F. Douglas, Tom Smith and Robert E. Whaley, 1997, Assessing The Goodness-Of-Fit Of Asset Pricing Models: The Distribution Of The Maximal R-Squared, Journal of Finance 52, 591-607.

Goetzmann,William N., and Philippe Jorion, 1993, Testing The Predictive Power Of Dividend Yields, Journal of Finance 48, 663-679.

Goyal, Amit, and Ivo Welch, 2003, Predicting Equity Premium with Dividend Ratios", Management Science 49, 639-654.

Grullon, Gustavo, and Roni Michaely, 2002, Dividends, Share Repurchases, and the Substitution Hypothesis, Journal of Finance 57, 1649-1684.

Harvey, Campbell R., 1989, Time-Varying Conditional Covariances in Tests of Asset Pricing Models, Journal of Financial Economics 24, 289-317.

Hodrick, Robert J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, Review of Financial Studies 5, 357-386.

Kim, Ji-Hyun, 1999, Spurious correlation between ratios with a common divisor, Statistics & Probability Letters 44, 383-386.

Kothari, S.P., and Jay Shanken, 1997, Book-to-market, Dividend Yield, and Expected Market Returns: A time-series Analysis, Journal of Financial Economics 44, 169-203.

Kronmal, Richard A., 1993, Spurious Correlation and the Fallacy of the Ratio Standard Revisted, Journal of Royal Statistical Society 156, Series A, 379-392.

Lamont, Owen, 1998, Earnings and Expected Returns, Journal of Finance 53, 1563-1587.

Lettau, Martin, and Sydney C. Ludvigson, 2005, Expected returns and expected dividend growth, Journal of Financial Economics 76, 583-626.

Lewellen, Jonathon, 2004, Predicting Returns With Financial Ratios, Journal of Financial Economics 74, 209-235.

Lintner, John, 1956, Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes, American Economic Review 46, 97-113.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-Snooping Biases In Tests Of Financial Asset Pricing Models, Review of Financial Studies 3, 431-468.

Marsh, Terry A., and Robert C. Merton, 1987, Dividend Behavior for the Aggregate Stock Market, Journal of Business 60, 1-40.

Nelson, Charles R, and Myung J. Kim, 1993, Predictable Stock Returns: The Role of Small Sample Bias, Journal of Finance 48, 641-661.

Neyman, Jerzy, 1952, *Lectures and Conferences on Mathematical Statistics and Probability*, 2nd edition, pp. 143-154, US Department of Agriculture, Washington D.C.

Pearson, Karl, 1897, Mathematical Contributions to the Theory of Evolution - On a Form of Spurious Correlation which may Arise when Indices are Used in the Measurement of Organs, Proceedings of the Royal Society 60, 489-498.

Pesaran, M. Hashem, and Allan Timmermann, 1995, Predictability Of Stock Returns: Robustness And Economic Significance, Journal of Finance 50, 1201-1228.

Pontiff, Jeffrey, and Lawrence D. Schall, 1998, Book-to-market Ratios as Predictors of Market Returns, Journal of Financial Economics 49, 141-160.

Robertson, Donald, and Stephen H. Wright, 2003, Dividends, total cashflows to shareholders and predictive return regression, Working Paper, Cambridge University.

Rozeff, Michael S., 1984, Dividend Yields Are Equity Risk Premiums, Journal of Portfolio Management 11, Fall, 68-75

Shiller, Robert J., 1983, Do Stock Prices Move Too Much To Be Justified By Subsequent Changes In Dividends?: Reply, American Economic Review 73, 236-237.

Stambaugh, Robert F., 1999, Predictive Regressions, Journal of Financial Economics 54, 375-421.

Sullivan, Ryan, Allan Timmermann and Halbert White, 1999, Data-Snooping, Technical Trading Rule Performance, And The Bootstrap, Journal of Finance 54, 1647-1691.

Sullivan, Ryan, Allan Timmermann and Halbert White, 2001, Dangers Of Data Mining: The Case Of Calendar Effects In Stock Returns, Journal of Econometrics 105, 249-286.

Tanner, John. M., 1949, Fallacy of Per-weight and Per-surface Area Standard, and their relation to Spurious Correlation, Journal of Applied Physiology 2, 1-15.

Touros, Walter, Valkanov, Rossen, and Shu Yan, 2004, On Predicting Stock Returns with Nearly Integrated Explanatory Variables, Journal of Business 77, 937- 966.

Valkanov, Rossen, 2003, Long-horizon regressions: theoretical results and applications, Journal of Financial Economics 68, 201-232.

Whitelaw, Robert F., 1994, Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns, Journal of Finance 49, 515-541.

Table I
Fourteen previous studies on predictability of the dividend yield

This table presents model specifications used in fourteen recent empirical studies on the predictability of the lagged dividend yield on stock index return. Under Sample Data, VW refers to CRSP value-weighted index, EW refers to CRSP equal-weighted index, and SP refers to S&P 500 (or composite) index. For sample period, we report the whole sample period used in the studies since some studies also examine sub-periods. For Frequency of Data, M refers to monthly data, Q refers to quarterly data and A refers to annual data. For Return Horizon, M (Q) denotes monthly (quarterly) return, 1 denotes one-year return, 2 denotes two-year return, and so on. Under Empirical Finding, '+' is used to signify a positive but insignificant relation, and '++' signifies a significant positive relation. While some studies only report R^2, others report $\bar{R}^2$ (adjusted R^2). NA denotes not available. Note that in Campbell and Shiller (1988), the sample period is 1887-1986 for SP and is 1926-1986 for VW.

Study	Sample Data	Sample Period	Freq. of Data	Return Horizon	Empirical Finding	R^2 or $\bar{R}^2$
Shiller (1984)	SP	1872-1983	A	1	++	0.06
Rozeff (1984)	SP	1926-1982	A	1	++	0.11
Campbell and Shiller (1988a)	SP VW	1871(1926) -1986	A	1	++	NA
Fama and French (1988)	VW EW	1927-1986	M/Q/A	M/Q/1/2/3/4	++	0.00-0.29
Balvers, Cosimano & McDonald (1990)	VW	1947-1987	A	1	++	0.069
Nelson and Kim (1993)	SP	1872-1986	A	1/3/10	+	0.016-0.267
Goetzmann and Jorion (1993)	SP	1927-1990	M	M/1/2/3/4	+	0.01-0.39
Kothari and Shanken (1997)	VW EW	1926-1991	A	1	++	0.014-0.075
Lamont (1998)	SP	1947-1994	Q	Q	++	0.05
Stambaugh (1999)	VW	1927-1996	M	M	+	NA
Goyal and Welch (2003)	VW	1926-2002	A	1	+	0.0431
Lewellen (2004)	VW EW	1946-2002	M	M	++	0.004-0.008
Lettau and Ludvigson (2005)	VW	1948-2001	A	1/2/3/4/5/6	+	0.08-0.23
Campbell and Yogo (2005)	VW SP	1880-2002	M/Q/A	M/Q/1	++	NA

Table II
Regressions of Annual Real Index Returns on Two Predictors

This table presents OLS regressions of annual real index return on two predictors:

$$r_{t+1} = b_0 + b_1 X_T + e_{t+1}$$

where r_{t+1} denotes annual real index return at time *t+1;* annual real index returns are based on two measures: *VWR*, CRSP value-weighted index return; *SPR*, S&P 500 index return; X_t variously denotes the time *t* real dividend yield (D_t / P_t), and the dividend yield from a constant dividend (c / P_t); D_t is the level of real annual dividends during the twelve months preceding time *t*, P_t is the real stock index level at time *t*, *C* is the unconditional average of the dividend level over the sample period. The sample begins in January 1927 and extends through December 2004. $\bar{R}^2$ denotes adjusted R^2. All data are obtained from CRSP. Regressions are estimated by OLS and *t*-statistics are adjusted for autocorrelation and heteroskedasticity using Newey-West (1987). Cut-off $\bar{R}^2$ is obtained using a simulation procedure where dependent and independent variables are uncorrelated but have the same autocorrelation properties as the actual data. Modified cut-off $\bar{R}^2$ recognizes the dependency of both the return and dividend yield variables on the share index and dividend levels.

Variable	n	b_0	$t(b_0)$	b_1	$t(b_1)$	$\bar{R}^2$	Cut-off $\bar{R}^2$	Cut-off $\bar{R}^2$ with Data Mining	Modified Cut-off $\bar{R}^2$
Panel A: $r_{t+1} = b_0 + b_1(D_t / P_t) + e_{t+1}$									
VWR	78	-0.009	-0.162	2.336	1.631	1.95%	3.95%	7.40%	19.77%
SPR	78	-0.002	-0.030	2.194	1.604	2.05%	4.09%	7.66%	19.51%
Panel B: $r_{t+1} = b_0 + b_1(C / P_t) + e_{t+1}$									
VWR	78	0.019	0.550	1.341	2.035	3.22%	3.86%	7.77%	20.56%
SPR	78	0.020	0.515	1.385	1.966	3.30%	3.87%	7.30%	20.39%

Table IIIA
Descriptive Statistics

Panel A presents summary statistics on the natural logarithm value of the real annual dividend (D_t), real dividend yield (D_t / P_t), real price index (P_t) and the dividend yield from a constant dividend (c / P_t) as well as the logarithmic real index return (Ln r_t) on the annual CRSP Value-weighted and S&P 500 Indices for the time period 1927 to 2004. D_t is the level of real annual dividends during the twelve months preceding time t, P_t is the real stock index level at time t, $\ln(r_t)$ denotes continuous index return at time t, and c is the unconditional average of the dividend level over the sample period. Panel B reports autocorrelations of the variables up to 6 lags, and correlations among the variables are presented in Panel C.

	CRSP Value-Weighted Index					S&P 500 Index				
	$Ln(D_t)$	$Ln(D_t/P_t)$	$Ln\ r_t$	$Ln(P_t)$	$Ln(C/P_t)$	$Ln(D_t)$	$Ln(D_t/P_t)$	$Ln\ r_t$	$Ln(P_t)$	$Ln(C/P_t)$
Panel A: Descriptive Statistics										
Mean	2.243	-3.286	0.065	5.529	-3.231	2.407	-3.291	0.067	5.698	-3.248
Median	2.332	-3.262	0.098	5.506	-3.209	2.500	-3.288	0.096	5.665	-3.216
Maximum	2.864	-2.363	0.441	6.922	-1.971	2.877	-2.293	0.421	7.326	-2.067
Minimum	1.329	-4.215	-0.461	4.268	-4.625	1.749	-4.478	-0.443	4.517	-4.876
Std. Dev.	0.359	0.393	0.190	0.676	0.676	0.324	0.447	0.194	0.702	0.702
Panel B: Autocorrelation										
1	0.894	0.784	0.033	0.880	0.880	0.895	0.807	0.029	0.885	0.885
2	0.794	0.594	-0.181	0.774	0.774	0.795	0.635	-0.153	0.780	0.780
3	0.715	0.467	-0.068	0.679	0.679	0.722	0.499	-0.024	0.678	0.678
4	0.663	0.347	-0.061	0.583	0.583	0.667	0.362	-0.056	0.577	0.577
5	0.632	0.287	0.020	0.511	0.511	0.629	0.276	0.009	0.501	0.501
6	0.600	0.219	0.028	0.448	0.448	0.598	0.221	0.041	0.436	0.436
Panel C: Correlations										
$Ln(D_t)$	1.000					1.000				
$Ln(D_t/P_t)$	-0.616	1.000				-0.648	1.000			
$Ln\ r_t$	-0.070	-0.320	1.000			-0.086	-0.262	1.000		
$Ln(P_t)$	0.889	-0.908	0.149	1.000		0.874	-0.936	0.127	1.000	
$Ln(C/P_t)$	-0.889	0.908	-0.149	-1.000	1.000	-0.874	0.936	-0.127	-1.000	1.000

Table IIIB
Descriptive Statistics

Panel A presents summary statistics on the value of the real annual dividend (D_t), real dividend yield (D_t / P_t), real price index (P_t) and the dividend yield from a constant dividend (c / P_t) as well as the logarithmic real index return (Ln r_t) on the annual CRSP Value-weighted and S&P 500 indices for the time period 1926 to 2004. D_t is the level of real annual dividends during the twelve months preceding time *t*, P_t is the real stock index level at time *t*, r_t denotes continuous index return at time *t*, and *C* is the unconditional average of the dividend level over the sample period. Panel B reports autocorrelations of the variables up to 6 lags, and correlations among the variables are presented in Panel C.

	CRSP Value-Weighted Index					S&P 500 Index				
	D_t	D_t / P_t	r_t	P_t	C / P_t	D_t	D_t / P_t	r_t	P_t	C / P_t
Panel A: Descriptive Statistics										
Mean	9.997	0.040	0.086	317.290	0.049	11.645	0.041	0.089	385.098	0.048
Median	10.298	0.038	0.103	246.284	0.040	12.185	0.037	0.100	288.760	0.040
Maximum	17.538	0.094	0.554	1,014.691	0.139	17.766	0.101	0.524	1,519.010	0.127
Minimum	3.776	0.015	-0.369	71.414	0.010	5.746	0.011	-0.358	91.516	0.008
Std. Dev.	3.261	0.015	0.196	235.413	0.031	3.397	0.017	0.201	311.244	0.031
Panel B: Autocorrelation										
1	0.878	0.715	0.007	0.840	0.881	0.901	0.733	0.016	0.859	0.886
2	0.791	0.471	-0.170	0.713	0.755	0.817	0.515	-0.140	0.724	0.776
3	0.714	0.354	-0.072	0.558	0.662	0.746	0.407	-0.030	0.541	0.694
4	0.651	0.281	-0.080	0.398	0.592	0.683	0.311	-0.073	0.370	0.623
5	0.598	0.258	0.033	0.299	0.560	0.628	0.251	0.019	0.271	0.581
6	0.547	0.239	0.011	0.249	0.516	0.567	0.241	0.022	0.224	0.531
Panel C: Correlations										
D_t	1.000					1.000				
D_t / P_t	-0.633	1.000				-0.660	1.000			
r_t	-0.070	-0.376	1.000			-0.092	-0.329	1.000		
P_t	0.842	-0.783	0.067	1.000		0.792	-0.778	0.059	1.000	
C / P_t	-0.867	0.853	-0.166	-0.764	1.000	-0.884	0.874	-0.151	-0.741	1.000

Table IV
Spurious Correlation Analysis

This table presents a comparison of cutoff correlation and actual correlation between index return and dividend yield (Kim, 1999). The analysis is conducted on both CRSP Value-weighted index and S&P 500 index over the period 1927-2004. The cutoff correlation between index return and lagged dividend yield variables ($r_{(X/Z,Y/Z)}$) takes into account the spurious correlation problem and is calculated as:

$$r_{(X/Z,\,Y/Z)} = \frac{V_{1/Z}^2}{\sqrt{[V_X^2(1+V_{1/Z}^2)+V_{1/Z}^2][V_Y^2(1+V_{1/Z}^2)+V_{1/Z}^2]}}$$

where X is the sum of level of real annual dividends at time $t+1$ (*i.e.* D_{t+1}) and the real stock index level at time $t+1$ ($P_{t+1}+D_{t+1}$); Y is the level of real annual dividends at time t (D_t); Z is the real stock index level at time t (P_t); V is the coefficient of variation which is equal to standard deviation of the variable dividend by expected value of the variable. Note that $r_{(X/Z,Y/Z)}$ assumes X, Y and Z are uncorrelated. $\bar{R}^2$ denotes adjusted R^2. Implied $\bar{R}^2$ is calculated using the equation $1-(1-R^2)\frac{T-1}{T-K}$, where R^2 is calculated as the square of correlation, T is the number of observations, and K is the number of independent variables in the regression. Actual correlation between index return and lagged dividend yield is observed from actual data.

Variable	$P_{t+1}+D_{t+1}$	D_t	P_t	$1/P_t$
Panel A: CRSP Value-Weighted Index				
Mean	327.286	9.852	306.906	0.005
Standard Deviation	238.165	3.171	225.430	0.003
Coefficient of Variation	0.728	0.322	0.735	0.623
Cutoff correlation between index return and lagged dividend yield	0.502			
Implied Modified Cutoff $\bar{R}^2$	24.25%			
Actual correlation between index return and lagged dividend yield	0.180			
Implied Actual $\bar{R}^2$	1.95%			
Panel B: S&P 500 Index				
Mean	396.743	11.504	372.592	0.004
Standard Deviation	313.941	3.367	301.253	0.003
Coefficient of Variation	0.791	0.293	0.809	0.628
Cutoff correlation between index return and lagged dividend yield	0.432			
Implied Modified Cutoff $\bar{R}^2$	17.56%			
Actual correlation between index return and lagged dividend yield	0.182			
Implied Actual $\bar{R}^2$	2.05%			

Table V
Regressions of Annual Real Index Returns on Stochastic Detrended Lagged Dividend Yield Variables

This table presents OLS regressions of annual real index return on a stochastic detrended dividend yield variable, as recommended by Ferson *et al* (2003b):

$$r_{t+1} = b_0 + b_1 X_T + e_{t+1}$$

where r_{t+1} denotes annual real index return at time $t+1$; annual real index returns are based on two measures: *VWR*, CRSP value-weighted index return; *SPR*, CRSP S&P 500 index return; X_t denotes the stochastic detrended dividend yield and is calculated as:

$$X_t = \frac{D_t}{P_t} - \frac{1}{t} \sum_{j=1,\dots,t} \frac{D_{t-j}}{P_{t-j}}$$

where D_t is the level of real annual dividends during the twelve months preceding time t and P_t is the real stock index level at time t. While different numbers of lags could be used in the detrending, a 12-month lag is used, as recommended by Ferson *et al* (2003b). The sample begins in January 1927 and extends through December 2004. All data are obtained from CRSP. Regressions are estimated by OLS and t-statistics are adjusted for autocorrelation and heteroskedasticity using Newey-West (1987). $\bar{R}^2$ denotes adjusted R^2. Cut-off $\bar{R}^2$ is obtained using a simulation procedure where dependent and independent variables are uncorrelated but have the same autocorrelation properties as the actual data. The final column of the table reports the modified cut-off $\bar{R}^2$ which recognizes the dependency of both the return and dividend yield variables on the share index and dividend levels.

Variable	b_0	$t(b_0)$	b_1	$t(b_1)$	$\bar{R}^2$	Cut-off $\bar{R}^2$	Cut-off $\bar{R}^2$ With Data Mining	Modified Cut-off $\bar{R}^2$
VWR	0.081	4.364	-2.295	-1.011	0.29%	3.94%	7.50%	6.32%
SPR	0.084	4.232	-1.808	-0.893	-0.24%	3.78%	7.20%	6.11%

Table VI
Simulation Results for the Boudoukh *et al* (2004) Model

This table presents simulation results for the Boudoukh *et al* (2004) dividend payout yield model. The model runs OLS regression of excess market returns at time t+1 on the dividend yield and the total payout yield at time t. Following Boudoukn et al. (2004), the analysis are conducted on two sample periods, 1926-2002 and 1926-1985. The excess market return is defined as the difference in the CRSP value-weighted total return and the return on a three-month Treasury-bill; Total payout (CF) yield is the sum of dividends and repurchases of common equity during the year divided by the year-end market capitalization. In the table, the coefficients, t-statistics and R^2 of the regressions are obtained from Table 2 of Boudoukh *et al* (2004). For the simulation, the properties of the excess market return, dividend yield and total payout (CF) yield use descriptive statistics of the variables provided in Table 1 of Boudoukh *et al* (2004). To obtain modified cut-off R^2, the properties of dividend and share price index are calculated using the annual CRSP Value-weighted Index. The total payout (CF) are estimated using the ratios of the aggregate expenditure on the purchase of common and preferred stocks and the aggregate amounts of dividends declared on the common stocks of US firms over the period 1972-2000 provided in Grullon and Michaely (2002). Following Boudoukh et al. (2004), the total payout (CF) yield is assumed equal to the dividend yield prior to 1972 because of negligible repurchase activity. $\bar{R}^2$ denotes adjusted R^2. Since Boudoukh et al. (2004) only reports R^2, $\bar{R}^2$ is derived using the equation, $1-(1-R^2)\frac{T-1}{T-K}$, where T is the number of observations and K is the number of independent variables in the regression. Cut-off $\bar{R}^2$ is obtained using a simulation procedure where dependent and independent variables are uncorrelated but have the same autocorrelation properties as the actual data. The final column of the table reports the modified cut-off $\bar{R}^2$ which recognizes the dependency of both the return and dividend yield variables on the share index and dividend levels.

Independent Variable	n	b_1	$t(b_1)$	R^2	$\bar{R}^2$	Cut-off $\bar{R}^2$	Cut-off $\bar{R}^2$ with Data Mining	Modified Cut-off $\bar{R}^2$
Panel A: 1926-2002								
Dividend Yield	77	0.133	1.522	4.10%	2.82%	4.80%	8.97%	18.75%
Total Payout (CF) Yield	77	0.200	2.699	7.20%	5.96%	4.84%	9.27%	19.86%
Panel B: 1926-1985								
Dividend Yield	60	0.294	3.502	12.90%	11.74%	6.09%	11.20%	22.43%
Total Payout (CF) Yield	60	0.281	3.473	11.50%	10.32%	6.06%	11.33%	20.22%

Table VII
Results for the Marsh and Merton (1987) and Lintner (1956) Dividend Models

This table presents regression results for the Marsh and Merton (1987) and Lintner (1956) dividend behaviour models. The Marsh and Merton (1987) model is tested using the annual CRSP value-weighted and S&P 500 indices for the time period 1927 to 2004, and the Lintner (1956) models are tested using the annual S&P 500 Index for the time period 1927 to 2004. Panel A of the table reports results for the Marsh and Merton (1987) model

$$\log(D_{t+1}/D_t) + D_t/P_{t-1} = y_0 + y_1 \log[(P_t + D_t)/P_{t-1}] + y_2 \log(D_t/P_{t-1}) + e_{t+1}$$

where D_t is the level of real annual dividends during the twelve months preceding time *t* and P_t is the real stock index level at time *t*. Panel B reports results for Lintner (1956) model 1

$$D_t = q_0 + q_1 E_{t-1} + q_2 D_{t-1} + e_t,$$

where E_t denotes the real level of earnings during the twelve months preceding time *t*. Panel C reports results for Lintner (1956) model 2

$$D_t - D_{t-1} = l_0 + l_1 E_{t-1} + l_2 D_{t-1} + e_t.$$

The models in Panels B and C are tested using the S&P 500 index only due to earnings data availability. Regressions are estimated by OLS and figures in parentheses are t-statistics using Newey-West (1987) standard errors. $\bar{R}^2$ denotes adjusted R^2. Modified cut-off $\bar{R}^2$ recognizes the dependency of the dependant and independent variables in the equation.

	Panel A: Marsh and Merton (1987) Model				
	$\log(D_{t+1}/D_t) + D_t/P_{t-1} = y_0 + y_1 \log[(P_t + D_t)/P_{t-1}] + y_2 \log(D_t/P_{t-1}) + e_{t+1}$				
Variable	y_0	y_1	y_2	$\bar{R}^2$	Modified Cut-off $\bar{R}^2$
CRSP Value-Weighted Index	-0.006	0.345	-0.011	33.95%	9.23%
	(-0.053)	(3.196)	(-0.330)		
S&P 500 Index	0.050	0.302	0.006	29.84%	12.04%
	(0.457)	(3.027)	(0.180)		
	Panel B: Lintner (1956) Model 1				
	$D_t = q_0 + q_1 E_{t-1} + q_2 D_{t-1} + e_t$				
Variable	q_0	q_1	q_2	$\bar{R}^2$	Modified Cut-off $\bar{R}^2$
S&P 500 Index	0.819	0.040	0.864	92.99%	84.48%
	(1.868)	(1.213)	(9.845)		
	Panel C: Lintner (1956) Model 2				
	$D_t - D_{t-1} = l_0 + l_1 E_{t-1} + l_2 D_{t-1} + e_t$				
Variable	l_0	l_1	l_2	$\bar{R}^2$	Modified Cut-off $\bar{R}^2$
S&P 500 Index	0.819	0.040	-0.136	3.13%	14.80%
	(1.868)	(1.213)	(-1.556)		

Table VIII
Results on a Reformulated Lintner (1956) Dividend First-Difference Model

This table presents results for the reformulated Lintner (1956) dividend first-difference regression model on the annual CRSP value-weighted and S&P 500 indices for the time period 1927 to 2004:

$$D_t - D_{t-1} = q_0 + q_1(P_{t-1} - P_{t-2}) + q_2(D_{t-1} - D_{t-2}) + (e_t - e_{t-1})$$

where D_t is the level of real annual dividends during the twelve months preceding time t and P_t is the real stock index level at time t. While regressions are estimated by OLS in both Panels A and B, t-statistics are estimated using Newey-West (1987) standard errors in Panel B. In the simulation procedure, the error term $(e_t - e_{t-1})$ is accommodated using a Heteroskedasticity and Autocorrelation Consistent Covariance estimation procedure within Generalized Method of Moments. $\bar{R}^2$ denotes adjusted R^2. Modified cut-off $\bar{R}^2$ recognizes the dependency of the dependant and independent variables in the equation.

Variable	q_0	$t(q_0)$	q_1	$t(q_1)$	q_2	$t(q_2)$	$\bar{R}^2$	Modified Cut-off $\bar{R}^2$
Panel A: Regression Results without Newey-West (1987) Correction								
CRSP Value-Weighted Index	0.075	0.885	0.004	2.665	0.190	1.681	12.17%	7.12%
S&P 500 Index	0.076	0.750	0.003	2.191	0.185	1.648	8.69%	6.91%
Panel B: Regression Results with Newey-West (1987) Correction								
CRSP Value-Weighted Index	0.075	0.822	0.004	1.790	0.190	1.301	12.17%	7.12%
S&P 500 Index	0.076	0.780	0.003	1.401	0.185	1.173	8.69%	6.91%