Dividends, Total Cashflow to Shareholders and Predictive Return Regressions*

Donald Robertson
University of Cambridge

Stephen Wright
Birkbeck College, University of London

March 9, 2005

Abstract

This paper provides new evidence on the predictive power of dividend yields for US aggregate stock returns. Following Miller & Modigliani (1961), we construct a measure of the dividend yield that includes all cashflows to shareholders. We show that this alternative "cashflow yield" has strong and stable predictive power for returns, and appears robust to a battery of tests that have been proposed in recent critiques of the predictability literature.

JEL Classifications: C32, C53, E44, G10, G14.

*We are grateful to the editor and two anonymous referees for helpful comments; also to Nellie Liang of the Division of Research and Statistics at the Federal Reserve Board and Kenneth Petrick of the Bureau of Economic Analysis for advice on data.

1. Introduction

There is a large body of research that claims to find evidence that the dividend yield predicts stock returns. More recently however an increasing body of research has cast doubt on the earlier evidence of predictability, attributing it to data mining or other statistical problems.[1] This paper suggests that the clear weaknesses of the dividend yield as a predictor may be due to mis-measurement. Miller & Modigliani (1961) showed that stock market value depends on investor valuation of all cashflows from firms, not just the dividend component. Since the propensity to pay dividends may vary due to for example taxation changes, dividends alone may at times be a poor proxy for true cashflow. In this paper we use a new "cashflow yield" that includes both dividend and non-dividend cashflows to shareholders and investigate its predictive power for aggregate stock returns.

In redefining dividends in this way our work is related to a number of papers that have investigated non-dividend cashflows in other contexts. Most firm-level studies (e.g., Fama & French, 2001; Grullon & Michaely, 2002; Liang & Sharpe,1999) have focussed on the growing importance of repurchases. However for the representative investor cash- or bond-financed acquisitions and new issues play an identical role in transferring cash from firms to shareholders (or vice versa in the case of new issues), and both have at times been quantitatively as important as dividends and repurchases. A number of authors (Bagwell & Shoven, 1989; Ackert & Smith, 1993; Mehra, 1998; Allen & Michaely, 2002) have noted the importance of treating all such non-dividend cashflows as being equivalent to dividends; but the implications for measures of total cashflow have however received distinctly less attention in econometric research.[2]

We use a new dataset (Wright, 2004) for the US non-financial corporate sector to construct an annual series for total corporate cashflow to shareholders since the start of the twentieth century. We then compare the resulting cashfow yield with standard yield measures, both from this dataset and the more commonly used S&P series. We show that, in contrast to conventional yield measures, the

[1] On predictability, see *eg* Fama & French, 1988; Jegadeesh (1990); Campbell and Shiller, 1988; 1998; Pesaran & Timmerman,1995; or the survey in Campbell *et al*, 1997. For revisionist critques see *eg* Goetzmann & Jorion 1993; Nelson & Kim, 1993; Kirby, 1997; Bossaerts & Hillion, 1999; Foster, Smith & Whaley, 1997; Stambaugh, 1999; Goyal & Welch, 2003, 2004, Ang and Bekaert 2004.

[2] A point stressed in Allen & Michaely's (*op cit*) recent comprehensive review article. The single notable exceptions that we are aware of are Ackert and Smith (*op cit*) and Mehra (*op cit*).

cashflow yield has strong and stable predictive power for returns at a range of horizons, and is robust to a battery of tests that have been proposed in recent critiques of the predictability literature.

2. Data

2.1. Data Sources and Construction

All data used in this paper come from a new annual dataset described in full in Wright (2004), that relates to the total non-financial US corporate sector (rather than the more commonly used subset of quoted companies) over the sample 1900-2002, using data from the Federal Reserve's Flow of Funds Tables, and Bureau of Economic Analysis data where these exist, and such historical sources as are available in earlier periods.

The core series[3] used in this paper are:

- **Real Market Value of Equities (V_t):** From 1945-2002 the nominal value of this series is taken directly from the Federal Reserve's Flow of Funds Accounts, for the market value of equities outstanding for the non-financial corporate sector (Table B102, line 34). It includes an adjustment that nets out intercorporate cross-holdings (Federal Reserve, 2000). Before 1945, Wright (*op cit*) describes the construction of this series using a combination of two proxies derived from S&P 500 and Cowles (1938) data on returns and dividend yields, in conjunction with the dividend and new issue data described below.

- **Real Dividends (D_t):** From 1946 onwards, the nominal value of this series equals non-farm, non-financial dividends from Flow of Funds Table F102, Line 3; and from 1929-45 the (virtually identical) series from the National Income and Product Accounts Table 1.16, for the total non-financial corporate sector. Before 1929, the series is constructed using data from Kuznets (1941), and Goldsmith (1955) and Cowles (1938) (for details, see Wright, *op cit*). All series net out inter-corporate dividend payments, and thus are consistent with market value data.

[3] All three series are deflated by the consumer price index (also taken from Wright, *op cit*'s database) to derive real values.

- **Real Net New Issues (N_t):** From 1946-2002, the nominal value of this series equals (net) non-financial corporate equity issues from Flow of Funds data Table R102 (line 11). In recent years, these figures have been consistently negative, implying net corporate purchases, due to the combined impact of repurchases and cash-financed mergers and acquisitions. Before 1946, data on these last two components are not available, but they are assumed to be empirically negligible.[4] Wright (*op cit*) constructs a series for new issues for this earlier period with data from various sources (Miller, 1963; *Historical Statistics*; and, for the first decade of the twentieth century, from editions of the *Commercial and Financial Chronicle).*

Wright (*op cit*) derives implied series for the aggregate real stock price index, the aggregate real return for the total non-financial corporate sector, and real dividends per share, (none of which are directly published), all of which can be derived from the three core series above. Total real cashflow can also be derived as $C_t = D_t - N_t$, and the real non-financial return by $1 + R_t = \frac{V_t + C_t}{V_{t-1}}$.[5]

2.2. Alternative Measures of the Dividend Yield

Figure 1 shows the conventional measure of the non-financial dividend yield (dividend per share to price ratio), alongside our alternative "cashflow yield" (defined as C_t/V_t) using our dataset, over the course of our sample. For comparison we also show the yield on the S&P Composite index. The two conventional measures, are, as might be expected, very similar.[6] The cashflow yield has a very similar mean to the conventional yield; but at times distinctively different properties. It is noteworthy that these differences were not just evident in the last two decades of the sample.

[4] Allen and Michaely (2002) note that before 1983 repurchases were barely legal, and as a result very uncommon. In the period of overlap with Fed data the alternative sources for new issues that we rely on in earlier periods yield very similar figures, suggesting that the omission of cash-financed acquisitions before 1946 is not empirically significant either.

[5] Following Miller & Modigliani (*op cit)* Wright (*op cit*) shows that this is identical to the more common definition using dividends per share.

[6] Following standard practice (eg, Shiller, 2000; Goyal & Welch, 2004) the S&P yield is extrapolated backwards before 1925 using the equivalent series from Cowles (*op cit).* The decline in the yield in the 1990s for all nonfinancial companies was not as marked as for the S&P 500 companies, largely due to a distinct divergence in tax incentives for smaller companies, which encouraged 100% payout ratios.

For much of the sample, the difference between the two series reflected distinct surges in new issues at certain periods (most strikingly in 1929, and in the early 1970s) that lowered the cashflow yield significantly, by lowering the net transfer of cash from firms to shareholders. In other periods (most notably the 1930s and early 1940s), new issues essentially collapsed to zero and the cashflow and conventional yields were nearly identical. However, in the last two decades of the century there was a distinct shift, with the difference between the two yields switching sign, as firms engaged in significant levels both of repurchases, and geared acquisitions, that more than offset minimal levels of new issues. The impact of the implied adjustment to the dividend yield in recent years is distinctly more significant than in estimates based solely on data for repurchases, as in, *e.g.*, Fama & French (2001); Liang & Sharpe (1999). While there are data coverage differences, the primary explanation is the impact of cash-financed acquisitions in the Fed data.[7]

The chart also shows that, while the cashflow yield is distinctly more volatile than the per share yield, it appears to have a stronger tendency to mean reversion than either of the two conventional measures. This is important because persistence of the dividend yield has been pointed to as a cause for the inferential problems in predictive regressions.

The downward drift in both conventional measures of the dividend yield in the latter part of the sample at least in part reflected the surge in the stock market during the course of the 1990s. Strikingly, however, this tendency was not evident in the cashflow yield, which, at the peak of the market in 2000, was close to its mean, since cashflow from the corporate sector to equity holders grew as rapidly as the stock market during the 1990s, due to the strength of M&A activity and repurchases.

Robertson and Wright (2004) argue that there are good theoretical grounds for expecting stronger evidence of mean reversion for the cashflow yield than for conventional measures. They show that the mean value of the conventional yield may be subject to permanent shocks if there are permanent shifts between dividend and non-dividend methods of cash transfer to shareholders, as Figure 1 strongly suggests has been the case. The mean cashflow yield will however be immune to such shifts. Empirically this seems to be borne out in our dataset: the

[7] The Federal Reserve do not publish a breakdown of net aggregate corporate equity purchases into new issues, repurchases and geared M&A. Allen & Michaely (*op cit*) provide aggregate data on all three elements for quoted companies that show similar patterns to those in the Fed data. Wright (*op cit*) provides comparative analysis of the Allen & Michaely dataset and Fed data.

AR1 coefficient for the S & P and non-financial dividend yield are 0.87 and 0.81 respectively, whereas that for the cashflow yield is only 0.63. Robertson & Wright (*op cit*) show that the unit root restriction cannot be rejected for conventional yields, but is strongly rejected for the cashflow yield.[8]

3. Predictive Return Regressions

Most tests of predictability rely on regressions of the form

$$r_{t,h} = \alpha + \beta_h x_t + u_{t,h} \tag{3.1}$$

$$x_t = \gamma + \lambda x_{t-1} + v_t \tag{3.2}$$

where $r_{t,h}$ is the h-period ahead log return at time t; and x_t is some predictor variable (most commonly the dividend yield) observed at t. The hypothesis of interest is typically a test of $H_0 : \beta_h = 0$, with a rejection being interpreted as evidence of predictability. A number of problems with inference in this framework have been pointed out. First, if $h > 1$ then $u_{t,h}$ will usually be serially correlated due to overlapping observations and this will affect estimates of the standard error of the estimate of β_h. Second, if we search for predictability at various different horizons h we need to take account of the multiplicity of tests and look at the implied joint hypothesis. Third, since Stambaugh (1999) it is necessary to take account of the time series properties of the predictor variable, since biases in the estimation of λ are transmitted to estimates of β_h if as typically happens there is a correlation between v_t and $u_{t,h}$. Together these problems mean that the conventional t-statistic will not be reliable since the point estimate may be biased and the OLS standard errors incorrect. Solutions to the problem of serial correlation in the residuals have been much discussed, usually by correcting the estimated standard errors through some Newey-West type adjustment - Ang and Bekaert (*op cit*) argue that Hodrick (1992) standard errors provide the most reliable inference and we present these. The estimates of β_h can also be bias-adjusted (see Stambaugh (1999); Lewellen (2003); Campbell & Yogo (2002)). In this paper however we follow Nelson and Kim (1993) and Ang and Bekaert (2004) in relying primarily on simulation methods to obtain p-values which will largely correct for these difficulties.

[8]See Appendix B for discussion of the impact on AR1 coefficients of imposing the null of no predictability.

In Table 1 we report the results of estimating the equation for $h = 1, \ldots, 10$ years where $r_{t,h} = \frac{1}{h}\Sigma_{i=1}^{h} r_{t+i}$ is the average $h-$period ahead return; $r_t = \log(1+R_t)$ is the one-period log return and and x_t is one of the three log dividend yield measures. We report p-values for the test of the null hypothesis $H_0 : \beta_h = 0$, using both OLS and Hodrick (1992) standard errors. We also report Monte Carlo derived p-values obtained by simulating the set of equations under the null $\beta_h = 0$ for all h, thus dealing simultaneously with the bias and overlapping observation problems, and also bootstrapped p-values where the actual residuals (under the null) are sampled to generate the simulated data. Given the known problems of focussing on results for individual horizons, the final column of Table 1 reports conventional and simulated p-values for joint tests of the null of no predictability at all horizons from 1 to 5, and from 1 to 10 years.[9]

Panels A and B of Table 1 replicate the known results on the fragility of the evidence that the conventional dividend yield predicts returns, with very similar results for both measures. The estimated coefficients show the well-known horizon effect - the OLS p-values drop with horizon such that a null-hypothesis of no predictability would be rejected strongly at conventional significance levels at longer horizons. However, p-values using Hodrick standard errors show the importance of controlling for serial dependence in the long horizon prediction regressions: the extent of predictability is brought down to at best marginal significance. The simulated p-values reinforce this conclusion, and indeed show that even the Hodrick correction understates the size distortion. Concerns about non-normality of the data are shown to be of little consequence as the Monte Carlo and bootstrapped p-values differ only marginally. For neither measure do simulated p-values fall below 5% at any individual horizon; and the more robust joint tests fail to reject the null of no predictability once the (again, very significant) size distortion is corrected.

In contrast, our alternative cashflow yield demonstrates much more robust performance. Monte Carlo results show that the size distortion of OLS, and Hodrick p-values remains, but even allowing for this the null of no predictability is strongly rejected at all individual horizons and in both joint tests. A further contrast with conventional yields is that the strongest predictive power is evident at shorter horizons.

Figures 2 and 3 use recursive OLS regressions with a sample starting with 20 observations and expanding up to the full sample to show that the predictive power of the cashflow yield has been evident on a consistent basis over a wide

[9] For details of simulations and the joint testing procedure, see Appendices B and C.

range of sample periods, and therefore does not appear to suffer from look-ahead bias.[10] Had data on the cashflow yield been available, it would have shown statistically significant evidence of predictive power from the 1950s onwards, and that increasing the available data has reinforced, rather than undermined this evidence. The chart shows that this is again in stark contrast to conventional yield measures. These results are also not due to the choice of starting point for the regressions, experiments with data post 1929, post 1945 and excluding the 1990s show a very similar pattern.

4. Conclusions

Recent research has significantly undermined the evidence that conventional dividend yields predict aggregate stock returns. This paper does not take issue with this revisionist view, but shows that robust evidence of predictability is restored if we use a new "cashflow" yield that aggregates dividend and non-dividend cashflows. Since this alternative measure is also clearly more in line with an economically meaningful measure of corporate cashflows to shareholders (as in Miller & Modigliani, 1961) we do not believe that we can be accused of redefining the data to achieve the desired result. It should be stressed that while the predictive power of the cashflow yield is stable and statistically significant, the associated predictive regressions do not have very high R^2 values, so that even supposing such predictability to reflect a degree of market inefficiency (which is of course an open question in itself) any implied trading strategy that exploited this predictability would itself be very risky. Nonetheless, rumours of the death of the predictability literature do appear to have been exaggerated.

[10] The recursive t-statistics shown in Figure 2 are monotonic transformations of Goyal and Welch's (2004) relative cumulative sum of squares statistics shown in Figure 3, but are arguably more readily interpretable. See Appendix D for further discussion of related issues.

Appendix

A. Correlations Across Different Datasets

The unconditional correlations between the key data series used in predictive regressions are as follows:

Unconditional Correlations					
	return	S&P return	cashflow yield	non-fin div yield	S&P div yield
return	1.000	0.976	-0.281	-0.355	-0.314
S&P return		1.000	-0.281	-0.341	-0.310
cashflow yield			1.000	0.632	0.475
non-fin div Yield				1.000	0.898
S&P div yield					1.000

Notes
(i) Sample period 1901-2001.
(ii) Return is log real return on our dataset, S&P return is log real return on S&P; yields are all in logs

The table above shows that the two measures of real returns are very highly correlated, and the two comparable measures of the conventional dividend yield only somewhat less so The second table, below, shows conditional correlations, i.e. between innovations to each of the 5 series, for simplicity assuming each to be generated by equations of the form as in (3.1), under the null of no predictability. This reveals that the conditional correlation between the two conventional yield series is even stronger than the unconditional correlation (which is lowered primarily by the somewhat lower persistence of the nonfinancial yield, as noted in the main text). It also shows the very high negative correlation between innovations to conventional yields and to returns; in contrast the conditional correlation between the cashflow yield and returns is distinctly less strong, thus reducing (albeit not entirely eliminating) the associated bias problems.

Correlations between Innovations					
	return	S&P return	cashflow yield	non-fin div yield	S&P div yield
return	1.000	0.976	-0.627	-0.849	-0.838
S&P return		1.000	-0.626	-0.858	-0.834
cashflow yield			1.000	0.683	0.717
non-fin div Yield				1.000	0.945
S&P div yield					1.000

B. Monte Carlo and Bootstrapped Simulations

The p-values for the point estimates of the β_h in (3.1) are derived by generating 10,000 samples, each of 103 annual observations, of the system of equations given by (3.1), for $h = 1$ and (3.2), with β_1 set equal to zero. The models for the data generating process under the null, using the three alternative measures of x_t, were estimated over the sample 1901-2002, and for each replication equations of the form in (3.1) were estimated for $h = 1..10$. The (two-sided) p-values reported are the proportion of samples in which the squared value of $\widehat{\beta}_h$ exceeded the value derived from the historic sample.

The advantage of the Monte Carlo approach is that it deals simultaneously with the bias and overlapping observation problems. One potential shortcoming of the Monte Carlo approach arises if the generating mechanism of the true data is not well approximated by that chosen (for instance if x_t is not in fact an AR1); however we found no evidence in our data that allowing for higher order processes for x_t made a substantial difference to our results.[11] The Monte Carlo approach also assumes there is no conditional heteroscedasticity in the underlying innovation sequence, but on our annual dataset the null of homoscedastic errors cannot be rejected. Finally the Monte Carlo simulations use normally generated error sequence (with covariance structure matching that of the data under the null hypothesis). There is some evidence of non-normality in the data so we also report Monte-Carlo using bootstrap residuals, that is the actual residuals from the system (3.1) are sampled randomly (with replacement) to generate the simulated data. This will ensure that the simulated data match any non-normality in the data.

It is worth noting that the estimated models used as the DGP in simulations all imply distinctly higher persistence of the three yield measures than when these

[11] The (partial) correlograms of all three yield series fall essentially to zero after the first lag.

are estimated as single equations. The AR1 parameters, for the cashflow, nonfinancial and S&P yield rise from (0.63,0.81,0.87) when estimated by single equation methods,to (0.79,0.90,0.95) when estimated jointly with return regressions under the null of no predictability. Thus imposing the null of no predictability is almost equivalent to imposing unit roots in conventional yields. This is what would be expected if (as is the case) the Campbell-Shiller (1988) loglinear approximation for returns is close to holding exactly, and if conventional dividend growth is close to being unforecastable (this reflected in the very strong negative conditional correlations between conventional yields and returns shown in the previous appendix). Even under the null of no predictability, however, the cashflow yield remains clearly stationary, since the cashflow yield has significant predictive power for cashflows, as well as for returns.

C. Joint Tests

Our joint test procedure exploits an equivalence between two ways of representing horizon return predictability. The standard approach in (3.1) is in terms of the h-period real return, $r_{t,h} = \frac{1}{h}\Sigma_{i=1}^{h} r_{t+i}$ where r_t is the one period real return, but this has the standard problem that the errors are MA processes. The alternative representation is

$$r_{t+h} = \gamma_h x_t + \omega_{t,h} \tag{C.1}$$

which has the advantage that under the null of no predictability the γ_h are zero and the $\omega_{t,h}$ are white noise. Since the implied coefficients in (3.1) can be derived by $\beta_h = \Sigma_{i=1}^{h}\gamma_i$, the joint null $H_0 : \beta_h = 0$ for all $h = 1..H$ is equivalent to the joint null $H_0 : \gamma_h = 0$ for all $h = 1..H$. We can estimate equations of the form in (C.1) as a system, and since all equations in the system have the same regressor, FIML, SUR and OLS estimates will be the same, and will imply identical estimates of the β_h to those derived by direct estimation. We therefore test the joint null by estimating restricted and unrestricted systems for a given H, and test the null by likelihood ratio. Given bias problems caused by the correlation of x_t with r_t, the size of the resulting test statistic will be incorrect, but we again run Monte Carlo and bootstrapped simulations under the null of no predictability (again with 10,000 replications) to find the true size.

D. Recursive and Sub-Sample Estimation

Whilst Table 1 shows that the cashflow yield maintains predictability even when we account for serial correlation properties, there are still possible problems with data mining by choice of sample period or look ahead bias. Would an investigator analysing this data earlier in the sample period have found significant predictive ability from the cashflow yield, or is the predictability evident in Table 1 an artificial construct of the period chosen? To investigate these possibilities we focus on one step ahead predictive regressions (i.e. the horizon h is one). We run recursive OLS estimations and Figure 2 shows the resulting recursive t-statistic on the predictor coefficient as the sample size increases (from an initial window of 20 observations). There is significant predictive ability from the cashflow yield throughout the sample, in contrast to the other measures of dividend yield. An alternative diagnostic comparing the predictive ability of each of the dividend measures relative to a constant expected mean forecast is shown in Figure 3. This shows the cumulative sum of squared one step ahead recursive residuals from an equation using the predictor variable (the conventional or cashflow yield) relative to the cumulative sum of squared recursive residuals from a regression on a constant (as suggested by Goyal and Welch 2004). The interpretation is that when the line rises the residuals from the predictor variable are smaller than those from the constant mean prediction, indicating additional predictive power. The S&P and conventional dividend yields gain very much (relative to the constant expected return benchmark) around the crash of 1929 but then lose this advantage in the 30s. They then produce smaller prediction errors until about 1953 when again all advantage is lost by the 70s. The 71-73 crash is again good for the conventional yield, but after that the performance is pretty feeble through to the end of the sample, essentially the prediction using the conventional dividend yield being little better than a constant expected mean prediction. This plot is very similar to that in Goyal and Welch (their Figure 1 Panel A particularly in the post 1940 period - though with a mean shift) even though they are looking at the equity premium rather than the real return. By contrast the cashflow yield also gains in 1929, but then produces roughly comparable prediction errors from 1930 to about 1970, again predicts rather better through the 71-73 crash and then gains almost monotonically in the post 1973 era (the upward sloping line indicating that the one step ahead prediction errors are almost uniformly smaller than those from the constant mean regression in this period). This general pattern is repeated if we carry out the same experiment at longer forecast horizons than one year.

In Table 2 we provide a full set of estimates of β_h coefficients, for $h = 1, .., 10$, in horizon regressions of the form in (3.1), and the same set of diagnostic statistics as in Panel C of Table 1, for three different sub-samples. The first two subsamples are chosen to reflect discontinuities in underlying data sources used in the Wright (*op cit*) dataset: BEA national income statistics only become available from 1929 onwards, and Flow of Funds statistics from 1946 onwards. Additionally, Panel C shows results for the commonly used sample (cf Lewellen, Goyal & Welch, Ang & Bekaert(*op cit)*) from 1963 onwards. Point estimates of coefficients at different horizons are very similar in all three subsamples to those shown in Table 1, Panel C, for the full sample. Monte Carlo and bootstrapped p-values are somewhat higher than over the full sample, but are still almost invariably significant for all horizons at conventional levels even in the shortest sample. The joint tests continue to reject the null of no predictability at at least the 5% level, except (barely) in the case of the very short sample shown in Panel C.

References

[1] Ackert, L and Smith, B (1993) " Stock price volatility, ordinary dividends, and other cash flows to shareholders", *Journal of Finance* 48 (4) 1147-1160

[2] Allen, F and Michaely, R (2002), "Payout Policy", in George Constantinides, Milton Harris and Rene Stulz (eds): *North-Holland Handbooks of Economics* (Elsevier, Amsterdam)

[3] Ang, A and Bekaert, G (2004) "Stock Return Predictability: Is it There?" Discussion Paper Columbia University

[4] Bagwell, L and Shoven, J (1989) " Cash Distributions to Shareholders", *Journal of Economic Perspectives*, 3 (3) pp 129-140

[5] Bossaerts, P and Hillion, P (1999), "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn", *Review of Financial Studies,* Summer 1999, vol 12, no. 2, pp 405-28

[6] Bureau of Economic Analysis (1977) *Historical Statistics of the United States, Colonial Times to 1970*

[7] Campbell, J and Shiller, R (1988) "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies*, 1. 195-227

[8] Campbell J Y and Shiller R J (1998) "Valuation Ratios and the Long-Run Stock Market Outlook", *Journal of Portfolio Management*, Fall 1998

[9] Campbell, J, Lo, A J and MacKinlay, A C (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ

[10] Campbell, J and Yogo, M (2002), "Efficient Tests of Stock Return Predictability", Harvard University *Mimeo*

[11] Cowles, A C (1938) *Common-Stock Indices* Principia Press, Bloomington Indiana

[12] Fama, E., and French K (1988) "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 22(1), 325.

[13] Fama, E and French, K (2001) "Disappearing Dividends: Changing firm characteristics or lower propensity to pay?", *Journal of Financial Economics*, 60, pp 3-43

[14] Board of Governors of the Federal Reserve System (2000), *Guide to the Flow of Funds Accounts*, Washington DC

[15] Foster, F D, Smith, T and Whaley, (1997) R "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2", *Journal of Finance,* Vol LI No 2.

[16] Goetzmann, W. N., and P. Jorion (1993) "Testing the Predictive Power of Dividend Yields", *Journal of Finance*, 48(2), pp 663-679.

[17] Goyal, A and I Welch (2003) "Predicting the Equity Premium with Dividend Ratios", *Management Science* 49(5), 639-654

[18] Goyal, A and I Welch (2004) "A comprehensive look at the empirical performance of equity premium prediction", NBER Working Paper No 10483.

[19] Goldsmith, R W (1955) *A Study of Saving in the United States*, Vol I, Princeton University Press, Princeton, New Jersey

[20] Grullon, G, and Michaely, R (2002), "Dividends, Share Repurchases and the Substitution Hypothesis", *Journal of Finance*, Vol 57, No 4, pp1649-1684

[21] Hodrick, R J (1992) "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement" *Review of Financial Studies*, 5, 3, 357-386

[22] Jegadeesh, N (1990), "Evidence of Predictable Behavior of Security Returns", *Journal of Finance* 45(3), pages 881-98.

[23] Kirby, C (1997), "Measuring the Predictable Variations in Stock and Bond Returns", *Review of Financial Studies,* Fall 1997 Vol 10, no. 3, pp 5-630

[24] Kuznets, S (1941) *National Income and Its Composition, 1919-1938* Vol I, National Bureau of Economic Research, New York

[25] Lewellen, J (2003), "Predicting Returns with Financial Ratios", forthcoming *Journal of Financial Economics*

[26] Liang, J N and Sharpe, S A (1999) "Share Repurchases and Employee Stock Options and their Implications for S&P 500 Share Retirements and Expected Returns" , Board of Governors of the Federal Reserve System Discussion Paper.

[27] Mehra, R (1998) "On the volatility of stock prices: an exercise in quantitative theory" *International Journal of Systems Science*, Vol 29, 11, pp 1203-1211

[28] Miller, M H (1963) "The Corporation Income Tax and Corporate Financial Policies" in *Commission and Money and Credit Stabilization Policies*, Prentice-Hall, Englewood Cliffs, NJ

[29] Miller, M H and Modigliani, F, (1961), "Dividend policy, growth, and the valuation of shares", *Journal of Business*, 34, pp 411-433.

[30] Nelson, C. R., and M. J. Kim (1993) "Predictable Stock Returns: The Role of Small Sample Bias", *Journal of Finance*, 48(2), 641661.

[31] Pesaran, M H & Timmermann,A F (1995) "Predictability of Stock Returns: Robustness and Economic Significance" *Journal of Finance*; 50(4), pp 1201-28

[32] Robertson, D and Wright, S (2004) "Miller & Modigliani, Predictive Return Regressions and Cointegration" *Mimeo*, University of Cambridge and Birkbeck College, University of London

[33] Shiller, R (2000) *Irrational Exuberance*, Princeton University Press

[34] Stambaugh, R F. (1999) "Predictive regressions" *Journal of Financial Economics* 54 pp 375-421

[35] Wright, S (2004) "Measures of Stock Market Value and Returns for the US non-financial Corporate Sector, 1900-2002", *Review of Income and Wealth* 50 (4) 561-584

Table 1. Testing Predictive Return Regressions at Different Horizons: 1901-2002

			p-values					
h	Obs	coeff	OLS	Hodrick	Monte-Carlo	Bootstrap	Joint Tests	
Panel A: S & P real return and dividend yield								
1	101	0.0853	0.087	0.059	0.136	0.133	$h = 1, \ldots, 5$	
2	100	0.100	0.007	0.046	0.083	0.079	p_{sur}	0.003
3	99	0.076	0.010	0.113	0.150	0.147	p_{mc}	0.108
4	98	0.067	0.012	0.173	0.182	0.185	p_{bs}	0.100
5	97	0.074	0.003	0.159	0.140	0.139		
6	96	0.070	0.003	0.187	0.154	0.153	$h = 1, \ldots, 10$	
7	95	0.066	0.002	0.167	0.162	0.158	p_{sur}	0.001
8	94	0.079	0.000	0.058	0.085	0.083	p_{mc}	0.154
9	93	0.079	0.000	0.037	0.077	0.074	p_{bs}	0.151
10	92	0.075	0.000	0.033	0.078	0.081		
Panel B: non-financial real return and dividend yield								
1	100	0.095	0.095	0.080	0.129	0.135	$h = 1, \ldots, 5$	
2	99	0.119	0.004	0.020	0.066	0.067	p_{sur}	0.006
3	98	0.090	0.004	0.045	0.121	0.120	p_{mc}	0.088
4	97	0.071	0.008	0.112	0.191	0.188	p_{bs}	0.084
5	96	0.072	0.003	0.093	0.171	0.168		
6	95	0.064	0.003	0.134	0.210	0.201	$h = 1, \ldots, 10$	
7	94	0.055	0.005	0.151	0.263	0.257	p_{sur}	0.006
8	93	0.062	0.001	0.047	0.205	0.202	p_{mc}	0.088
9	92	0.059	0.001	0.044	0.213	0.209	p_{bs}	0.084
10	91	0.053	0.001	0.070	0.250	0.245		
Panel C: non-financial real return and cashflow yield								
1	101	0.144	0.001	0.005	0.003	0.003	$h = 1, \ldots, 5$	
2	100	0.151	0.000	0.000	0.001	0.001	p_{sur}	0.000
3	99	0.112	0.000	0.001	0.003	0.003	p_{mc}	0.004
4	98	0.086	0.000	0.002	0.012	0.012	p_{bs}	0.004
5	97	0.082	0.000	0.000	0.011	0.010		
6	96	0.071	0.000	0.000	0.019	0.017	$h = 1, \ldots, 10$	
7	95	0.061	0.000	0.000	0.040	0.033	p_{sur}	0.000
8	94	0.063	0.000	0.000	0.025	0.020	p_{mc}	0.009
9	93	0.060	0.000	0.000	0.027	0.021	p_{bs}	0.009
10	92	0.057	0.000	0.000	0.026	0.022		

Table 1: Estimates of the h-period return regression $r_{t,h} = \alpha + \beta_h x_t + u_{t,h}$ where $r_{t,h} = \frac{1}{h}\Sigma_{i=1}^{h} r_{t+i}$ is the average real log return over the next h years, r_t is the one period real log return and x_t is one of the three yield measures. We report the OLS estimate of the coefficient β_h at various horizons, and the the p-values of the test of the null hypothesis $H : \beta_h = 0$ using OLS standard errors, Hodrick (1992) autocorrelation corrected standard errors, p-values from Monte Carlo simulations of the model under the null where x_t follows an AR1 $x_t = \gamma + \lambda x_{t-1} + v_t$ and the residuals are assumed normal with covariance structure matching the data, and finally a bootstrap simulation where the residuals are sampled with replacement from the actual equation residuals. The simulations are based on 10,000 repetitions. The final column reports p-values of the joint tests that $\beta_h = 0$ for horizons $h = 1, \ldots, H$ for $H = 5$ and 10 as described in Appendix C: p_{sur} is the conventional p-value for the LR statistic $(\chi^2 (H))$ when the the restricted and unrestricted systems are estimated by SUR; p_{mc} and p_{bs} are implied p-values from Monte Carlo and bootstrapped simulations of the null model as given above.

Table 2. Testing Predictive Return Regressions Using Cashflow Yield over Subsamples

			p-values					
h	Obs	coeff	OLS	Hodrick	Monte-Carlo	Bootstrap	Joint Tests	
Panel A: 1929-2002								
1	73	0.154	0.001		0.006	0.005	$h = 1, \ldots, 5$	
2	72	0.152	0.000		0.002	0.002	p_{sur}	0.000
3	71	0.119	0.000		0.006	0.007	p_{mc}	0.009
4	70	0.089	0.000		0.024	0.026	p_{bs}	0.009
5	69	0.082	0.000		0.026	0.030		
6	68	0.068	0.000		0.053	0.055	$h = 1, \ldots, 10$	
7	67	0.062	0.000		0.065	0.068	p_{sur}	0.000
8	66	0.065	0.000		0.042	0.045	p_{mc}	0.023
9	65	0.062	0.000		0.043	0.046	p_{bs}	0.021
10	64	0.061	0.000		0.037	0.039		
Panel B: 1946-2002								
1	56	0.142	0.002		0.010	0.017	$h = 1, \ldots, 5$	
2	55	0.136	0.000		0.008	0.014	p_{sur}	0.000
3	54	0.113	0.000		0.016	0.022	p_{mc}	0.020
4	53	0.090	0.000		0.035	0.045	p_{bs}	0.021
5	52	0.090	0.000		0.027	0.034		
6	51	0.085	0.000		0.027	0.035	$h = 1, \ldots, 10$	
7	50	0.084	0.000		0.023	0.028	p_{sur}	0.000
8	49	0.083	0.000		0.019	0.025	p_{mc}	0.021
9	48	0.084	0.000		0.013	0.020	p_{bs}	0.020
10	47	0.085	0.000		0.010	0.015		
Panel C: 1963-2002								
1	39	0.149	0.006		0.027	0.033	$h = 1, \ldots, 5$	
2	38	0.146	0.000		0.015	0.023	p_{sur}	0.001
3	37	0.123	0.000		0.024	0.033	p_{mc}	0.044
4	36	0.097	0.000		0.055	0.061	p_{bs}	0.020
5	35	0.095	0.000		0.057	0.053		
6	34	0.090	0.000		0.057	0.054	$h = 1, \ldots, 10$	
7	33	0.088	0.000		0.039	0.050	p_{sur}	0.000
8	32	0.084	0.000		0.040	0.050	p_{mc}	0.058
9	31	0.086	0.000		0.029	0.039	p_{bs}	0.058
10	30	0.086	0.000		0.023	0.029		

Table 2: Estimates of the h-period return regression $r_{t,h} = \alpha + \beta_h x_t + u_{t,h}$ where x_t is the cashflow yield, over different sub-samples. The elements of the table are defined as in Panel C of Table 1.

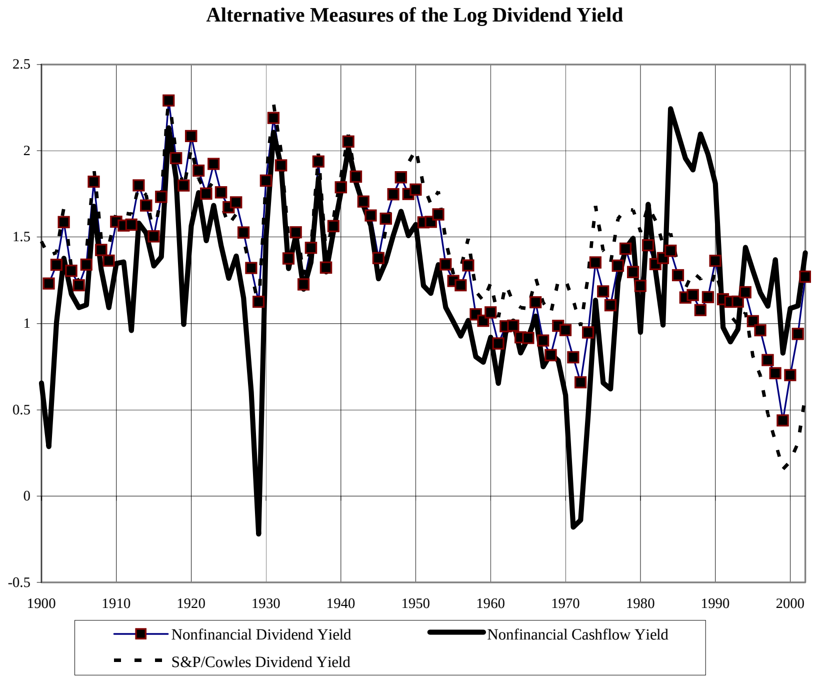

Figure 1 Alternative measures of the dividend yield

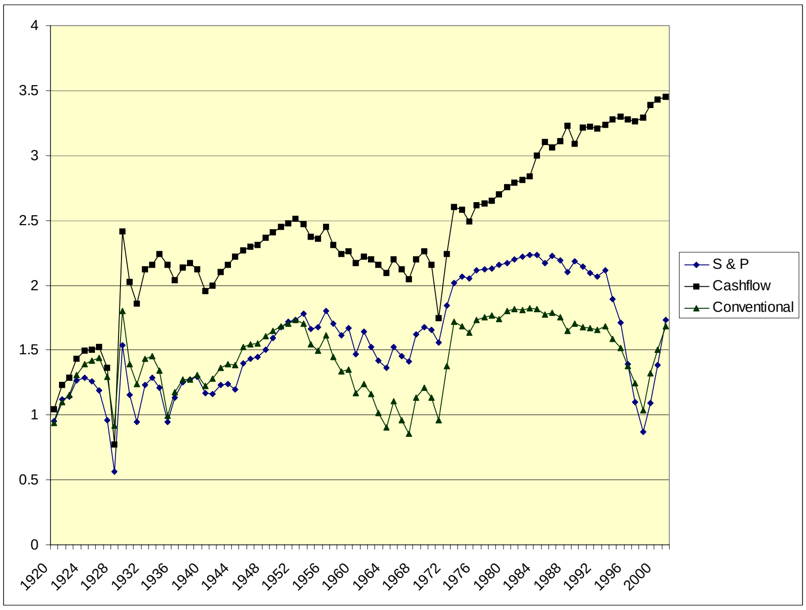

Figure 2 Recursive t-statistics from predictive regressions

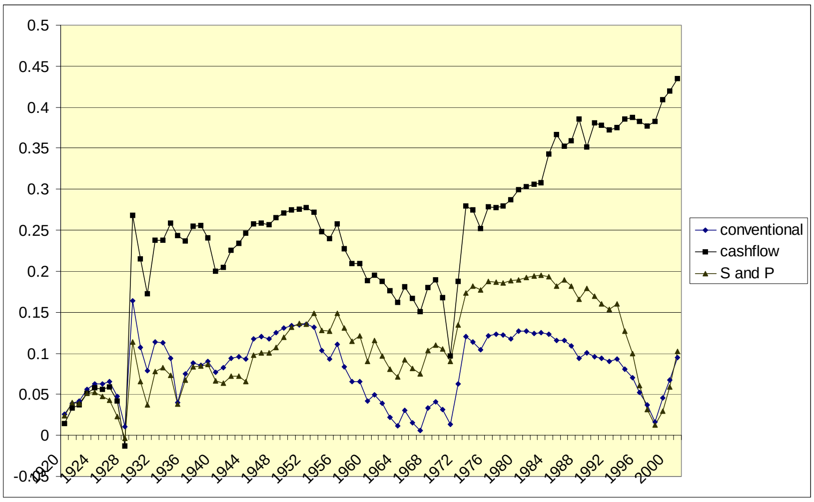

Figure 3 Recursive sums of squared residuals from predictive regressions relative to constant expected mean predictions