Dividend yields, dividend growth, and return predictability in the cross-section of stocks

Paulo Maio\textsuperscript{1} \hspace{1cm} Pedro Santa-Clara\textsuperscript{2}

First version: June 2012
This version: November 2012\textsuperscript{3}

\textsuperscript{1}Hanken School of Economics. E-mail: paulofmaio@gmail.com.
\textsuperscript{2}Nova School of Business and Economics, NBER, and CEPR. E-mail: psc@novasbe.pt.
\textsuperscript{3}We thank John Cochrane for comments on a preliminary version of this paper, and Kenneth French for providing stock return data on his website.
Abstract

There is a generalized conviction that variation in dividend yields is exclusively related to expected returns and not to expected dividend growth—e.g. Cochrane’s presidential address (Cochrane (2011)). We show that this pattern, although valid for the stock market as a whole, is not true for small and value stocks portfolios where dividend yields are related mainly to future dividend changes. Thus, the variance decomposition associated with aggregate dividend yields (commonly used in the literature) has important heterogeneity in the cross-section of equities. Our results are robust for different forecasting horizons, econometric methodology used (direct long-horizon regressions or first-order VAR), and also confirmed by a Monte-Carlo simulation.

Keywords: asset pricing; predictability of stock returns; dividend-growth predictability; long-horizon regressions; dividend yield; VAR implied predictability; present-value model; size premium; value premium; cross-section of stocks

JEL classification: C22; G12; G14; G17; G35
1 Introduction

There is a generalized conviction that variation in dividend yields is exclusively related to expected returns and not to expected dividend growth—e.g. Cochrane’s presidential address (Cochrane (2011)). We extend the analysis conducted in Cochrane (2008, 2011) to equity portfolios sorted on size and book-to-market (BM). Our goal is to assess whether the results obtained in these studies extends to disaggregated portfolios sorted on size and BM. Indeed this is true for the stock market as a whole. However, we find the opposite pattern for some categories of stocks (e.g., small and value stocks).

Following Cochrane (2008, 2011), we compute the dividend yield variance decomposition based on direct estimates from long-horizon regressions at several forecasting horizons, leading to a term-structure of predictive coefficients at horizons between one and 20 years in the future. Our results show that what explains time-variation in the dividend-to-price ratio of small stocks is predictability of future dividend growth, while in the case of big stocks it is all about return predictability, especially at longer horizons. The bulk of variation in the dividend yield of value stocks is related to dividend growth predictability, while in the case of growth stocks, both return and dividend growth predictability drive the variation in the dividend-to-price ratio. Our conclusions are qualitatively similar if we compute the variance decomposition for the dividend yield based on the implied estimates from a first-order VAR, as is usually done in the literature. Thus, the claim from Cochrane (2008, 2011) that return predictability is the key driver of variation in the dividend yield of the market portfolio does not hold for small and value stocks. We conduct a Monte-carlo simulation to analyze the finite-sample joint distribution of the return and dividend predictive coefficients at multiple horizons, based on the first-order VAR. The results show that we cannot reject dividend growth predictability for small and value stocks.

Our results, although simple, have important implications not only for the return predictability literature but for the asset pricing literature, in general. Specifically, many applications in asset pricing or portfolio choice assume that the dividend-to-price ratio (or similar financial ratios) are a good proxy for expected returns (discount rates).1 Our findings show that while this

---

1For example, in the conditional asset pricing literature, the dividend yield is frequently used as an instrument to proxy for a time-varying price of risk or time-varying betas [e.g., Harvey (1989), Ferson and Harvey (1999),
might represent a good approximation for the value-weighted market index or some categories of stocks, it is certainly not the case for other categories of stocks.

Our work is related with the large literature that use aggregate equity financial ratios like the dividend yield, earnings yield, or book-to-market to forecast stock market returns. Specifically, we are closely related with a smaller and growing literature that analyzes predictability from the dividend-to-price ratio by incorporating the restrictions associated with the Campbell and Shiller (1988a) present-value relation: Cochrane (1992, 2008, 2011), Lettau and Van Nieuwerburgh (2008), Chen (2009), Binsbergen and Koijen (2010), Lacerda and Santa-Clara (2010), Ang (2012), Chen, Da, and Priestley (2012), Engsted, Pedersen, and Tanggaard (2012), among others. Koijen and Van Nieuwerburgh (2011) provides a survey on this area of research. The basic idea of this branch of the return predictability literature is simple: stock return predictability driven by the dividend yield cannot be analyzed in isolation, but must instead be studied jointly with dividend growth predictability since the dividend yield should forecast either or both variables. This literature emphasizes the advantages in terms of statistical power and economic significance of analyzing the return/dividend growth predictability at very long horizons, contrary to the traditional studies of return predictability, which usually use long-horizon regressions up to a limited number of years ahead [see Cochrane (2008) for a discussion]. One reason for the lower statistical power at short and intermediate horizons is that the very large persistence of the annual dividend-to-price ratio overshadows the return/dividend growth predictability at those horizons.

The paper is organized as follows. In Section 2, we describe the data and methodology.

Petkova and Zhang (2005), Maio (2012a), among others]. In the Intertemporal CAPM (ICAPM) literature, the dividend yield is used in some models as a state variable that proxies for future investment opportunities [e.g., Campbell (1996), Petkova (2006), Maio and Santa-Clara (2012), among others]. In the portfolio choice literature, expected stock returns, and thus dynamic portfolio rules, are often linear in the dividend-to-price ratio [see, for example, Campbell and Viceira (1999), Campbell, Chan, and Viceira (2003), and Brandt and Santa-Clara (2006)].


3Other papers analyze the predictability from alternative financial ratios (e.g., earnings yield, book-to-market ratio, payout yield, etc) also in relation with present-value decompositions [e.g., Cohen, Polk, and Vuolteenaho (2003), Larrain and Yogo (2008), Chen, Da, and Priestley (2012), Kelly and Pruitt (2012), Maio (2012b, 2012c)].

4Chen, Da, and Priestley (2012) also look at the return-dividend growth predictability among portfolios, but they use different portfolio sorts and only analyze the very long-run predictability, that is, they do not look at short-run and intermediate term predictability. Moreover, their long-run coefficients are implied from a first-order VAR, while we also compute the long-run coefficients directly from weighted long-horizon regressions.
Section 3 presents the dividend yield variance decomposition for portfolios sorted on size and BM from long-horizon regressions. In Section 4, we conduct an alternative variance decomposition based on a first-order VAR. Section 5 presents the results from Monte-Carlo simulations. Section 6 concludes.

2 Data and methodology

2.1 Methodology

Unlike some of the previous work [e.g., Lettau and Van Nieuwerburgh (2008), Chen (2009), Binsbergen and Koijen (2010), Chen, Da, and Priestley (2012)], in our benchmark analysis the variance decomposition for the dividend yield is based on direct weighted long-horizon regressions, rather than implied estimates from a first-order VAR. The slope estimates from the long-horizon regressions may be different than the implied VAR slopes if the correlation between the log dividend-to-price ratio and future multi-period returns or dividend growth is not fully captured by the first-order VAR. This might happen, for example, if there is a gradual reaction of returns or dividend growth to shocks in the current dividend yield. Thus, the long-horizon regressions provides more correct estimates of the long-horizon predictive relations in the sense that they do not depend on the restrictions imposed by the short-run VAR. On the other hand, the VAR may have better finite-sample properties, that is, there might exist a tradeoff between power and misspecification. In Section 4, we present a variance decomposition based on the first-order VAR, and in Section 5, we analyze the finite-sample distribution of the slopes from the long-horizon regressions.

Following Campbell and Shiller (1988a), the dynamic accounting identity for \( d - p \) can be represented as

\[
d_t - p_t = -\frac{c(1 - \rho^K)}{1 - \rho} + \sum_{j=1}^{K} \rho^{j-1} r_{t+j} - \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} + \rho^K (d_{t+K} - p_{t+K}),
\]

where \( c \) is a log-linearization constant that is irrelevant for the forthcoming analysis; \( \rho \) is a (log-linearization) discount coefficient that depends on the mean of \( d - p \); and \( K \) denotes the
forecasting horizon. Under this present-value relation, the current log dividend-to-price ratio \((d - p)\) is positively correlated with future log returns \((r)\) and the future dividend yield at time \(t + K\), and negatively correlated with future log dividend growth \((\Delta d)\).

Following Cochrane (2008, 2011) we estimate weighted long-horizon regressions of future log returns, log dividend growth, and log dividend-to-price ratio on the current dividend-to-price ratio,

\[
\sum_{j=1}^{K} \rho_{t+j}^{K-1} r_{t+j} = a^r_{t} + b^r_{t} (d_t - p_t) + \epsilon_{t+K}^r, \quad (2)
\]

\[
\sum_{j=1}^{K} \rho_{t+j}^{K-1} \Delta d_{t+j} = a^d_{t} + b^d_{t} (d_t - p_t) + \epsilon_{t+K}^d, \quad (3)
\]

\[
\rho^K (d_{t+K} - p_{t+K}) = a^K_{dp} + b^K_{dp} (d_t - p_t) + \epsilon_{t+K}^{dp}, \quad (4)
\]

where the \(t\)-statistics for the direct predictive slopes are based on Newey and West (1987) standard errors with \(K\) lags.\(^6\)

Similarly to Cochrane (2011), by combining the present-value relation with the predictive regressions above, we obtain an identity involving the predictability coefficients associated with \(d - p\), at horizon \(K\),

\[
1 = b^r_{t} - b^d_{t} + b^{dp}_{t}, \quad (5)
\]

which can be interpreted as a variance decomposition for the log dividend yield. The predictive coefficients \(b^r_{t}\), \(b^d_{t}\), and \(b^{dp}_{t}\) represent the fraction of the variance of current \(d - p\) attributable to return, dividend growth, and dividend yield predictability, respectively.\(^7\)

\(^6\)An alternative estimation of the long-run predictive coefficients relies on a weighted sum of the forecasting slopes for each horizon, \(\sum_{j=1}^{K} \rho_{t+j}^{K-1} b^r_{t}\), where \(b^r_{t}\) is estimated from the following long-horizon regression:

\[
r_{t+j} = a^r_{t} + b^r_{t} (d_t - p_t) + \epsilon_{t+j}^r, j = 1, ..., K.
\]

The difference relative to the first method is that this approach allows for more usable observations in the predictive regression for each forecasting horizon. Unreported results show that the two methods yield qualitatively similar results.

\(^7\)Cohen, Polk, and Vuolteenaho (2003) derive a similar \(K\)-period variance decomposition for the log book-to-market ratio.
2.2 Data and variables

We estimate the predictive regressions using annual data for the 1928–2010 period. The return data on the value-weighted (VW) stock index, with and without dividends, are obtained from CRSP. As in Cochrane (2008), we construct the annual dividend-to-price ratio and dividend growth by combining the series on total return and return without dividends. The estimate for the log-linearization parameter, \( \rho \), for the stock index is 0.965. The descriptive statistics in Table 1 show that the aggregate dividend growth has minor negative autocorrelation, while the log dividend-to-price ratio is highly persistent (0.94).

In the empirical analysis conducted in the following sections we use portfolios sorted on size and book-to-market (BM) available from Kenneth French’s webpage. For each characteristic we use the portfolio containing the bottom 30% of stocks (denoted by L) and the portfolio with the top 30% of stocks (H). The reason for not using a greater number of portfolios within each sorting variable (for example, deciles) is that for some of the more disaggregated portfolios there exist months with no dividends, which invalidates our analysis.

Figure 1 shows the dividend-to-price ratios (in levels) for the size and BM portfolios. We can see that the dividend-to-price ratios were generally higher in the first half of the sample, and have been declining sharply since the 80’s. The dividend yields for big capitalization stocks tend to be higher than for small stocks, although in the first-half of the sample there are some periods where both small and big stocks have similar price multiples. With the exception of the 30’s, value stocks tend to have significant higher dividend yields than growth stocks, although the gap has vanished significantly in recent years. We can also see that the decline in dividend yields since the 80s was more severe for big and value stocks in comparison to small and growth stocks, respectively.

From Table 1 (Panel C), we can see that the log dividend yield of small stocks is more volatile than the corresponding ratio for big stocks (standard deviation of 0.72 versus 0.44), while big stocks have a significantly more persistent dividend-to-price ratio (0.95 versus 0.83). On the other hand, the log dividend yield of value stocks is slightly more volatile than for growth stocks (standard deviation of 0.60 versus 0.54), while growth stocks have a more persistent multiple (0.95 versus 0.86). The estimates for \( \rho \) in the case of the “small” and “big” portfolios are 0.979
and 0.965, respectively, while the corresponding estimates for the growth and value portfolios are 0.972 and 0.963, respectively.

In Figure 2 we have the time-series for portfolio (gross) dividend growth rates. We can see that dividend growth was quite volatile during the great depression, especially for small and value stocks. The standard deviation calculations in Table 1 (Panel B) show that small and value stocks exhibit much more volatile dividend growth than big and growth stocks, respectively. Dividend growth is negatively autocorrelated for small and growth stocks, while for value stocks we have a small positive autocorrelation.

3 Predictability of size and book-to-market portfolio

3.1 Size portfolios

The term-structure of predictive coefficients, and respective \(t\)-statistics, for the variance decompositions associated with small and large stocks are shown in Figure 3. In the case of small stocks (Panel A) the share associated with dividend growth predictability approaches 70\% at the 20-year horizon, while the fraction of return predictability never exceeds 30\% (which is achieved for horizons between 6 and 8 years in the future).\(^8\) For big stocks (Panel C), the share of return predictability is clearly dominant and goes above 100\% for horizons beyond 15 years. The reason for this “overshooting” is that the long-horizon predictability of the dividend yield has the “wrong” sign (about -20\% at \(K = 20\)).

The analysis of the \(t\)-statistics shows that the slopes in the dividend growth regressions for small stocks are statistically significant at the 5\% level for horizons beyond 10 years, but these coefficients are insignificant (at all horizons) in the case of the big portfolio. In contrast, the coefficients in the return equation are statistically significant at all horizons for the big portfolio, while in the case of the small portfolio there is also statistical significance for horizons beyond three years (although the magnitudes of the \(t\)-ratios are smaller). The coefficients associated with the future dividend yield are statistically significant at shorter horizons (less than 10 years).

---

\(^8\)Rangvid, Schmeling, and Schrimpf (2011) also perform long-horizon regressions of dividend growth on the dividend yield for portfolios sorted on size (deciles). However, they do not compute the variance decomposition of the dividend yield for each size portfolio at multiple horizons, that is, they do not quantify the fraction of dividend yield variation attributable to return, dividend growth, or future dividend yield predictability.
for both portfolios.

As a reference point, we present the variance decomposition for the stock index in Panels E and F. We can see that the plots of the term-structure of predictive slopes for the market and the big portfolio look quite similar, showing that the bulk of variation in the aggregate dividend-to-price ratio is return predictability, which confirms previous evidence based on implied VAR estimates [see Cochrane (2008, 2011) and Chen (2009)]. This result is also consistent with the fact that the value-weighted index is tilted toward big capitalization stocks. For all portfolios the accuracy of the identity involving the predictive coefficients in equation (5) is quite good as shown by the curve labeled “sum”, which exhibits values very close to 100% at all forecasting horizons.

In sum, these results indicate that the predictability decomposition for the market index hide some significant and interesting differences among stocks with different market capitalization: what explains time-variation in the dividend-to-price ratio of small stocks is predictability of future dividend growth, while in the case of big stocks it is all about return predictability, especially at long-horizons.

### 3.2 Book-to-market portfolios

Next, we conduct a similar analysis for book-to-market portfolios. The analysis of the term-structure of predictive coefficients in Figure 4 shows some significant differences between growth and value portfolios. In the case of growth stocks (Panels A and B) the shares of return and dividend growth predictability are very similar at long horizons (around 40% each at $K = 20$). Simultaneously, there is some predictability about the future dividend yield for very long horizons (around 30%), thus indicating that the dividend-to-price ratio of growth stocks is quite persistent. The term-structure of $t$-ratios shows that for growth stocks the dividend growth slopes are significant for horizons beyond four years, while the return coefficients are not significant (at the 5% level) both at very short and long horizons.

In the case of value stocks (Panels C and D) the pattern of direct long-horizon coefficients shows that dividend growth predictability achieves weights close to 100% at the 20-year horizon. In comparison, the share of long-run return predictability never exceeds 40% (which is obtained for $K = 8$) and decreases to values around 20% at the 20-year horizon. The slopes associated
with future $d - p$ converge to zero much faster than in the case of the growth portfolio, turning negative for horizons greater than nine years. The $t$-statistics indicate that the dividend growth slopes are highly significant at longer horizons, while there are some periods (shorter horizons) for which the return coefficients are not statistically significant at the 5% level.

The variance decomposition associated with the stock index (Panels E and F) indicates more return predictability and less dividend growth predictability than the growth portfolio. This suggests that the predictability pattern estimated for the value-weighted index might be the result of some large growth stocks.

Overall, the results for the BM portfolios can be summarized as follows. First, the bulk of variation in the dividend yield of value stocks is related to dividend growth predictability. Second, in the case of growth stocks, both return and dividend growth predictability drive equally the variation in the current dividend yield.

Why is future dividend growth the main influence on the current dividend yield of value stocks, while for growth stocks future return predictability also plays an important role? A possible explanation is that growth stocks are firms with a long duration of cash flows, which are not expected to have positive cash flows for several periods in the future [see Cornell (1999) and Lettau and Wachter (2007) for a discussion]. Thus, changes in their current valuation are more likely to reflect changes in discount rates (or in future dividend yields) rather than news about future dividends in the upcoming periods, given the virtually zero expected dividend growth rates for these stocks in the short-term. Put differently, the dividend yield of growth stocks is more sensitive to variations in discount rates than to changes in dividend growth.9

9In related work, Campbell and Vuolteenaho (2004) argue that the returns of value stocks are more sensitive to shocks in aggregate cash flows than the returns of growth stocks, that is, value stocks have greater cash flow betas. On the other hand, they find that growth stocks have higher discount rate betas. Campbell, Polk, and Vuolteenaho (2010) decompose further the cash flow and discount rate betas into the parts attributable to specific cash flows and discount rates of growth/value stocks. They find that the cash flows of value stocks drive their cash flow betas, and similarly, the cash flows of growth stocks determine their discount rate betas.
4 VAR-based results

4.1 Methodology

In this section, we conduct an alternative variance decomposition for portfolio dividend yields based on a first-order VAR, as in Cochrane (2008, 2011), Lettau and Van Nieuwerburgh (2008), Chen, Da, and Priestley (2012), among others.

Following Cochrane (2008), we base the long-run predictability statistics on the following first-order restricted VAR,

\[ r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1}, \]  
\[ \Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{d,t+1}, \]  
\[ d_{t+1} - p_{t+1} = a_{dp} + \phi(d_t - p_t) + \varepsilon_{dp,t+1}, \]

where the \( \varepsilon \)s represent error terms. The VAR above is estimated by OLS (equation-by-equation) with Newey and West (1987) \( t \)-statistics (computed with one lag).

By combining the VAR above with the Campbell and Shiller (1988a) present-value relation, we obtain an identity involving the predictability coefficients associated with \( d - p \), at horizon \( K \),

\[ 1 = b^K_r - b^K_d + b^K_{dp}, \]  
\[ b^K_r \equiv b_r(1 - \rho^K \phi^K), \]  
\[ b^K_d \equiv b_d(1 - \rho^K \phi^K), \]  
\[ b^K_{dp} \equiv \rho^K \phi^K, \]

which represents the variance decomposition shown in Cochrane (2008, 2011) or Engsted and Pedersen (2010). The \( t \)-statistics associated with the predictive coefficients in (9) are computed from the \( t \)-statistics for the VAR slopes by using the Delta method (details are available in Appendix A). This decomposition differs from the variance decomposition used in the previous section to the extent that the long-horizon coefficients are not estimated directly from the long-horizon regressions but rather implied by the VAR estimates. If the first-order VAR does not
fully capture the dynamics of the data generating process for $r$, $d-p$, and $\Delta d$, then this variance decomposition will be a poor approximation of the true decomposition for the dividend yield.

Similarly to Cochrane (2008, 2011), we also compute the variance decomposition for an infinite horizon ($K \to \infty$):

$$1 = b^{lr}_r - b^{lr}_d,$$

$$b^{lr}_r \equiv \frac{b_r}{1 - \rho \phi},$$

$$b^{lr}_d \equiv \frac{b_d}{1 - \rho \phi}.$$

In this decomposition, all the variation in the dividend yield is tied with either return or dividend growth predictability, since the predictability of the future dividend yield vanishes out for a very long horizon.

The $t$-statistics for the long-run coefficients, $b^{lr}_r, b^{lr}_d$, are based on the standard errors of the one-period VAR slopes by using the Delta method. We compute $t$-statistics for two null hypotheses: the first null assumes that there is only dividend growth predictability,

$$H_0: b^{lr}_r = 0, b^{lr}_d = -1,$$

while the second null hypothesis assumes that there is only return predictability:

$$H_0: b^{lr}_r = 1, b^{lr}_d = 0.$$

### 4.2 Size portfolios

The VAR estimation results for the size portfolios are presented in Table 2, Panel A. The return slope for the small portfolio is 0.05 and this estimate is not significant at the 10% level. The $R^2$ estimate for the return equation is only 1.34%. The dividend growth coefficient has a relatively large magnitude (-0.14), although this estimate is also not significant at the 10% level, which should be related with the high volatility of the dividend growth of small stocks, as referred in Section 2. The forecasting ratio in the dividend growth equation is 6.18%, about four times as large as the fit in the return equation. The dividend yield of the small portfolio is much less
persistent than the market index with an autoregressive coefficient of 0.83 versus 0.95.

The results for the big portfolio are qualitatively different than the findings for small stocks. The return coefficient is 0.09, almost twice as large as the estimate for the small portfolio. This point estimate is significant at the 10% level, while the coefficient of determination is 4.17%. The dividend growth slope has the wrong sign, but it is highly insignificant. The estimate for $\phi$ is 0.95, showing that the dividend yield of big stocks is significantly more persistent than for small stocks. The VAR estimation results for the market index, displayed in Table 2, Panel C, show that the predictive slopes and forecasting ratios are relatively similar to the corresponding estimates associated with big caps.

From Table 2 we can also see that in all cases the slope estimates (and associated $t$-statistics calculated under the Delta method) implied from the one-period variance decomposition,

$$1 = b_r - b_d + \rho \phi,$$

are very similar to the direct VAR estimates, thus showing that this present-value decomposition works relatively well at the one-year horizon.

Overall, these results show that big capitalization stocks have significantly greater short-run return predictability from the dividend yield than small stocks. Second, small caps have some relevant short-run dividend growth predictability from $d - p$.

The term-structure of the VAR-based variance decomposition for the size portfolios is displayed in Figure 5. In the case of small stocks (Panel A), even at long horizons the bulk of $d - p$ variation is associated with dividend growth predictability (almost 80%) rather than return predictability, which does not go above 30%. In the case of large stocks (Panel C), we have an opposite pattern with return predictability at long-horizons representing about 90% of the variance of the current dividend-to-price ratio, while dividend growth predictability has the wrong sign (positive slopes). The VAR-based variance decomposition for the stock index (Panel E) is quite similar to the corresponding decomposition for large caps, as in the last section.

The analysis of the $t$-statistics show that the VAR-based return slopes are never statistically significant for the “small” portfolio, but they are significant at the 5% level for horizons beyond two years in the case of the “big” portfolio. In contrast, the VAR-based dividend coefficients
are statistically significant for horizons greater than six years in the case of small stocks, but largely insignificant in the case of big stocks.

To provide a better picture of the predictability mix at very long horizons, in the case of small stocks the return and dividend growth long-run (infinite horizon) coefficients are 0.27 and -0.74, respectively, as shown in Table 2. Moreover, we cannot reject the null of no return predictability ($t$-stat=1.08), whereas we strongly reject the null of no dividend predictability ($t$-stat=-2.97). In the case of big stocks we have a totally different picture: the estimates for $b_{lr}^r$ and $b_{lr}^d$ are 1.08 and 0.12, respectively, that is, more than 100% of the variation in the dividend yield is associated with return predictability in the long-run since the slope for dividend growth has the wrong sign. We reject (at the 5% level) the null of no return predictability while we cannot reject the null of no dividend growth predictability ($t$-ratios close to zero).

By comparing with the benchmark variance decomposition estimated in the last section, the results are relatively similar in the sense that in both methodologies it is the case that return predictability is the main driver of variation in the dividend yield of big caps, while for small caps the bulk of the variation is related to dividend growth predictability.

4.3 Book-to-market portfolios

The VAR estimation results for the value and growth portfolios are depicted in Table 2, Panel B. Growth stocks have a return predictive slope of 0.05 and a forecasting ratio of 1.75%. For value stocks the amount of return predictability is smaller with a coefficient of 0.04 and a $R^2$ estimate of only 0.83%. Regarding dividend growth predictability, the slope has the right sign in the case of growth stocks (-0.03), but this estimate is highly insignificant and the explanatory ratio is quite small (0.71%). In contrast, for value stocks we have a coefficient estimate of -0.13, which is both economically and statistically significant (10% level). The associated forecasting ratio is 4.67%, much higher than the fit for the return equation of the value portfolio. Moreover, the dividend yield of growth stocks is significantly more persistent than the corresponding ratio for value stocks, with autoregressive slopes of 0.95 and 0.86, respectively.

The VAR-based variance decomposition for the BM portfolios at several horizons is presented

\[10\] Similarly to Cochrane (2008), the $t$-statistics for the return and dividend growth long-run coefficients are similar, although they are not numerically equal.
in Figure 6. In the case of growth stocks, return predictability is the dominant source of \(d - p\) variance, although the weights are significantly lower than the corresponding estimates for large stocks, for example. The reason is that there is some dividend growth predictability at long-horizons (about 20%). The term-structure of \(t\)-ratios shows that despite the fact that the VAR-based return predictability represents the major source of variation in the dividend yield of growth stocks, the respective coefficients are not statistically significant at the 5% level at any horizon. By comparing to the variance decomposition associated with the stock index (Panel E), the growth portfolio has less return predictability and more dividend growth predictability.

For value stocks most of the variation in the current dividend yield is a result of long-run dividend growth predictability (around 70%), while the share attached to return predictability never goes above 30% even at long horizons. Interestingly, the plot for value stocks looks quite similar to the one for small stocks in Figure 5. The \(t\)-statistics indicate that the dividend growth predictive coefficients for value stocks are statistically significant (at the 5% level) at nearly all horizons, while the return slopes are not significant at any horizon.

The long-run return and dividend slopes for growth stocks (Table 2, Panel B) are 0.66 and -0.35, respectively, confirming that long-run return predictability is the main driver of variation in the dividend yield of those stocks. However, due to large standard errors of the VAR slopes we cannot reject the null of no return predictability at the 5% level (\(t\)-stat=1.50). In contrast, for value stocks the estimates for \(b^r_l\) and \(b^d_l\) are 0.22 and -0.75, respectively, indicating that the key driver of the dividend yield is long-run dividend predictability. We do not reject the null of no return predictability by a big margin, while the null of no dividend predictability is rejected at the 5% level.

By comparing the VAR-based variance decomposition with the benchmark decomposition analyzed in the last section, we detect a similar pattern of predictability: for growth stocks the main driver of variation in the dividend yield is return predictability (although there is also some predictability of dividend growth and future dividend yield), while for value stocks return predictability is relatively marginal.
5 Monte-Carlo simulation

In this section, we conduct a Monte-carlo simulation to analyze the finite-sample joint distribution of the return and dividend predictive coefficients at multiple horizons, based on the first-order VAR estimated in the last section.

Following Cochrane (2008), the first Monte-Carlo simulation is based on the null hypothesis of no return predictability, that is, the data generating process is simulated under the hypothesis that what drives the variation in the dividend yield is only dividend growth predictability:

\[
\begin{pmatrix}
    r_{t+1} \\
    \Delta d_{t+1} \\
    d_{t+1} - p_{t+1}
\end{pmatrix} =
\begin{pmatrix}
    0 \\
    \rho \phi - 1 \\
    \phi
\end{pmatrix}
(\Delta d_{t} - p_{t}) +
\begin{pmatrix}
    \varepsilon_{t+1}^d \\
    \varepsilon_{t+1}^d - \rho \varepsilon_{t+1}^{dp} \\
    \varepsilon_{t+1}^{dp}
\end{pmatrix}.
\]

In the second Monte-Carlo experiment, we simulate the first-order VAR by imposing the restrictions (in the predictive slopes and residuals) consistent with the null of no dividend growth predictability, that is, what drives the dividend yield is only return predictability:

\[
\begin{pmatrix}
    r_{t+1} \\
    \Delta d_{t+1} \\
    d_{t+1} - p_{t+1}
\end{pmatrix} =
\begin{pmatrix}
    1 - \rho \phi \\
    0 \\
    \phi
\end{pmatrix}
(\Delta d_{t} - p_{t}) +
\begin{pmatrix}
    \varepsilon_{t+1}^r \\
    \varepsilon_{t+1}^r + \rho \varepsilon_{t+1}^{dp} \\
    \varepsilon_{t+1}^{dp}
\end{pmatrix}.
\]

Notice that in both simulations we are imposing the one-period variance decomposition for the dividend yield, \(1 = b_r - b_d + \rho \phi\).

Following Cochrane (2008), in drawing the VAR residuals (10,000 times) we assume that they are jointly normally distributed and use their covariances from the original sample. The dividend yield for the base period is simulated as \(d_0 - p_0 \sim N\left[0, \text{Var}(\varepsilon_{t+1}^{dp})/(1 - \phi^2)\right]\). Equipped with the artificial data we compute the fractions of simulated estimates for the return/dividend growth coefficients that are higher or lower than the estimates found in the data. The probabilities for the one-year and 20-year coefficients, under both null hypotheses discussed above, are presented in Table 3.

Figure 7 presents the term-structure of the probabilities that either the return or dividend growth coefficients are higher than the corresponding sample estimates, under the null of no
return predictability. We also display the probabilities that both slopes are jointly higher than the corresponding sample estimates at each horizon (curve labeled $r + d$).

In the case of the small portfolio, the p-values for the return slopes are well above 10% at all horizons, thus indicating that, according to the marginal distribution, these coefficients are not statistically significant. However, we reject (at the 10% level) the joint null of dividend predictability/no return predictability for short-horizons as indicating by the probabilities that both coefficients are joint higher than the respective sample estimates (probability of 5% at the one-year horizon). Nevertheless, for longer horizons ($K > 10$), we can no longer reject the joint null hypothesis (probability of 15% at the 20-year horizon).

In the case of the big portfolio, the probabilities associated with the joint null are clearly below 5% at any horizon, thus indicating, that we reject the null of dividend predictability/no return predictability (these probabilities are 2% and 1% at the one- and 20-year horizons, respectively). These results are consistent with the findings in Cochrane (2008) that the joint null of dividend predictability/no return predictability is rejected (at the one-year horizon) for the value-weighted market portfolio (which is tilted towards bigger stocks).

Regarding the BM portfolios the marginal p-values associated with the return slopes clearly indicate no return predictability for both portfolios. However, we do not reject the joint null hypothesis (at the 10% level) at short horizons and for both growth and value stocks (probabilities of 9% and 7% for growth and value stocks, respectively, at the one-year horizon).

Figure 8 presents the term-structure of the probabilities that either the return or dividend coefficients are lower than the sample estimates, under the null of no dividend predictability. Similarly to the first Monte-Carlo simulation, we also present the probabilities that both slopes are jointly lower than the corresponding sample estimates.

We can see quite opposite results for small and big stocks: in the case of small stocks the marginal p-values associated with the dividend slopes are clearly below 10%, and actually lower than 5% at most horizons. In contrast, for big stocks the p-values for the dividend slopes are quite large, thus showing that we cannot reject the null of no dividend predictability. The probabilities for the joint null of no dividend predictability/return predictability point to the same qualitative decisions as the marginal p-values for the dividend slopes, that is, the joint null is clearly rejected for small stocks, but not rejected for big stocks. From Table 3, the
corresponding probabilities for small stocks are 0% at both $K = 1$ and $K = 20$, while for big stocks these probabilities are 29% and 50%, respectively.

The analysis for the BM portfolios also shows quite different results for growth and value stocks: while for growth stocks the p-values for the dividend growth slopes are well above 10%, thus showing no dividend predictability, in the case of value stocks the probabilities associated with the dividend coefficients are quite close to zero, especially at long horizons.

Overall, the results of this section largely confirm the asymptotic statistical inference in the last section: we cannot reject return predictability for big and growth stocks, and on the other hand, we cannot reject dividend growth predictability for small and value stocks.

6 Conclusion

We provide additional evidence for the predictability associated with the dividend yield for future stock returns and dividend growth. We extend the analysis conducted in Cochrane (2008, 2011) to equity portfolios sorted on size and book-to-market (BM). Our results show that what explains time-variation in the dividend-to-price ratio of small stocks is predictability of future dividend growth, while in the case of big stocks it is all about return predictability, especially at longer horizons. The bulk of variation in the dividend yield of value stocks is related to dividend growth predictability, while in the case of growth stocks, both return and dividend growth predictability drive the variation in the dividend-to-price ratio.

Our conclusions are qualitatively similar if we compute the variance decomposition for the dividend yield based on the implied estimates from a first-order VAR, as is usually done in the literature. Thus, the claim from Cochrane that return predictability is the key driver of variation in the dividend yield of the market portfolio does not hold for small and value stocks. We conduct a Monte-carlo simulation to analyze the finite-sample joint distribution of the return and dividend predictive coefficients at multiple horizons, based on the first-order VAR. The results show that we cannot reject dividend growth predictability for small and value stocks.
References


Campbell, J., and L. Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, Quarterly Journal of Economics 114, 434–495.


A  Delta method for standard errors of long-run coefficients

To compute the asymptotic standard errors of the long-run predictive coefficients, \( b^K \equiv (b^K_r, b^K_d, b^K_{dp})' \), we use the delta method:

\[
\text{Var}(b^K) = \frac{\partial b^K}{\partial b'} \text{Var}(b) \frac{\partial b^K}{\partial b},
\]

(A.1)

where \( b \equiv (b_r, b_d, \phi)' \) denotes the vector of VAR slopes. The matrix of derivatives is given by:

\[
\frac{\partial b^K}{\partial b'} = \begin{bmatrix}
\frac{\partial b^K_{r}}{\partial \phi} & \frac{\partial b^K_{d}}{\partial \phi} & \frac{\partial b^K_{dp}}{\partial \phi} \\
\frac{\partial b^K_{r}}{\partial b_r} & \frac{\partial b^K_{d}}{\partial b_r} & \frac{\partial b^K_{dp}}{\partial b_r} \\
\frac{\partial b^K_{r}}{\partial b_d} & \frac{\partial b^K_{d}}{\partial b_d} & \frac{\partial b^K_{dp}}{\partial b_d} \\
\frac{\partial b^K_{r}}{\partial \phi} & \frac{\partial b^K_{d}}{\partial \phi} & \frac{\partial b^K_{dp}}{\partial \phi}
\end{bmatrix} = \begin{bmatrix}
\frac{(1-\rho^K\phi)}{1-\rho^K} & 0 & -Kb_r\rho^K\phi^{K-1}(1-\rho^K) + \rho_b(1-\rho^K) \\
0 & \frac{(1-\rho^K\phi)}{1-\rho^K} & -Kb_d\rho^K\phi^{K-1}(1-\rho^K) + \rho_b(1-\rho^K) \\
0 & 0 & \frac{(1-\rho^K\phi)}{K\rho^K(\phi)}
\end{bmatrix}.
\]

(A.2)

B  Derivation of the variance decomposition for \( d-p \) with excess returns

Following Cochrane (2008), by multiplying both sides of the augmented Campbell and Shiller (1988a) decomposition based on excess returns by \( d_t - p_t - E(d_t - p_t) \), and taking unconditional expectations we obtain the following variance decomposition for \( d_t - p_t \):

\[
\text{Var}(d_t - p_t) = \text{Cov} \left( \sum_{j=1}^{K} \rho^{j-1}r_{t+j}^e, d_t - p_t \right) - \text{Cov} \left( \sum_{j=1}^{K} \rho^{j-1}\Delta d_{t+j}, d_t - p_t \right) + \text{Cov} \left( \sum_{j=1}^{K} \rho^{j-1}r_{f,t+j}, d_t - p_t \right) + \text{Cov} \left[ \rho^K(d_{t+K} - p_{t+k}), d_t - p_t \right],
\]

(B.3)

and by dividing both sides by \( \text{Var}(d_t - p_t) \), we have:

\[
1 = \beta \left( \sum_{j=1}^{K} \rho^{j-1}r_{t+j}^e, d_t - p_t \right) - \beta \left( \sum_{j=1}^{K} \rho^{j-1}\Delta d_{t+j}, d_t - p_t \right) + \beta \left( \sum_{j=1}^{K} \rho^{j-1}r_{f,t+j}, d_t - p_t \right) + \beta \left[ \rho^K(d_{t+K} - p_{t+k}), d_t - p_t \right] \equiv
\]

\[
1 = b^K_r - b^K_d + b^K_f + b^K_{dp},
\]

(B.4)

where \( \beta(y, x) \) denotes the slope from a regression of \( y \) on \( x \). This expression represents the variance decomposition for \( d-p \) when the predictive slopes are obtained directly from long-horizon regressions.
**Table 1: Descriptive statistics**

This table reports descriptive statistics for the log market return ($r$), log dividend growth ($\Delta d$), and log dividend-to-price ratio ($d - p$). The equity portfolios consist of the value-weighted index ($VW$); small stocks ($SL$); big stocks ($SH$); growth stocks ($BML$); and value stocks ($BMH$). The sample corresponds to annual data for the 1928–2010 period. $\phi$ designates the first-order autocorrelation.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Stdev.</th>
<th>Min.</th>
<th>Max.</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A ($r$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$VW$</td>
<td>0.09</td>
<td>0.20</td>
<td>-0.59</td>
<td>0.45</td>
<td>0.05</td>
</tr>
<tr>
<td>$SL$</td>
<td>0.11</td>
<td>0.31</td>
<td>-0.77</td>
<td>0.93</td>
<td>0.12</td>
</tr>
<tr>
<td>$SH$</td>
<td>0.09</td>
<td>0.19</td>
<td>-0.57</td>
<td>0.43</td>
<td>0.07</td>
</tr>
<tr>
<td>$BML$</td>
<td>0.08</td>
<td>0.20</td>
<td>-0.45</td>
<td>0.39</td>
<td>0.02</td>
</tr>
<tr>
<td>$BMH$</td>
<td>0.12</td>
<td>0.26</td>
<td>-0.82</td>
<td>0.79</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Panel B ($\Delta d$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$VW$</td>
<td>0.04</td>
<td>0.14</td>
<td>-0.38</td>
<td>0.37</td>
<td>-0.07</td>
</tr>
<tr>
<td>$SL$</td>
<td>0.08</td>
<td>0.39</td>
<td>-1.89</td>
<td>1.47</td>
<td>-0.27</td>
</tr>
<tr>
<td>$SH$</td>
<td>0.04</td>
<td>0.14</td>
<td>-0.33</td>
<td>0.32</td>
<td>-0.06</td>
</tr>
<tr>
<td>$BML$</td>
<td>0.04</td>
<td>0.16</td>
<td>-0.34</td>
<td>0.43</td>
<td>-0.09</td>
</tr>
<tr>
<td>$BMH$</td>
<td>0.06</td>
<td>0.36</td>
<td>-2.08</td>
<td>1.16</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Panel C ($d - p$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$VW$</td>
<td>-3.33</td>
<td>0.43</td>
<td>-4.50</td>
<td>-2.63</td>
<td>0.94</td>
</tr>
<tr>
<td>$SL$</td>
<td>-3.85</td>
<td>0.72</td>
<td>-5.98</td>
<td>-2.70</td>
<td>0.83</td>
</tr>
<tr>
<td>$SH$</td>
<td>-3.31</td>
<td>0.44</td>
<td>-4.56</td>
<td>-2.60</td>
<td>0.95</td>
</tr>
<tr>
<td>$BML$</td>
<td>-3.55</td>
<td>0.54</td>
<td>-4.85</td>
<td>-2.51</td>
<td>0.95</td>
</tr>
<tr>
<td>$BMH$</td>
<td>-3.25</td>
<td>0.60</td>
<td>-5.32</td>
<td>-2.43</td>
<td>0.86</td>
</tr>
</tbody>
</table>
This table reports the one-month restricted VAR estimation results for portfolios sorted on size (Panel A) and book-to-market (Panel B). The equity portfolios represent small/big stocks and growth/value stocks. Panel C shows the results for the value-weighted stock index (VW). The variables in the VAR are the log stock return (r), log dividend growth (Δd), and log dividend-to-price ratio (d − p). b(φ) denote the VAR slopes associated with lagged d − p, while t denote the respective Newey and West (1987) t-statistics (calculated with one lag). b(φi) denote the slope estimates implied from the variance decomposition for d − p, and t denote the respective asymptotic t-statistics computed under the Delta method. $R^2(\%)$ is the coefficient of determination for each equation in the VAR, in %. $\delta^R$ denote the long-run coefficients (infinite horizon). $t(b_{lr}^R = 0)$ and $t(b_{lr}^R = 1)$ denote the t-statistics associated with the null hypotheses $H_0 : b_{lr}^R = 0, b_{lr}^R = -1$ and $H_0 : b_{lr}^R = 1,b_{lr}^R = 0$, respectively. The sample corresponds to annual data for 1928–2010. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>b(φ)</th>
<th>t</th>
<th>b(φ)</th>
<th>t</th>
<th>$R^2(%)$</th>
<th>$\delta^R$</th>
<th>t(b_{lr}^R = 0)</th>
<th>t(b_{lr}^R = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A (Size)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.050</td>
<td>1.22</td>
<td>0.048</td>
<td>1.17</td>
<td>1.34</td>
<td>0.273</td>
<td>1.08</td>
<td>-2.88</td>
</tr>
<tr>
<td>Δd</td>
<td>-0.135</td>
<td>-1.49</td>
<td>-0.133</td>
<td>-1.46</td>
<td>6.18</td>
<td>-0.740</td>
<td>1.05</td>
<td>-2.97</td>
</tr>
<tr>
<td>d − p</td>
<td>0.834</td>
<td><strong>9.99</strong></td>
<td>0.832</td>
<td><strong>9.97</strong></td>
<td>68.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.089</td>
<td>1.87</td>
<td>0.093</td>
<td><strong>1.94</strong></td>
<td>4.17</td>
<td>1.079</td>
<td><strong>2.43</strong></td>
<td>0.18</td>
</tr>
<tr>
<td>Δd</td>
<td>0.010</td>
<td>0.28</td>
<td>0.007</td>
<td>0.18</td>
<td>0.10</td>
<td>0.121</td>
<td><strong>2.43</strong></td>
<td>0.26</td>
</tr>
<tr>
<td>d − p</td>
<td>0.951</td>
<td><strong>20.36</strong></td>
<td>0.954</td>
<td><strong>20.25</strong></td>
<td>88.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B (BM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.048</td>
<td>1.10</td>
<td>0.048</td>
<td>1.07</td>
<td>1.75</td>
<td>0.661</td>
<td>1.50</td>
<td>-0.77</td>
</tr>
<tr>
<td>Δd</td>
<td>-0.026</td>
<td>-0.75</td>
<td>-0.025</td>
<td>-0.74</td>
<td>0.71</td>
<td>-0.349</td>
<td>1.45</td>
<td>-0.78</td>
</tr>
<tr>
<td>d − p</td>
<td>0.953</td>
<td><strong>23.03</strong></td>
<td>0.953</td>
<td><strong>23.02</strong></td>
<td>90.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.039</td>
<td>0.63</td>
<td>0.045</td>
<td>0.74</td>
<td>0.83</td>
<td>0.224</td>
<td>0.65</td>
<td>-2.23</td>
</tr>
<tr>
<td>Δd</td>
<td>-0.130</td>
<td>-1.92</td>
<td>-0.136</td>
<td>-1.95</td>
<td>4.67</td>
<td>-0.745</td>
<td>0.76</td>
<td>-2.22</td>
</tr>
<tr>
<td>d − p</td>
<td>0.857</td>
<td><strong>16.95</strong></td>
<td>0.863</td>
<td><strong>16.82</strong></td>
<td>71.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C (VW)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.090</td>
<td>1.79</td>
<td>0.093</td>
<td><strong>1.83</strong></td>
<td>3.72</td>
<td>1.027</td>
<td><strong>2.40</strong></td>
<td>0.06</td>
</tr>
<tr>
<td>Δd</td>
<td>0.005</td>
<td>0.13</td>
<td>0.002</td>
<td>0.06</td>
<td>0.02</td>
<td>0.054</td>
<td><strong>2.42</strong></td>
<td>0.12</td>
</tr>
<tr>
<td>d − p</td>
<td>0.945</td>
<td><strong>20.65</strong></td>
<td>0.947</td>
<td><strong>20.57</strong></td>
<td>87.78</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Probability values for return and dividend growth slopes
This table reports the fractions of Monte-Carlo simulated values of the return \( (b^K) \) and dividend growth coefficients \( (b^K_d) \) that are higher/lower than the respective estimates from the original sample \( (b^K_{r,0}, b^K_{d,0}) \). The predictive variable is the dividend yield and the forecasting horizon is one year (Panels A and C) and 20 years ahead (Panels B and D). The sample corresponds to annual data for 1928–2010. The Monte-Carlo simulation is conducted under two different null hypotheses, no return and no dividend growth predictability. The equity portfolios consist of small stocks, big stocks, growth stocks, and value stocks. For details on the Monte-Carlo simulation see Section 5.

<table>
<thead>
<tr>
<th>Panel</th>
<th>(( b^K_r &lt; b^K_{r,0}, b^K_d &gt; b^K_{d,0} ))</th>
<th>(( b^K_r &gt; b^K_{r,0}, b^K_d &gt; b^K_{d,0} ))</th>
<th>(( b^K_r &lt; b^K_{r,0}, b^K_d &lt; b^K_{d,0} ))</th>
<th>(( b^K_r &gt; b^K_{r,0}, b^K_d &lt; b^K_{d,0} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A (( K = 1 ), no return predictability)</td>
<td>Small 0.06 0.05 0.70 0.19</td>
<td>Big 0.00 0.02 0.77 0.21</td>
<td>Growth 0.01 0.09 0.61 0.30</td>
<td>Value 0.05 0.07 0.69 0.19</td>
</tr>
<tr>
<td>Panel B (( K = 20 ), no return predictability)</td>
<td>Small 0.01 0.15 0.84 0.00</td>
<td>Big 0.00 0.01 0.96 0.03</td>
<td>Growth 0.00 0.13 0.76 0.11</td>
<td>Value 0.00 0.16 0.79 0.05</td>
</tr>
<tr>
<td>Panel C (( K = 1 ), no dividend predictability)</td>
<td>Small 0.00 0.94 0.00 0.06</td>
<td>Big 0.02 0.34 0.29 0.35</td>
<td>Growth 0.02 0.60 0.10 0.27</td>
<td>Value 0.00 0.90 0.00 0.10</td>
</tr>
<tr>
<td>Panel D (( K = 20 ), no dividend predictability)</td>
<td>Small 0.00 1.00 0.00 0.00</td>
<td>Big 0.01 0.33 0.50 0.16</td>
<td>Growth 0.01 0.70 0.18 0.11</td>
<td>Value 0.00 0.99 0.00 0.00</td>
</tr>
</tbody>
</table>
Figure 1: Dividend-to-price ratios across portfolios
This figure plots the time-series for the portfolio dividend-to-price ratios. The portfolios are sorted on size (Panel A) and book-to-market (Panel B). L denotes the 30% of stocks with lowest size or BM, while H stands for the 30% of stocks with greatest size or BM. The sample is 1928–2010.
Figure 2: Dividend growth rates across portfolios
This figure plots the time-series for the portfolio gross dividend growth rates. The portfolios are sorted on size (Panel A) and book-to-market (Panel B). L denotes the 30% of stocks with lowest size or BM, while H stands for the 30% of stocks with greatest size or BM. The sample is 1928–2010.
Figure 3: Term structure of long-run coefficients: Size portfolios
This figure plots the term structure of the long-run predictive coefficients, and respective t-statistics, for the case of size portfolios. The predictive slopes are associated with the log return \(r\), log dividend growth \(d\), and log dividend-to-price ratio \(dp\). The forecasting variable is the log dividend-to-price ratio in all three cases. “Sum” denotes the value of the variance decomposition, in %. Panels A and B present the results for small stocks, while Panels C and D show the results for big stocks. Panels E and F are related to the value-weighted stock market index (VW). The long-run coefficients are measured in %, and \(K\) represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The sample is 1928–2010.
Figure 4: Term structure of long-run coefficients: Book-to-market portfolios
This figure plots the term structure of the long-run predictive coefficients, and respective t-statistics, for the case of BM portfolios. The predictive slopes are associated with the log return ($r$), log dividend growth ($d$), and log dividend-to-price ratio ($dp$). The forecasting variable is the log dividend-to-price ratio in all three cases. “Sum” denotes the value of the variance decomposition, in %. Panels A and B present the results for growth stocks, while Panels C and D show the results for value stocks. Panels E and F are related to the value-weighted stock market index (VW). The long-run coefficients are measured in %, and $K$ represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The sample is 1928–2010.
Figure 5: VAR-based term structure of long-run coefficients: Size portfolios

This figure plots the VAR-based term structure of the long-run predictive coefficients, and respective \( t \)-statistics, for the case of size portfolios. The predictive slopes are associated with the log return \( (r) \), log dividend growth \( (d) \), and log dividend-to-price ratio \( (dp) \). The forecasting variable is the log dividend-to-price ratio in all three cases. “Sum” denotes the value of the variance decomposition, in \%. Panels A and B present the results for small stocks, while Panels C and D show the results for big stocks. Panels E and F are related to the value-weighted stock market index (VW). The long-run coefficients are measured in \%, and \( K \) represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The sample is 1928–2010.
Figure 6: VAR-based term structure of long-run coefficients: Book-to-market portfolios
This figure plots the VAR-based term structure of the long-run predictive coefficients, and respective \( t \)-statistics, for the case of BM portfolios. The predictive slopes are associated with the log return \( (r) \), log dividend growth \( (d) \), and log dividend-to-price ratio \( (dp) \). The forecasting variable is the log dividend-to-price ratio in all three cases. “Sum” denotes the value of the variance decomposition, in %. Panels A and B present the results for growth stocks, while Panels C and D show the results for value stocks. Panels E and F are related to the value-weighted stock market index (VW). The long-run coefficients are measured in \( \% \), and \( K \) represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The sample is 1928–2010.
Figure 7: Term-structure of simulated p-values: no return predictability
This figure plots the simulated p-values for the return ($r$) and dividend growth ($d$) slopes from a Monte-Carlo simulation with 10,000 replications under the null of no return predictability. The predictive variable is the dividend yield. The numbers indicate the fraction of pseudo samples under which the return or dividend growth coefficient is higher than the corresponding estimates from the original sample. The line labeled $r + d$ shows the fraction of samples in which both coefficients are jointly greater than the corresponding sample estimates. $K$ represents the number of years ahead. The sample is 1928–2010. For details on the Monte-Carlo simulation see Section 5.
Figure 8: Term-structure of simulated p-values: no dividend predictability
This figure plots the simulated p-values for the return \((r)\) and dividend growth \((d)\) slopes from a Monte-Carlo simulation with 10,000 replications under the null of no dividend growth predictability. The predictive variable is the dividend yield. The numbers indicate the fraction of pseudo samples under which the return or dividend growth coefficient is lower than the corresponding estimates from the original sample. The line labeled \(r + d\) shows the fraction of samples in which both coefficients are jointly lower than the corresponding sample estimates. \(K\) represents the number of years ahead. The sample is 1928–2010. For details on the Monte-Carlo simulation see Section 5.