Testing the predictive power of dividend yields: non-parametric evidence from the G5

Francis Breedon  Marco Bianchi  Darren Sharma

Bank of England, Threadneedle Street, London EC2R 8AH.

The views in this paper are those of the author and not necessarily those of the Bank of England. We thank William Perraudin and Gerald Dwyer for comments.

Issued by the Bank of England, London, EC2R 8AH to which requests for individual copies should be addressed: envelopes should be marked for the attention of the Publications Group. (Telephone: 0171-601-4030).

Bank of England 1997
ISSN 0142-6753
Contents

Abstract 5
1 Introduction 7
2 Background 7
3 The predictive power of dividend yields 8
4 Non-linear estimation 12
5 Testing the significance of yield return relationship 14
6 Out-of-sample performance of trading rules 16
7 Conclusions 19
Appendix 20
References 21
Abstract

This paper extends US evidence on the ability of current dividend yields to predict future equity returns to the G5. Also, using non-parametric methods, we find evidence of a similar non-linear structure in all the countries analysed. This casts doubt on the linear framework adopted in earlier studies. Our tests find that there is a strong relationship between extremes of dividends and future returns (i.e., very low/high dividends do predict low/high returns whilst intermediate levels of dividends do not). This non-linear structure strengthens the statistical evidence of a relationship between dividend yields and future returns and may help explain why previous studies have found mixed evidence.
1 Introduction

The ability of the current level of dividends to predict future equity returns is a deeply researched subject in financial economics. However, as seems often to be the case, this research has ranged quite narrowly by focusing almost exclusively on US data and a linear estimation framework despite no prior evidence that this is either a linear or purely US phenomenon.

By looking at a range of countries in a non-parametric framework we find that there appears to be a very similar non-linear structure in all cases, we also find that doubts that have been cast on the significance of the dividend returns relationship may relate more to the inadequacy of the linear framework that has been used than to the lack of an underlying relationship.

The paper is organised as follows. The next section gives a brief review of previous work on the dividend yield/returns relationship. Section 3 gives results for the standard linear approach to testing this relationship for the G5. Sections 4, 5 and 6 then use non-linear estimation techniques and examines tests of significance within this non-linear framework. Section 7 concludes.

2 Background

The most commonly cited work on the predictive power of current dividend yield for future equity returns is Fama and French (1988) (though study of this phenomenon has long tradition, dating back at least to Dow (1920)). They found that dividend yields were significant predictors of future equity returns and that the explanatory power of dividends increased with the time horizon of returns. However, more recent studies have cast some doubt on these results and have cited two related estimation problems. First, although Fama and French used estimators that are asymptotically robust to the moving-average error problems associated
with long-horizon forecasts, many have questioned the validity of these estimators, even over very large samples. Second, since dividend yields are strongly influenced by current share prices there may be an implicit lagged dependent variable problem in estimation (see for example, Stambaugh (1986)).

In fact, dividend yield regressions have become an important test case for analysing the estimation problems implicit in overlapping forecast horizons. For example, Hodrick (1992) and Nelson and Kim (1993) use a VAR approach with standard errors derived from Monte Carlo simulations and find the predictive power of dividends is largely confirmed. Goetzmann and Jorion (1993 and 1995), on the other hand, find that they cannot reject the hypothesis of no predictive power using a modified bootstrap approach described below.

Alongside the purely empirical question of significance, a number of papers have proposed alternative explanations for this relationship between current dividends and future returns. The most widely accepted approach is that of Rozeff (1984) who suggested that the dividend yield gives an indication of expected risk premia and so how required returns change over time. Many extensions of this basic idea have been proposed (see, for example, Timmermann (1993) who relates this premium to cycles in income). The alternative approach is that of Shiller (1984) who proposes that this relationship reflects the presence of noise trading. If noise traders can temporarily drive prices away from fundamentals, low yields may be associated with temporarily overvalued prices and thus have some predictive power.

3 The predictive power of dividend yields

The standard approach to testing the predictive power of dividend yields is to estimate equation (1)

\[ R_{t,t+i} = \alpha + \beta Y_t + \epsilon_{t,t+i}, \]  

(1)
where $R_{t,t+i}$ is the compound total return on equity from $t$ to $t+i$ (ie $(P_{t+i} - P_t + d_{t+i})/P_t$ where $d_{t+i}$ is compounded dividends received between $t$ and $t+i$ dividend plus capital gain) and $Y_t$ is the compounded dividends paid in the year up to time $t$ divided by the equity price at $t$ (ie $D_{t-1,t}/P_t$). As well as using nominal returns in equation (1) different studies have assessed both real and excess returns and found largely similar results.

As is well established, unless the data frequency is equal to or greater than $t, t+i$ the errors from equation 1 will contain a moving-average process due to overlapping forecast horizons. In this case, Ordinary Least Squares (OLS hereafter) standard errors are invalid and so some form of robust estimator is commonly used instead. Of course these estimates rely on both series being stationary which Augmented Dickey-Fuller tests confirm in all cases except dividends in the United States (However, the low power of these tests is well known; moreover, visual inspection of the time series behaviour would rather point out Japan as more problematic case than the United States).\textsuperscript{1} Table A reports estimates of equation (1) and their robust $t$-statistics for G5 equity indices using both nominal and excess returns.

Table A confirms the results found for the United States for the other G5 countries. Dividend yields do seem to contain information for predicting future nominal and excess equity returns (though the results for Germany in particular are quite weak) and significance is generally increasing in the forecasting horizon as has been found in other studies. However, although the consistency of result across countries is striking the estimation problems described above and the fact that developments across different countries can be strongly related (see for example King and Wadhani (1990)) and so these are not truly independent measures mean that we cannot conclude convincingly from Table A that the relationship is not spurious.

\textsuperscript{1}See Table in the Appendix.
Figure 1: Dividend yields (solid line) and twelve-month returns (dashed line) in G5 countries (for data descriptions see notes to Table A).
Table A: Dividend yield regressions. Estimates of $\beta$ from equation (1) for nominal and excess returns. Legend: N = nominal; E = excess; 1, 12, 24, and 36 = horizon in months.

Notes:
2) USA - NYSE Common Stock Index (up to 3,000 constituents approx), Japan - Tokyo Stock Exchange Section 1 (up to 1,500 constituents approx), Germany Index der Aktienkurse (30 constituents), UK - FT Ordinary share price index (30 constituents). All data end month and all indices value-weighted except UK which is equal-weighted.
3) Excess returns defined relative to short-term official interest rates from the BIS database.
4) For Germany and the United Kingdom prior to 1963 only an end year dividend series available. A monthly series was constructed by assuming dividends grew at a constant rate equal to the growth rate between end-year last year and year before that. Over the period that both monthly data and monthly estimates based on annual data were available for the United Kingdom the correlation between the two series was relatively high at 0.86.
4 Non-linear estimation

Despite having no strong theory to tell us the precise form of the relationship between dividends and future returns, all studies in this area have used a form of the standard linear framework of equation (1). A more appealing approach is to fit the data with some general non-linear model based on the notion of scatterplot smoothing. In this section we apply such an approach, focusing on twelve-month ahead nominal returns as the dependent variable.

According to non-parametric methods, scatterplot smoothing is performed using no parametric assumption about the functional form linking the predictor \( x \) and the response variable \( y \). The philosophy of the approach is to ‘let the data show the effective functional form’. The model is

\[
y_j = g(x_j) + \epsilon_j, \quad j = 1, \ldots, T,
\]

where \( g(x_j) \) is an unknown function to be estimated from the data, with \( E(\epsilon_j|x_j) = 0 \). A smoother \( \hat{g}(x_j) \) is an estimate of the conditional mean of the response, which is obtained by local averaging of the response variable at any given neighbourhood of the predictor. There are a variety of smoothers available (see Hastie and Tibshirani, 1990, page 10 for a complete list) but in this paper, we apply the \textit{LOcally WEighted Scatterplot Smooth}er (Lowess), a popular smoother developed by Cleveland (1979), which is probably suitable for this problem given its robustness to outliers. The Lowess smoother, fitted at a given point, is derived by locally averaging the data in a neighbourhood of that point. A polynomial is fitted to the data using iterative weighted least squares, with the weights computed according to a ‘tri-cube’ weight function. The estimator is constructed through the following steps (see Hastie and Tibshirani, 1990, sec. 2.11; Cleveland 1993, pages 94–101):

(i) Given the value \( x \), the \( k \) nearest neighbours of \( x \) are identified, denoted by \( N(x) \).
(ii) $\Delta(x) = \max_{N(x)} |x - x_j|$ is computed, the distance of the farthest near-neighbour from $x$.

(iii) Weights $w_j$ are assigned to each point in $N(x)$, using the so-called ‘tri-cube’ weight function

$$W \left( \frac{|x - x_j|}{\Delta(x)} \right)$$

where, for any $u$,

$$W(u) = \begin{cases} 
(1 - u^3)^3 & \text{for } 0 \leq u < 1 \\
0 & \text{otherwise}
\end{cases}$$

(iv) The fitted Lowess curve at $x$, $\hat{y}(x)$, is the value of a polynomial of $d$-th degree fitted to the data using iterative weighted least squares, with the weights computed as in (iii). So, if $d = 1$ the values of $a$ and $b$, $\hat{a}$ and $\hat{b}$ respectively, are found that minimise

$$\sum_{j=1}^{n} w_j(x)(y_j - a - bx_j)^2.$$ 

Then, the fit at $x$ is $\hat{y}(x) = \hat{a} + \hat{b}x$. If $d = 2$, the values of $a$, $b$ and $c$ — $\hat{a}$, $\hat{b}$ and $\hat{c}$ respectively — are found that minimise

$$\sum_{j=1}^{n} w_j(x)(y_j - a - bx_j - cx_j^2)^2.$$ 

The fitted Lowess curve at $x$ in this case is $\hat{y}(x) = \hat{a} + \hat{b}x + \hat{c}x^2$.

Clearly, there are two parameters to be selected when using Lowess: the smoothing parameter $k$, that is the number of points in the neighbourhood (‘nearest neighbours’), and the polynomial degree $d$. The choice of $k$ is related to the degree of smoothness desired. Larger or smaller values of $k$ imply different trade-offs between
the bias and the variance of the estimator. The selection of the degree of polynomial, on the other hand, depends on the underlying pattern of the data. Quadratic fitting is used where the scatterplot is characterised by multiple maxima and minima. If the scatterplot has a gentle curvature with few local maxima and minima, then local linear fitting is usually appropriate. Both $k$ and $d$ are usually selected "based on a combination of judgement and of trial and error" (Cleveland, 1993, page 96) bearing in mind certain guidelines.\footnote{We also acknowledge that the use of Lowess implies only minimal assumptions on the errors $\epsilon_i$ in equation (2), which are only required to be zero mean. No normality assumption is needed, nor homoskedasticity of errors, due to the robustness of the fit achieved by Lowess thanks to the use of iterative weighted least squares.}

For our data, we select a smoothing parameter of $\alpha := k/T = 0.65$ for all countries except Japan, that is at any given evaluation point $x$ we construct a fitted value by averaging 65% of the points in the scatterplot which are nearest neighbours to $x$. For Japan, we select a value of $0.45$. As for the degree of polynomial, we select $d = 1$. The choice of the above parameters is supported by the standard diagnostic plots (see Chambers and Hastie, 1992); these are available upon requests from the authors. The estimation results are presented in Figure 2, which shows the existence of a non-linear relationship between yields and future returns.\footnote{Furthermore, both the hypothesis of no relationship and that of a simple linear relationship are rejected at the 95% level on the basis of confidence intervals for the non-linear estimator derived from a standard bootstrap of the residuals (see Efron and Tibshirani, 1993, section 9.5).}

5 Testing the significance of yield return relationship

As was noted above, although standard tests clearly reject the hypothesis of no relationship between current dividends and future returns, these tests may not be appropriate in this case. One
problem that has been highlighted is the implicit lagged dependent variable in this regression. Since dividends themselves are highly autoregressive, most of the variation in yields is coming from price movements. Although there have been a number of suggested solutions to this problem, only one consistently overturns the result of a statistically significant relationship — the bootstrap approach proposed by Goetzmann and Jorion (1993). Since our non-parametric approach appears to give a stronger relationship between yields and future returns than the standard linear approach, it is interesting to see if Goetzmann and Jorion’s tests can still give a rejection in the non-linear case.

Basically, the bootstrap approach of Goetzmann and Jorion is a variation of the standard bootstrap which aims to keep the autoregressive properties of the yield series even after the data has been re-ordered. They start by randomly sampling total one month returns $R^*$ and subtracting the dividend income component. The resultant series can then be used to generate a pseudo-price-level series $P^*$. This is used to create a pseudo-dividend yield $Y^*$, where $Y^* = D/P^*$ and $D$ is the actual (i.e. not re-ordered) dividend series.

By randomising returns, this procedure removes any relationship between returns and dividends, but by using actual dividends, the autoregressive structure of the yield series is maintained. By creating 1,000 pseudo twelve-month returns and pseudo yield series, error bands based on the hypothesis of no relationship can be estimated by fitting Lowess curves to each of these series.

Figure 3 shows these 95% error bands and the actual estimated relationship. It is clear that, in every case, the fitted curves are significantly different from ones that could have been estimated if there had been no relationship.
Figure 2: Lowess smoother for G5 countries plotted with the scatterplot. Note: for the United Kingdom, a few outlier observations are omitted from the scatterplot for which dividend yields were higher than 8.5%.
Figure 3: Goetzman and Jorion 95% confidence intervals (dashed lines) with Lowess curve estimated on the original series of dividend yields and returns in G5 countries.
6 Out-of-sample performance of trading rules

While the non-linearity discovered in the previous analysis may be statistically significant using standard tests, it may not be numerically significant. Thus investors may not care about the difference, since it only occurs during extreme dividend regimes, which themselves occur relatively infrequently. It is useful to know, therefore, how much extra (or less) expected return is predicted by the non-linear model, relative to a (positive) linear (and even flat), model in each regime.

To analysis this question we use the simple trading rule approach proposed by Fuller and Kling (1990). Their simulated trading rule starts with a single unit of funds. At the end of each month the investor estimates his prediction of future equity returns using either the Lowess or OLS model using data available up to that point. He then compares this prediction with the 20-year bond yield to see which is higher. If the bond yield is higher, he puts all of his portfolio in Treasury Bills for a month, otherwise he invests in equity. If this involves a switch of markets since last month, he pays a commission of 0.5%.

Using data for United Kingdom, United States, Germany and Japan, for the period April 1984 to April 1994 we obtain the final value of the investors portfolio $V_t$ as shown in Table B. This shows that the Lowess based rule gives a better return than the OLS rule in three of the four countries analysed largely because it leaves the portfolio in equity more often than the OLS rule (this explains the higher volatility of returns under the Lowess rule). The cases of Germany and Japan, however, involve only one switch in the portfolio from one market to the other whilst in the United States and United Kingdom about 5 switches occur.

4 The rule abstracts from price pressure considerations, that is the fact that actual prices could have been affected by implementing a rule.
Table B: Value of investment at terminal point and standard deviation of $V_t$.

<table>
<thead>
<tr>
<th></th>
<th>Lowess</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$V^*$</td>
<td>SD</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2.38</td>
<td>0.065</td>
</tr>
<tr>
<td>United States</td>
<td>2.94</td>
<td>0.050</td>
</tr>
<tr>
<td>Germany</td>
<td>1.34</td>
<td>0.044</td>
</tr>
<tr>
<td>Japan</td>
<td>1.44</td>
<td>0.121</td>
</tr>
</tbody>
</table>

7 Conclusion

Analysing the relationship between dividend yields and future returns is important for two reasons. First, the relationship has proved to be an important test case for various testing procedures used in financial economics. Second, the relationship – if valid – may help us understand the underlying behaviour of economic agents. This paper has demonstrated not only that this relationship seems to be present in a number of countries, but that in each case the relationship appears to have a similar non-linear structure.

In terms of using dividend yield regressions as a test case, our results suggest that this relationship may be too subtle to be a good basis for testing linear estimation techniques. In fact, our results may explain some of the divergences between techniques found by other authors. As an economic relationship, although we have not presented any economic rationale for our results, we hope that the non-linear structure we have uncovered may become a useful ‘stylised fact’ for future researchers to explain.
Table 1: ADF tests for yields and returns. Critical values for $T(Y - 1)$ if ADF with no trend (Hamilton, 1994, case 2): 5% = -14; 10% = -11.2. For Japan a trend is included (Hamilton, 1994, case 4): 5% = -21; 10% = -18.
REFERENCES


Dow, C (1920), 'Scientific stock speculations', The Magazine of Wall Street.


Publication date in italics

1. Real interest parity, dynamic convergence and the European Monetary System (June 1992)
   Andrew G Haldane
   Mahmood Pradhan

2. Testing real interest parity in the European Monetary System (July 1992)
   Andrew G Haldane
   Mahmood Pradhan

3. Output, productivity and externalities—the case of banking (August 1992)
   R J Colwell
   E P Davis

4. Testing for short-termism in the UK stock market (October 1992)
   David Miles

5. Financial deregulation and household saving
   (October 1992)
   Tamim Bayoumi

6. An investigation of the effect of funding on the slope of the yield curve (January 1993)
   D M Egginton
   S G Hall

7. A simple model of money, credit and aggregate demand (April 1993)
   Spencer Dale
   Andrew G Haldane

8. Bank credit risk (April 1993)
   E P Davis

   Paul Fisher
   Suzanne Hudson
   Mahmood Pradhan

10. The effect of official interest rate changes on market rates since 1987 (April 1993)
    Spencer Dale

11. Tax specific term structures of interest rates in the UK government bond market (April 1993)
    Andrew J Derry
    Mahmood Pradhan

12. Regional trading blocs, mobile capital and exchange rate co-ordination (April 1993)
    Tamim Bayoumi
    Gabriel Sterne

    Gabriel Sterne
    Tamim Bayoumi

14. House prices, arrears and possessions: A three equation model for the UK (June 1993)
    F J Breedon
    M A S Joyce
15 Tradable and non-tradable prices in the UK and EC: measurement and explanation *(June 1993)*  
C L Melliss

16 The statistical distribution of short-term libor rates under two monetary regimes *(September 1993)*  
Bahram Pesaran  
Gary Robinson

17 Interest rate control in a model of monetary policy *(September 1993)*  
Spencer Dale  
Andrew G Haldane

18 Interest rates and the channels of monetary transmission: some sectoral estimates *(September 1993)*  
Spencer Dale  
Andrew G Haldane

19 The effect of futures trading on cash market volatility: evidence from the London stock exchange *(December 1993)*  
Gary Robinson

20 M0: causes and consequences *(December 1993)*  
F J Breeden  
P G Fisher

21 An empirical analysis of M4 in the United Kingdom *(December 1993)*  
P G Fisher  
J L Vega

22 A model of building society interest rate setting *(June 1994)*  
Joanna Paisley

23 Deriving estimates of inflation expectations from the prices of UK government bonds *(July 1994)*  
Mark Deacon  
Andrew Derry

24 Estimating the term structure of interest rates *(July 1994)*  
Mark Deacon  
Andrew Derry

25 Potential credit exposure on interest rate swaps *(August 1994)*  
Ian Bond  
Gareth Murphy  
Gary Robinson

26 New currencies in the Former Soviet Union: a recipe for hyperinflation or the path to price stability? *(September 1994)*  
C L Melliss  
M Cornelius

27 Inflation, inflation risks and asset returns *(November 1994)*  
Jo Corkish  
David Miles
28 The construction of RPIY (February 1995)  R Beaton
P G Fisher

29 Pricing deposit insurance in the United Kingdom (March 1995)  David Maude
William Perraudin


31 Measuring core inflation (April 1995)  Danny T Quah
Shaun P Vahey

32 An assessment of the relative importance of real interest rates, inflation and term premia in determining the prices of real and nominal UK bonds (April 1995)  David G Barr
Bahram Pesaran

33 Granger causality in the presence of structural changes (May 1995)  Marco Bianchi

34 How cyclical is the PSBR? (May 1995)  Joanna Paisley
Chris Salmon

35 Money as an Indicator (May 1995)  Mark S Astley
Andrew G Haldane

36 Testing for convergence: evidence from nonparametric multimodality tests (June 1995)  Marco Bianchi

37 Wage interactions: comparisons or fall-back options (August 1995)  Jennifer C Smith

38 The microstructure of the UK gilt market (September 1995)  James Proudman

39 Valuation of underwriting agreements for UK rights issues: evidence from the traded option market (September 1995)  Francis Breedon
Ian Twinn

40 Rules, discretion and the United Kingdom's new monetary framework (November 1995)  Andrew G Haldane

41 Optimal commitment in an open economy credibility vs flexibility (December 1995)  Sylvester Eijffinger
Eric Schaling
<table>
<thead>
<tr>
<th>No.</th>
<th>Topic</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>Bidding and information: Evidence from gilt-edged auctions</td>
<td>Francis Breedon Joe Ganley</td>
</tr>
<tr>
<td></td>
<td>(January 1996)</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>International bank lending to LDCs—an information-based approach</td>
<td>Prasanna Gai</td>
</tr>
<tr>
<td></td>
<td>(February 1996)</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>A comparison of methods for seasonal adjustment of the monetary</td>
<td>Marco Bianchi</td>
</tr>
<tr>
<td></td>
<td>aggregates (March 1996)</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Base money rules in the United Kingdom (March 1996)</td>
<td>Andrew G Haldane Bennett T McCallum Chris Salmon</td>
</tr>
<tr>
<td></td>
<td>A market for intra-day funds: does it have implications for</td>
<td>Spencer Dale Marco Rossi</td>
</tr>
<tr>
<td></td>
<td>monetary policy? (March 1996)</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Measurement bias in price indices: an application to the UK’s RPI</td>
<td>Alastair W F Cunningham</td>
</tr>
<tr>
<td></td>
<td>(March 1996)</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>The construction of the Bank’s new UK commodity price index</td>
<td>Andrew Logan Lucy O’Carroll</td>
</tr>
<tr>
<td></td>
<td>(March 1996)</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Title</td>
<td>Author(s)</td>
</tr>
<tr>
<td>-----</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>56</td>
<td>Inflation forecast targeting: implementing and monitoring inflation targets</td>
<td>Lars E O Svensson</td>
</tr>
<tr>
<td>57</td>
<td>Why do the LIFFE and DTB bund futures contracts trade at different prices?</td>
<td>Francis Breedon</td>
</tr>
<tr>
<td>58</td>
<td>The determinants of UK business cycles</td>
<td>Allison Holland, Andrew Scott</td>
</tr>
<tr>
<td>59</td>
<td>Which inter-dealer market prevails? An analysis of inter-dealer trading in opaque markets</td>
<td>Victoria Saporta</td>
</tr>
<tr>
<td>60</td>
<td>Testing the predictive power of dividend yields: non-parametric evidence from the G5</td>
<td>Francis Breedon, Marco Bianchi, Darren Sharma</td>
</tr>
</tbody>
</table>