The Predictive Power of Dividend Yields for Stock Returns: Risk Pricing or Mispricing? *

Glenn Boyle†
Department of Economics and Finance
University of Canterbury

Yanhui Li
Department of Economics and Finance
University of Canterbury

November 20, 2013

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT QUOTE

*Inspiration for this study came from John Cochrane’s 2012 JB Condliffe Memorial Lecture at the University of Canterbury. We are grateful to Martin Lally for providing us with essential data.
†Corresponding author: Private Bag 4800, Christchurch 8140, New Zealand. Phone: 64-3-364-3479. Email: glenn.boyle@canterbury.ac.nz
The Predictive Power of Dividend Yields for Stock Returns: Risk Pricing or Mispricing?

Abstract

If observed predictability is primarily due to market mispricing, then the relationship between dividend yields and future stock returns should be strongest in market environments that face the greatest ‘limits to arbitrage’. However, using 1931-2012 data from New Zealand, we find that the proportion of dividend yield volatility attributable to expected return variation is greater in the post-liberalisation period, when there should be less scope for mispricing. Although other factors could also be at work, this result poses a challenge to mispricing explanations of stock return predictability.

JEL classification: G12, G14

Keywords: return predictability; discount rates; limits to arbitrage
The Predictive Power of Dividend Yields for Stock Returns: Risk Pricing or Mispricing?

1 Introduction

There is much evidence to suggest that the dividend-price ratio has strong predictive power for future stock returns – see, for example, Cochrane (2008, 2011). Less clear, however, is the proper interpretation of this relationship. On the one hand, a low dividend-price ratio may indicate that perceived future risk is low, causing investors to set prices so that risk premia, and hence expected returns, are low. Subsequent average returns then turn out to be low, as ‘predicted’ by the earlier dividend-price ratio. On the other hand, a low dividend-price ratio may simply reflect excessive optimism on the part of investors who ‘over-price’ stocks relative to their intrinsic value. As the mispricing dissipates, subsequent returns are low. Either story is consistent with the observed power of dividend-price ratios to forecast future returns.¹

In this paper, we propose a simple approach for distinguishing between these two hypotheses. If observed predictability is primarily due to mispricing, then the relationship between dividend yields and future stock returns should be strongest in market environments most conducive to mispricing, i.e., markets in which frictions, regulations, illiquidity, and financing constraints impose ‘limits to arbitrage’ and so inhibit the incorporation of information in prices. Markets in which information is more easily incorporated in prices, resulting in fewer mispricing opportunities, should, ceteris paribus, see weaker return predictability. If, on the other hand, observed predictability is primarily due to time varying risk premia, then the relationship between dividend yields and future stock returns should be unaffected by any change in the potential for mispricing.

We implement this idea using 1931–2012 stock market data from New Zealand (NZ). Prior to July 1984, New Zealand had limited and highly regulated financial markets. Short-selling was prohibited, as were company stock repurchases; there were also restrictions on interest rates, on private overseas borrowing, on foreign-owned companies access to domestic financial markets, and on the ability of NZ residents to purchase foreign exchange for investment purposes. All of these regulations effectively restricted information flows and arbitrage, and so encouraged mispricing. By March 1985, most had been removed; in addition, open-market operations in government securities had begun, the exchange rate had been floated, and the banking sector deregulated.²

¹The risk and mispricing stories are often associated with 2013 Nobel Laureates Eugene Fama and Robert Shiller respectively. For a nice summary, see the Prize Committee’s citation (Royal Swedish Academy of Sciences, 2013). For an even better (and shorter) summary, see John Cochrane’s blog post: http://johnhcochrane.blogspot.co.nz/2013/10/bob-shillers-nobel.html
²The restrictions on short selling and stock repurchases were not rescinded until the early 1990s. For a
The speed of this liberalisation provides an ideal setting for comparing return predictability in markedly different environments within the same market. The potential for mispricing would seem to be considerably greater in the pre-reform period, and Groenwald (1997) reports evidence (albeit fairly weak) that this is indeed the case. Consequently, if return predictability in the NZ market is primarily due to mispricing, then the estimated magnitude of this predictability should be lower in the post-reform period.

In the next section, we set out our approach in more detail and describe our data. Section 3 contains our results and section 4 offers some concluding remarks.

2 Methods and Data

We follow the approach of Cochrane (2011). The Campbell-Shiller (1988) return approximation yields:

\[ dp_t \approx \sum_{j=1}^{k} \rho^{j-1} r_{t+j} - \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} + \rho^k dp_{t+k} \]

where \( dp \) is the log dividend-price ratio, \( r \) is the log of the real return factor (i.e., the log of \( \frac{1+\text{nominal stock return}}{1+\text{inflation rate}} \)), \( \Delta d \) is log real dividend growth, and \( \rho = \frac{e^{-E[dp]} - 1}{e^{-E[dp]} + 1} \) is a constant of approximation. This in turn implies:

\[ \text{var}(dp_t) \approx \text{cov}(dp_t, \sum_{j=1}^{k} \rho^{j-1} r_{t+j}) + \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} + \rho^k \text{cov}(dp_t, dp_{t+k}) \]

or, dividing through by \( \text{var}(dp_t) \),

\[ 1 \approx b_r^{(k)} + b_{\Delta d}^{(k)} + \rho^k b_{dp}^{(k)} \tag{1} \]

where \( b_r^{(k)} \equiv \frac{\text{cov}(dp_t, \sum_{j=1}^{k} \rho^{j-1} r_{t+j})}{\text{var}(dp_t)} \) is the coefficient from a regression of weighted \( k \)-period future real returns \( \sum_{j=1}^{k} \rho^{j-1} r_{t+j} \) on the current dividend yield \( dp_t \); \( b_{\Delta d}^{(k)} \) and \( b_{dp}^{(k)} \) are defined similarly. Equation (1) tells us that \( b_r^{(k)} \), \( b_{\Delta d}^{(k)} \) and \( \rho^k b_{dp}^{(k)} \) are the proportions of dividend yield volatility attributable to time varying expected returns, time varying expected (negative) dividend growth, and dividend yield persistence respectively.

A considerable literature reports estimates \( \hat{b}_r^{(k)} > 0 \) and \( \hat{b}_{\Delta d}^{(k)} = 0 \), i.e., returns are predictable, but dividend growth is not.\(^3\) However, much disagreement remains over the source of return predictability: is it time variation in risk premia, or irrational mispricing?

To make this distinction more concrete, let \( \bar{r} \) be the component of expected (real) return justified by risk considerations and \( z \) denote the mispricing component. Then \( b_r^{(k)} = b_{\bar{r}}^{(k)} + b_z^{(k)} \), i.e., return predictability is due to risk pricing (\( b_{\bar{r}}^{(k)} \)) or mispricing (\( b_z^{(k)} \)) or both. In practice, of course, the econometrician observes only \( \hat{b}_r^{(k)} \), not the individual components. However, if the relative strength of the two components varies over time, then it potentially

---

\(^3\)For summaries of this evidence, see Cochrane (2008, 2011) and Koijen and Nieuwerburgh (2011)
becomes possible to infer something about the importance of each from \( \hat{b}^{(k)} \) alone. Suppose, for example, that rational pricing involves setting expected returns equal to a constant. Then rational return predictability is, by definition, zero, and any return predictability observed by the econometrician must be the result of mispricing. That is

\[
\hat{b}^{(k)} = \hat{b}^{(k)} > 0; \quad \hat{b}^{(k)} = 0
\]

But for return predictability to be due to mispricing, an obvious prerequisite is that mispricing actually occur. So if the market environment then changes in a way such that mispricing is completely or substantially eliminated, then \( \hat{b}^{(k)} \) must fall, and so therefore does the estimate of \( b^{(k)} \).4

While this example is extreme, it illustrates a general point: if mispricing-induced return predictability is the null, then a reduction in the capacity for mispricing (i.e., fewer limits to arbitrage, as in Shleifer and Vishny, 1997) should see a fall in \( \hat{b}^{(k)} \).5 On the other hand, if the null is that return predictability is primarily due to risk pricing, then the same reduction in the capacity for mispricing should have no effect on \( \hat{b}^{(k)} \) (or possibly a positive effect if market liberalisation allows for more accurate risk pricing).

We investigate this idea using 1931–2012 data from NZ. As explained in the Introduction, NZ experienced a comprehensive, and rapid, market liberalisation in 1984–1985, so we compare estimates of \( \hat{b}^{(k)} \) obtained from the 1931–1984 period with those from the 1985–2012 period. Annual data on stock market returns, bond returns, dividend yields, and dividend growth between 1931 and 2002 come from Lally and Marsden (2004) and were kindly supplied to us by Martin Lally. Subsequent stock market data were obtained from the NZX Company Research database and bond and inflation data from the Reserve Bank of NZ.6

3 Results

To facilitate comparison with existing research, we first use the full 1931–2012 sample to estimate \( b^{(k)}_r \) and \( b^{(k)}_\Delta d \) for various values of \( k \). For \( k > 1 \), we report both direct and VAR-implied estimates, with the latter given by (see Cochrane, 2008, 2011)

\[
\hat{b}^{(k)}_j = \hat{b}^{(1)}_j \left[ \frac{1 - (\hat{\phi} \rho)^k}{1 - \rho \hat{\phi}} \right]
\]

where \( \rho = 0.95 \) and \( \hat{\phi} = 0.81 \) is the log dividend yield autocorrelation estimate. Standard errors are Newey-West and, in the case of the implied estimates, obtained using the delta method.

---

4By (1), \( \hat{b}^{(k)}_r + \rho^k \hat{b}^{(k)}_\Delta d \) must correspondingly increase.
5For a recent summary of the limits-to-arbitrage literature, see Gromb and Vayanos (2010).
Table 1: Forecasting Regressions: Full Sample

Predictability coefficient estimates from the full 1931–2012 data sample. $b_r^{(k)}$ is the coefficient from a regression of weighted $k$-period future log returns $\sum_{j=1}^k \rho^{j-1} r_{t+j}$ on the current dividend yield $d_{pt}$, where $\rho = 0.95$ is a constant of approximation; $b_{er}^{(k)}$ is the analogous coefficient from a regression where the dependent variable is excess log returns. Each implied estimates equals $1 - (\hat{\phi}^k \hat{\phi})$ and its corresponding 1-year estimates (Cochrane, 2008, 2011), where $\hat{\phi} = 0.81$ is the log dividend yield autocorrelation estimate. Terms in brackets are $t$-statistics based on Newey-West standard errors; implied standard error estimates are obtained using the delta method.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$b_r^{(k)}$</th>
<th>$b_{er}^{(k)}$</th>
<th>$b_r^{(k)} \Delta d_{br}$</th>
<th>$b_{er}^{(k)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.27</td>
<td>-0.03</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.2)</td>
<td>(-0.5)</td>
<td>(2.9)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.53</td>
<td>0.08</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.1)</td>
<td>(0.7)</td>
<td>(2.3)</td>
<td></td>
</tr>
<tr>
<td>3 (implied)</td>
<td>0.63</td>
<td>-0.08</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.5)</td>
<td>(-0.5)</td>
<td>(3.4)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.64</td>
<td>0.07</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.3)</td>
<td>(0.4)</td>
<td>(2.1)</td>
<td></td>
</tr>
<tr>
<td>5 (implied)</td>
<td>0.84</td>
<td>-0.11</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.5)</td>
<td>(-0.5)</td>
<td>(3.3)</td>
<td></td>
</tr>
<tr>
<td>$\infty$ (implied)</td>
<td>1.15</td>
<td>-0.15</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.3)</td>
<td>(-0.5)</td>
<td>(3.1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 confirms the principal result of Cochrane (2008, 2011) in NZ data: at short-, medium- and long-term horizons, real stock returns are predictable but real dividend growth is not. All estimates of the return predictability coefficient are economically large and significantly different to zero at the 0.005 level or better, while the dividend predictability estimates are small, often have the wrong sign, and are not remotely close to being statistically significant. The return point estimates are somewhat higher than Cochrane finds for the US: 1-year expected return variation accounts for approximately 27% of dividend yield volatility, rising to 53% over three years, 64% over five years, and to 115% in the long-run; implied estimates are somewhat higher still, and have bigger $t$-statistics.

The last column of Table 1 reports excess return predictability coefficients, where excess return is the log of $\frac{1+\text{nominal stock return}}{1+\text{nominal riskless bond return}}$. Unlike Cochrane (2008) who reports stronger return predictability for excess returns, our estimates are slightly smaller than their real return counterparts and less precisely estimated. This is almost certainly due to the noise in our bond return series. All that we have available are yields on the NZ Government 10-year bond, so we calculate 1-year bond returns as the beginning-of-year yield, an assumption that is obviously sensitive to term structure shifts. Nevertheless, the same picture emerges: (excess) returns are predictable, but (excess) dividend growth is not.

The main question of interest is whether or not the observed return predictability represents market risk pricing or mispricing. Table 2 attempts to shed light on this issue by

---

7 The 10-year bond is the only NZ government bond to exist for the entire 1931–2012 period.
Table 2: Forecasting Regressions: Pre- and Post-Reform Sub-Periods

Return predictability coefficient estimates from the pre-reform 1931–1984 and post-reform 1985–2012 periods. \( b_{rp}^{(k)} \) is the excess return predictability coefficient assuming the riskless bond 1-year return equals its beginning-of-period yield; \( b_{erp}^{(k)} \) is the excess return predictability coefficient assuming the riskless bond is a perpetuity. \(^*\) indicates that the post-reform coefficient is greater than the pre-reform coefficient at the 0.05 level or better, based on a Welch t test.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.23</td>
<td>0.50</td>
<td>0.23</td>
<td>0.50</td>
<td>0.30</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(2.7)</td>
<td>(2.4)</td>
<td>(3.1)</td>
<td>(2.4)</td>
<td>(3.6)</td>
<td>(2.6)</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>1.01*</td>
<td>0.51</td>
<td>1.07*</td>
<td>0.67</td>
<td>1.17*</td>
</tr>
<tr>
<td></td>
<td>(2.3)</td>
<td>(5.6)</td>
<td>(3.1)</td>
<td>(5.9)</td>
<td>(3.2)</td>
<td>(6.4)</td>
</tr>
<tr>
<td>3 (implied)</td>
<td>0.56</td>
<td>1.01</td>
<td>0.56</td>
<td>1.02</td>
<td>0.72</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>(3.0)</td>
<td>(2.6)</td>
<td>(3.5)</td>
<td>(2.5)</td>
<td>(4.1)</td>
<td>(2.8)</td>
</tr>
<tr>
<td>5</td>
<td>0.63</td>
<td>1.08</td>
<td>0.61</td>
<td>1.25*</td>
<td>0.78</td>
<td>1.60*</td>
</tr>
<tr>
<td></td>
<td>(3.0)</td>
<td>(3.2)</td>
<td>(3.2)</td>
<td>(3.8)</td>
<td>(3.0)</td>
<td>(5.1)</td>
</tr>
<tr>
<td>5 (implied)</td>
<td>0.77</td>
<td>1.21</td>
<td>0.77</td>
<td>1.22</td>
<td>0.99</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>(3.2)</td>
<td>(2.4)</td>
<td>(3.7)</td>
<td>(2.4)</td>
<td>(4.3)</td>
<td>(2.6)</td>
</tr>
<tr>
<td>( \infty ) (implied)</td>
<td>1.13</td>
<td>1.33</td>
<td>1.12</td>
<td>1.34</td>
<td>1.44</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(2.9)</td>
<td>(2.2)</td>
<td>(3.4)</td>
<td>(2.2)</td>
<td>(3.4)</td>
<td>(2.3)</td>
</tr>
</tbody>
</table>

comparing the return predictability estimates for the pre-reform 1931–84 period with those for the post-reform 1985–2012 period. If the Table 1 results primarily reflect mispricing, then these estimates should be lower in the latter period.

The first four columns of Table 2 reveal no such trend. In fact, at all horizons, the return predictability estimate is strictly greater in the post-reform period, in some cases at the 0.05 significance level despite the relatively small number of observations in the post-reform period making it difficult to reject the null of no difference. For 1- and 3-year horizons, and for both real and excess returns, the predictability estimates approximately double following liberalisation, while at 5 years the increase is 60–70%; only in the long run do the estimates come close to converging, but even then the difference remains strictly in favour of the later period.

As noted above, our riskless return proxy is likely to contain significant measurement error due to our assumption that the 10-year bond for which we have data is identical to a 1-year bond. To check whether this has any effect on our results, we move to the other end of the scale and assume the 10-year bond is identical to a perpetuity, since this also allows us to calculate an approximate annual return.\(^8\) The results from calculating excess returns in this alternative way appear in the last two columns of Table 2, and are essentially indistinguishable from those in the previous four columns: return predictability estimates are much higher in the post-reform period.

\(^8\)Specifically, in this case the bond return between \( t \) and \( t + 1 \) is \( y_t + \frac{y_t - y_{t+1}}{y_t} \), where \( y_t \) is the bond yield at the beginning of year \( t \).
4 Concluding Remarks

Based on the null of mispricing-induced return predictability, we expect to see weaker predictability in the post-reform period when limits to arbitrage are less apparent. But we do not. Instead, we see considerably stronger return predictability following market liberalisation.

One factor potentially contributing to this phenomenon is lower dividend yield persistence after 1984 (\(\hat{\phi} = 0.66\) versus 0.83 in the pre-reform period), as the present value identity tells us that the sum of return and dividend growth predictability must then rise mechanically as a result. However, our results show that the slack is entirely taken up by returns, and none by dividend growth. To reconcile this with the mispricing explanation of return predictability, one would have to believe either that the NZ liberalisation increased limits to arbitrage, or that lower limits to arbitrage resulted in larger and longer-lasting mispricing. Both are hard stories to tell.

In a similar vein, Maio and Santa-Clara (2013) report that return predictability is much weaker amongst small stocks. As small stocks are more likely to encounter arbitrage difficulties, this result also implies that reducing limits to arbitrage is associated with stronger return predictability. Again, this seems difficult to reconcile with a mispricing story.

Nevertheless, we recommend some caution. Our post-reform sample is relatively short, and contains two major market corrections. As more data become available, thus lessening the impact of these two events, it is possible that return predictability estimates for the post-1984 period could fall substantially.

References


