Dividend Yields, Dividend Growth, and Return Predictability in the Cross Section of Stocks

Paulo Maio and Pedro Santa-Clara*

Abstract

There is a generalized conviction that variation in dividend yields is exclusively related to expected returns and not to expected dividend growth, for example, Cochrane's (2011) presidential address. We show that this pattern, although valid for the aggregate stock market, is not true for portfolios of small and value stocks, where dividend yields are related mainly to future dividend changes. Thus, the variance decomposition associated with the aggregate dividend yield has important heterogeneity in the cross section of equities. Our results are robust to different forecasting horizons, econometric methodology (long-horizon regressions or first-order vector autoregression), and alternative decomposition based on excess returns.

I. Introduction

There is a generalized conviction that variation in dividend yields is exclusively related to expected returns and not to expected dividend growth, for example, Cochrane's (2011) presidential address. We extend the analysis conducted in Cochrane (2008), (2011) to equity portfolios sorted on size and book-to-market (BM) ratio. Our goal is to assess whether the results obtained in these studies extend to disaggregated portfolios sorted on these characteristics. Indeed this finding 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 weighted regressions at several forecasting horizons, leading to a term structure of predictive

coefficients at horizons between 1 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 long-run return and dividend growth predictability drive the variation in the respective dividend-to-price ratio. 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.

These conclusions are qualitatively similar if we compute the variance decomposition for the dividend yield based on the implied estimates from a first-order vector autoregression (VAR), as is usually done in the related literature. We conduct a Monte Carlo simulation to analyze the size and power of the asymptotic t-statistics associated with the VAR-based predictive slopes and also to analyze the finite-sample distribution of these coefficients. The results show that the VAR-based asymptotic t-statistics exhibit reasonable size and power, and moreover, we cannot reject dividend growth predictability for both small and value stocks.

Our benchmark results based on the long-horizon regressions remain reasonably robust when we conduct several alternative tests such as computing a bootstrap-based inference, estimating the variance decomposition for the postwar period, and estimating an alternative variance decomposition based on excess returns and interest rates instead of nominal stock returns. We also conduct a variance decomposition for double-sorted equity portfolios: small-growth, small-value, large-growth, and large-value. We find that the large dividend growth predictability observed for small stocks seems concentrated on small-value stocks, since for small-growth stocks, dividend growth predictability plays no role in explaining the current dividend-to-price ratio. On the other hand, the large share of long-run return predictability observed for large stocks holds only for large-growth stocks, not for large-value stocks, in which case cash flow predictability is the key driving force at long horizons. Moreover, the large share of dividend growth predictability (and small amount of return predictability) verified for the value portfolio is robust on size, that is, it holds for both small-value and big-value stocks. On the other hand, while there is no cash flow predictability for small-growth stocks, it turns out that for large-growth stocks, dividend growth predictability plays a significant role at long horizons.

The results in this paper, although simple, have important implications not only for the stock 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) is a good proxy for expected stock returns (discount rates).[1] Our findings show that while this might represent a good approximation for the value-weighted (VW) market

---

[1]For 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), Petkova and Zhang (2005), and Maio (2013a), among others). In the intertemporal CAPM 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), and Maio and Santa-Clara (2012),

index or some categories of stocks, it is certainly not the case for other categories of stocks.

Our work is related to the large amount of literature that uses aggregate equity financial ratios such as dividend yield, earnings yield, or BM ratio to forecast stock market returns.[2] Specifically, our work is closely related to 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), Van Binsbergen and Koijen (2010), Engsted and Pedersen (2010), Lacerda and Santa-Clara (2010), Ang (2012), Chen, Da, and Priestley (2012), and Engsted, Pedersen, and Tanggaard (2012), among others. Koijen and Van Nieuwerburgh (2011) provide a survey on this area of research.[3] 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; instead, it must 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.

Among the papers that analyze predictability from the dividend yield at the equity portfolio level, Cochrane ((2011), Appendix B.4) conducts forecasting panel regressions for portfolios sorted on size and BM. However, he reports only the average predictive slopes; thus, his analysis does not show the different degrees of predictive performance across the different portfolios (which cannot be detected from the cross-sectional average slopes). Thus, Cochrane (2011) does not show which portfolios (within each sorting group) exhibit larger return or dividend growth predictability from the respective dividend yield, which represents the core of our analysis. Moreover, his estimates are based on a single-period forecasting regression, while we conduct multiple-horizon forecasting regressions on the dividend yield to infer how the forecasting patterns change across the forecasting horizon. Chen et al. (2012) also look at the return/dividend growth predictability among portfolios, but they use different portfolio sorts than size and BM. Moreover, they analyze only the very long-run (infinite horizon) predictability (i.e., they do not compute the dividend yield variance decomposition at short

---

among others). In the portfolio choice literature, expected stock returns, and thus dynamic portfolio rules, are often linear in the dividend-to-price ratio (e.g., Campbell and Viceira (1999), Campbell, Chan, and Viceira (2003), and Brandt and Santa-Clara (2006)).

[2]An incomplete list includes Campbell and Shiller (1988a), (1988b), Fama and French (1988), (1989), Cochrane (1992), Hodrick (1992), Goetzmann and Jorion (1993), Kothari and Shanken (1997), Pontiff and Schall (1998), Lewellen (2004), Campbell and Yogo (2006), and Ang and Bekaert (2007).

[3]Other papers analyze the predictability from alternative financial ratios (e.g., earnings yield, BM ratio, and payout yield) also in relation with present-value decompositions (e.g., Cohen, Polk, and Vuolteenaho (2003), Larrain and Yogo (2008), Chen et al. (2012), Kelly and Pruitt (2013), Maio (2013b), and Maio and Xu (2014)).

and intermediate horizons). Additionally, their long-run coefficients are implied from a first-order VAR, while we also compute the long-horizon coefficients directly from weighted long-horizon regressions. Kelly and Pruitt (2013) use equity portfolio dividend yields to forecast returns and dividend growth but for the market portfolio rather than disaggregated portfolios.

The paper is organized as follows: In Section II, we describe the data and methodology. Section III presents the dividend yield variance decomposition for portfolios sorted on size and BM from long-horizon weighted regressions. In Section IV, we conduct an alternative variance decomposition based on a first-order VAR. In Section V, we conduct several robustness checks. Section VI presents the results from Monte Carlo simulations, and Section VII concludes.

II.   Data and Methodology

A.   Methodology

Unlike some of the previous related work (e.g., Chen (2009), Chen et al. (2012), and Rangvid, Schmeling, and Schrimpf (2014)), 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.[4] 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 multiperiod 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 either returns or dividend growth to shocks in the current dividend yield. Thus, the long-horizon regressions provide 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 statistical power and misspecification. In Section IV, we present a variance decomposition based on the first-order VAR, and in Section V, we analyze the finite-sample distribution of the slopes from the long-horizon regressions by conducting a bootstrap simulation.

Following Campbell and Shiller (1988a), the dynamic accounting identity for dp can be represented as

$$(1) \qquad \mathrm{dp}_t \;\; = \;\; -\frac{c(1-\rho^K)}{1-\rho} + \sum_{j=1}^{K} \rho^{j-1} r_{t+j} - \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} + \rho^K \mathrm{dp}_{t+K},$$

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

---

[4]Cochrane (2008), (2011) and Maio and Xu (2014) use a similar approach.

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,

$$(2) \qquad \sum_{j=1}^{K} \rho^{j-1} r_{t+j} = a_r^K + b_r^K \mathrm{dp}_t + \varepsilon_{t+K}^r,$$

$$(3) \qquad \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} = a_d^K + b_d^K \mathrm{dp}_t + \varepsilon_{t+K}^d,$$

$$(4) \qquad \rho^K \mathrm{dp}_{t+K} = a_{\mathrm{dp}}^K + b_{\mathrm{dp}}^K \mathrm{dp}_t + \varepsilon_{t+K}^{\mathrm{dp}},$$

where the t-statistics for the direct predictive slopes are based on Newey and West (1987) standard errors with K lags.[5]

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 dp, at horizon K,

$$(5) \qquad 1 = b_r^K - b_d^K + b_{\mathrm{dp}}^K,$$

which can be interpreted as a variance decomposition for the log dividend yield. The predictive coefficients b_r^K, $-b_d^K$, and b_{dp}^K represent the fraction of the variance of current dp attributable to return, dividend growth, and dividend yield predictability, respectively.[6]

B. Data and Variables

We estimate the predictive regressions using annual data for the 1928–2010 period. The return data on the VW stock index, with and without dividends, are obtained from the Center for Research in Security Prices. 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, ρ, associated with the stock index is 0.965. The descriptive statistics in Table 1 show that the aggregate dividend growth has a 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 BM available from Kenneth French's Web page (http://mba .tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). For each characteristic, we use the portfolio containing the bottom 30% of stocks (denoted by L)

---

[5]An alternative estimation of the multiple-horizon predictive coefficients relies on a weighted sum of the forecasting slopes for each forecasting horizon, $\sum_{j=1}^{K} \rho^{j-1} b_r^j$, where b_r^j is estimated from the following long-horizon regression:

$$r_{t+j} = a_r^j + b_r^j \mathrm{dp}_t + \varepsilon_{t+j}^r, \quad j = 1, \dots, K.$$

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

[6]Cohen et al. (2003) derive a similar K-period variance decomposition for the log BM ratio.

TABLE 1

Descriptive Statistics

Table 1 reports descriptive statistics for the log stock return (r), log dividend growth (Δd), and log dividend-to-price ratio (dp). 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. ϕ designates the first-order autocorrelation.

	Mean	Stdev.	Min.	Max.	ϕ
Panel A. r					
VW	0.09	0.20	−0.59	0.45	0.05
SL	0.11	0.31	−0.77	0.93	0.12
SH	0.09	0.19	−0.57	0.43	0.07
BML	0.08	0.20	−0.45	0.39	0.02
BMH	0.12	0.26	−0.82	0.79	0.03
Panel B. Δd					
VW	0.04	0.14	−0.38	0.37	−0.07
SL	0.08	0.39	−1.89	1.47	−0.27
SH	0.04	0.14	−0.33	0.32	−0.06
BML	0.04	0.16	−0.34	0.43	−0.09
BMH	0.06	0.36	−2.08	1.16	0.18
Panel C. dp					
VW	−3.33	0.43	−4.50	−2.63	0.94
SL	−3.85	0.72	−5.98	−2.70	0.83
SH	−3.31	0.44	−4.56	−2.60	0.95
BML	−3.55	0.54	−4.85	−2.51	0.95
BMH	−3.25	0.60	−5.32	−2.43	0.86

and the portfolio with the top 30% of stocks (H). The reason for not using a greater number of portfolios within each sorting variable (e.g., 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 1980s. The dividend yields for big capitalization stocks tend to be higher than those 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 1930s, value stocks tend to have significantly 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 1980s is significantly more severe for big and value stocks in comparison to small and growth stocks, respectively.

From Panel C of Table 1, we can see that the log dividend yield of small stocks is more volatile than the corresponding log ratio for big stocks (standard deviation of 0.72 vs. 0.44), while big stocks have a significantly more persistent dividend-to-price ratio (0.95 vs. 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 vs. 0.54), while growth stocks have a more persistent multiple (0.95 vs. 0.86). The estimates for ρ 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 present the time series for portfolio dividend growth rates (in levels). We can see that dividend growth was quite volatile during the great

FIGURE 1

Dividend-to-Price Ratios across Portfolios

Figure 1 plots the time series for the portfolio dividend-to-price ratios (in levels). The portfolios are sorted on size (Graph A) and BM ratio (Graph 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.
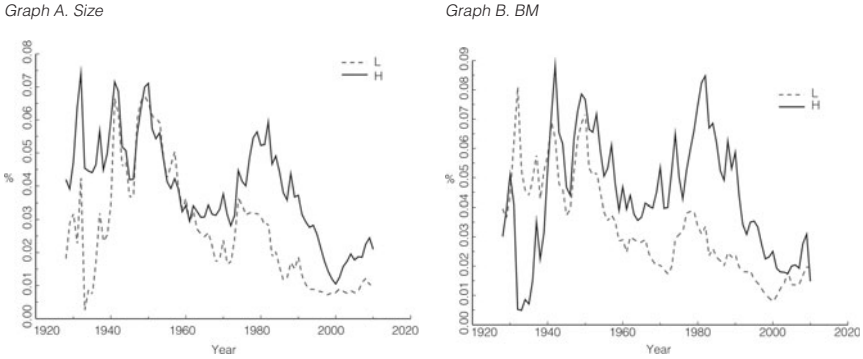
Graph A. Size

Graph B. BM

FIGURE 2

Dividend Growth Rates across Portfolios

Figure 2 plots the time series for the portfolio dividend growth rates (in levels). The portfolios are sorted on size (Graph A) and BM ratio (Graph 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.

Graph A. Size

Graph B. BM
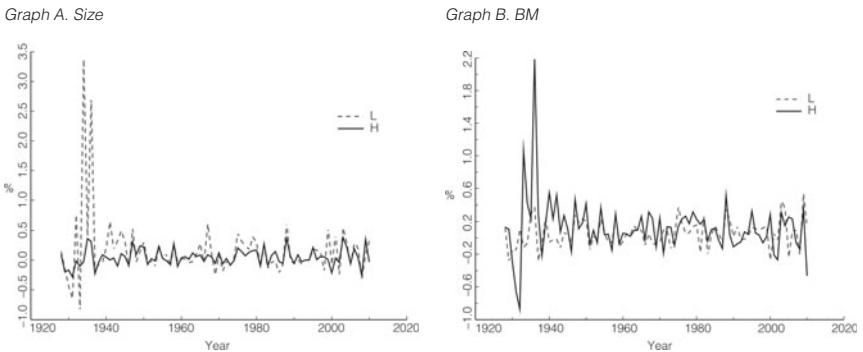
depression, especially for small and value stocks. The standard deviation calculations in Panel B of Table 1 show that small and value stocks exhibit much more volatile dividend growth than big and growth stocks, respectively. Log dividend growth is weakly and negatively autocorrelated for small stocks (-0.27), while for value stocks, we have a small positive autocorrelation.

III. Predictability of Size and BM Portfolios
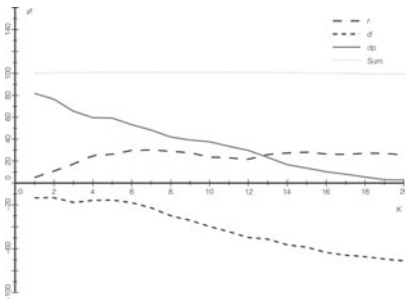
A. Size Portfolios

The term structure of predictive coefficients and respective *t*-statistics for the variance decompositions associated with small and large stocks is shown

in Figure 3. In the case of small stocks (Graph 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 forecasting horizons between 6 and 8 years). For big stocks (Graph 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
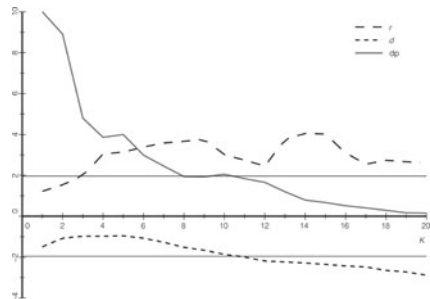
FIGURE 3

Term Structure of Coefficients: Size Portfolios

Figure 3 plots the term structure of the long-horizon 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 percent. The long-run coefficients are measured in percent, 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.
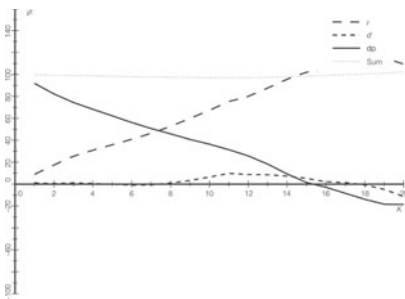
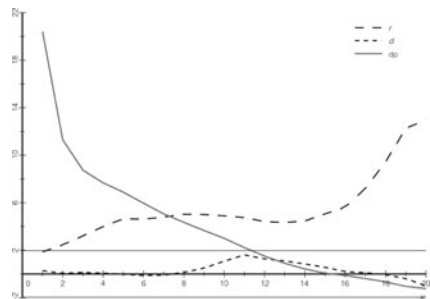Graph A. Small: Slopes

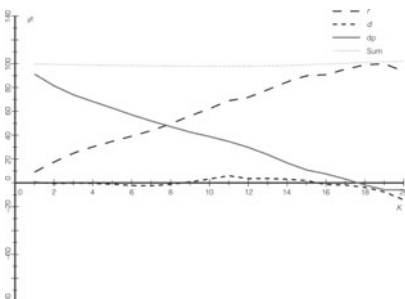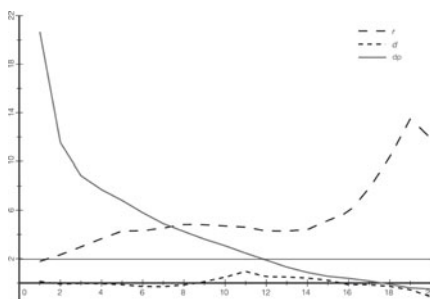Graph B. Small: t-Statistics

Graph C. Large: Slopes

Graph D. Large: t-Statistics

Graph E. Market: Slopes

Graph F. Market: t-Statistics

long-horizon predictability of the future dividend yield has the "wrong" sign (about -20% at $K = 20$).[7]

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 regressions 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 3 years (although the magnitudes of the *t*-ratios are smaller than for the large portfolio, especially at long horizons). The coefficients associated with the future dividend yield are statistically significant at short and intermediate horizons (until 10 years) for both portfolios but become insignificant at long horizons.

As a reference point, we present the variance decomposition for the stock index in Graphs E and F of Figure 3. We can see that the plots of the term structure of predictive slopes associated with 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 (see Cochrane (2008), (2011), Chen (2009)). This result is also consistent with the fact that the VW 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 hides 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, and these patterns are especially notable at long horizons.

B.   BM Portfolios

Next, we conduct a similar analysis for the BM 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 (Graphs 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 confirming 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 4 years, while the return coefficients are not significant (at the 5% level) at very short horizons and also at some of the longer horizons.
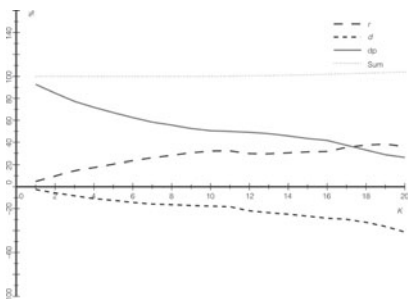
In the case of value stocks (Graphs C and D of Figure 4), the pattern of direct predictive coefficients shows that dividend growth predictability achieves

---

[7]By conducting forecasting panel regressions for aggregate country portfolios, Rangvid et al. (2014) find that there is less dividend growth predictability in countries where the size of the average firm is larger.
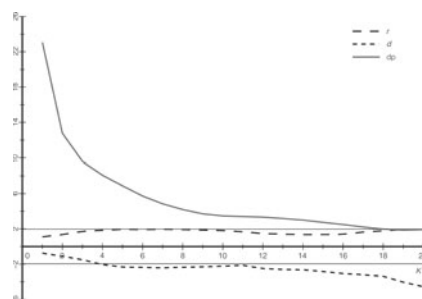
FIGURE 4

Term Structure of Coefficients: BM Portfolios

Figure 4 plots the term structure of the long-horizon 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 percent. The long-run coefficients are measured in percent, 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.
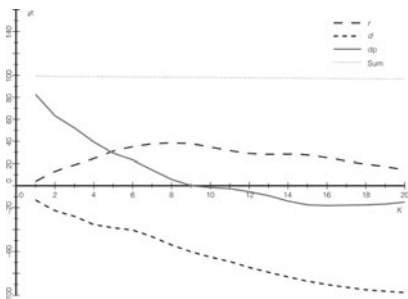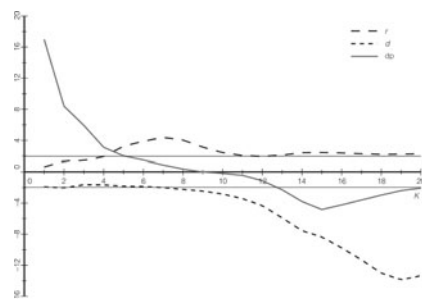
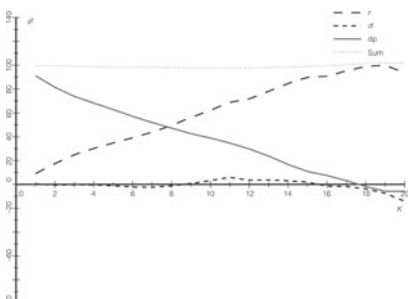Graph A. Growth: Slopes

Graph B. Growth: t-Statistics

Graph C. Value: Slopes

Graph D. Value: t-Statistics

Graph E. Market: Slopes

Graph F. Market: t-Statistics

weights close to 100% at the 20-year horizon. In comparison, the share of return predictability never exceeds 40% (which is obtained at $K = 8$) and decreases to values around 20% at the 20-year horizon. The slopes associated with future dp converge to 0 much faster than in the case of the growth portfolio, turning negative for horizons beyond 9 years. The *t*-statistics indicate that the dividend growth slopes are highly significant at intermediate and long horizons, while there are

some periods (short horizons) for which the return coefficients are not statistically significant at the 5% level.

The variance decomposition associated with the stock index (Graphs E and F of Figure 4) indicates significantly more return predictability and less dividend growth predictability than that associated with the growth portfolio. This suggests that the predictability pattern estimated for the VW index might be the result of some large growth stocks. On the other hand, the variance decomposition for the stock market looks basically the opposite of the decomposition associated with the value portfolio.

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 at most horizons, while the predictability of the future dividend yield also plays an important role.[8]

Why is future dividend growth the main influence on the current dividend yield of value stocks, while for growth stocks, return predictability also plays an important role? A possible explanation is that many growth stocks are firms with a longer duration of cash flows, which are not expected to deliver positive cash flows for several periods in the future (see Cornell (1999), Lettau and Wachter (2007) for a discussion). Thus, changes in their current valuations 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 (close to) 0 expected dividend growth rates for these stocks in the short and medium term. Put differently, the dividend yield of many growth stocks is more sensitive to variations in discount rates than to changes in dividend growth.[9]

IV.   VAR-Based Results

A.   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), Engsted and Pedersen (2010), Chen et al. (2012), among others.

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

$$(6) \qquad\qquad r_{t+1} \;=\; a_r + b_r \mathrm{dp}_t + \varepsilon^r_{t+1},$$

$$(7) \qquad\qquad \Delta d_{t+1} \;=\; a_d + b_d \mathrm{dp}_t + \varepsilon^d_{t+1},$$

$$(8) \qquad\qquad \mathrm{dp}_{t+1} \;=\; a_{\mathrm{dp}} + \phi \mathrm{dp}_t + \varepsilon^{\mathrm{dp}}_{t+1},$$

---

[8]By using a different methodology, Rytchkov (2010) also finds more return predictability for growth versus value stocks.

[9]In related work, Campbell and Vuolteenaho (2004) argue that the returns of value stocks are more sensitive to shocks in future aggregate cash flows than are the returns of growth stocks; that is, value stocks have greater cash flow betas. On the other hand, these authors 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; similarly, the cash flows of growth stocks determine their discount rate betas.

where the εs represent error terms. The VAR above is estimated by ordinary least squares (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 dp, at every horizon K,

$$
\begin{aligned}
1 &= b_r^K - b_d^K + b_{dp}^K, \\
b_r^K &\equiv \frac{b_r(1 - \rho^K \phi^K)}{1 - \rho\phi}, \\
b_d^K &\equiv \frac{b_d(1 - \rho^K \phi^K)}{1 - \rho\phi}, \\
b_{dp}^K &\equiv \rho^K \phi^K,
\end{aligned}
\tag{9}
$$

which represents the variance decomposition shown in Cochrane (2008), (2011) or Engsted and Pedersen (2010). The t-statistics associated with the predictive coefficients in expression (9) are computed from the t-statistics for the VAR slopes by using the delta method (details are available in the Online Appendix at www.jfqa.org). 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 weighted regressions but rather implied from the VAR estimates. If the first-order VAR does not fully capture the multiperiod dynamics of the data-generating process for r, dp, and Δd, 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 ($K \to \infty$):

$$
\begin{aligned}
1 &= b_r^{LR} - b_d^{LR}, \\
b_r^{LR} &\equiv \frac{b_r}{1 - \rho\phi}, \\
b_d^{LR} &\equiv \frac{b_d}{1 - \rho\phi}.
\end{aligned}
\tag{10}
$$

In this long-run decomposition, all the variation in the current dividend yield is tied to either return or dividend growth predictability, since the predictability of the future dividend yield vanishes out at a very long horizon.

The t-statistics for the long-run coefficients, b_r^{LR}, b_d^{LR}, 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,

$$H_0 : b_r^{LR} = 0, \quad b_d^{LR} = -1,$$

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

$$H_0 : b_r^{LR} = 1, \quad b_d^{LR} = 0.$$

B. Size Portfolios

The VAR estimation results for the size portfolios are presented in Panel A of Table 2. The return slope for the small portfolio is 0.05, and this estimate is not statistically 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 indicated in Section II. The R^2 in the dividend growth equation is 6.18%, about 4 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 slope of 0.83 versus 0.95.

TABLE 2

VAR Estimates

Table 2 reports the one-period (restricted) VAR estimation results for portfolios sorted on size (Panel A) and BM ratio (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 (dp). $b(\phi)$ denotes the VAR slopes associated with lagged dp, while t denotes the respective Newey and West (1987) t-statistics (calculated with one lag). $b^i(\phi^i)$ denotes the slope estimates implied from the variance decomposition for dp, and t denotes the respective asymptotic t-statistics computed under the delta method. R^2(%) is the coefficient of determination for each equation in the VAR, in percent. b^{LR} denotes the long-run coefficients (infinite horizon). $t(b_r^{LR} = 0)$ and $t(b_r^{LR} = 1)$ denote the t-statistics associated with the null hypotheses $H_0 : b_r^{LR} = 0$, $b_d^{LR} = -1$ and $H_0 : b_r^{LR} = 1$, $b_d^{LR} = 0$, respectively. The sample corresponds to annual data for 1928–2010. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

		$b(\phi)$	t	$b^i(\phi^i)$	t	R^2(%)	b^{LR}	$t(b_r^{LR} = 0)$	$t(b_r^{LR} = 1)$
Panel A. Size									
Small	r	0.050	1.22	0.048	1.17	1.34	0.273	1.08	−2.88***
	Δd	−0.135	−1.49	−0.133	−1.46	6.18	−0.740	1.05	−2.97***
	dp	0.834	9.99***	0.832	9.97***	68.81	—	—	—
Big	r	0.089	1.87*	0.093	1.94*	4.17	1.079	2.43**	0.18
	Δd	0.010	0.28	0.007	0.18	0.10	0.121	2.43**	0.26
	dp	0.951	20.36***	0.954	20.25***	88.74	—	—	—
Panel B. BM									
Growth	r	0.048	1.10	0.048	1.07	1.75	0.661	1.50	−0.77
	Δd	−0.026	−0.75	−0.025	−0.74	0.71	−0.349	1.45	−0.78
	dp	0.953	23.03***	0.953	23.02***	90.74	—	—	—
Value	r	0.039	0.63	0.045	0.74	0.83	0.224	0.65	−2.23**
	Δd	−0.130	−1.92*	−0.136	−1.95*	4.67	−0.745	0.76	−2.22**
	dp	0.857	16.95***	0.863	16.82***	71.30	—	—	—
Panel C. VW									
	r	0.090	1.79*	0.093	1.83*	3.72	1.027	2.40**	0.06
	Δd	0.005	0.13	0.002	0.06	0.02	0.054	2.42**	0.12
	dp	0.945	20.65***	0.947	20.57***	87.78	—	—	—

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 R^2 is 4.17%. The dividend growth slope has the wrong sign, but this estimate is clearly insignificant. The estimate for ϕ is 0.95, confirming that the dividend yield of big stocks is significantly more persistent than that for small stocks. The VAR estimation results for the market index, displayed in Panel C of

Table 2, show that the predictive slopes and R^2 estimates 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,

$$(11) \qquad\qquad\qquad 1 \;=\; b_r - b_d + \rho\phi,$$

are very similar to the direct VAR estimates, showing that this present-value decomposition works quite well at the 1-year horizon.

Overall, these results show that big capitalization stocks have significantly greater short-run return predictability from the dividend yield than do small stocks. Second, small caps have some relevant short-run dividend growth predictability from dp, while the same does not occur with large stocks.

The term structure of the VAR-based variance decomposition for the size portfolios is displayed in Figure 5. In the case of small stocks (Graph A), even at long horizons, the bulk of variation in dp is associated with dividend growth predictability (almost 80%) rather than return predictability, which does not go above 30%. In the case of large stocks (Graph 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 (Graph E) is quite similar to the corresponding decomposition for large caps, as in the last section.

The analysis of the t-statistics shows 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 2 years in the case of the big portfolio. In contrast, the dividend coefficients are statistically significant for horizons greater than 6 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-statistic $= 1.08$), whereas we strongly reject the null of no dividend growth predictability (t-statistic $= -2.97$).[10] In the case of big stocks, we have a totally different picture: the estimates for b_r^{LR} and b_d^{LR} 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 do not reject the null of no dividend growth predictability (t-ratio close to 0).
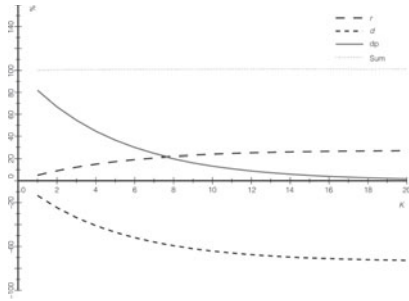
When compared with the benchmark variance decomposition estimated in the last section, the results are relatively similar in the sense that with both

---

[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.
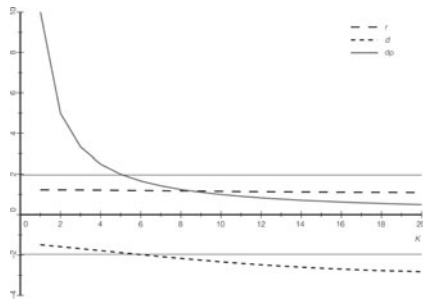
FIGURE 5

VAR-Based Term Structure of Coefficients: Size Portfolios

Figure 5 plots the VAR-based term structure of the long-horizon 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 percent. The long-run coefficients are measured in percent, 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.
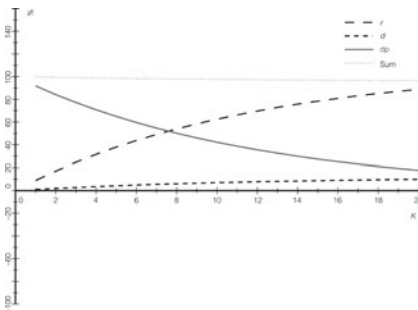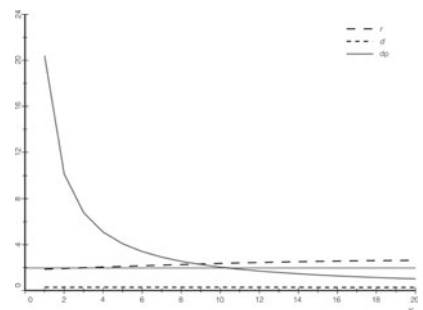
Graph A. Small: Slopes
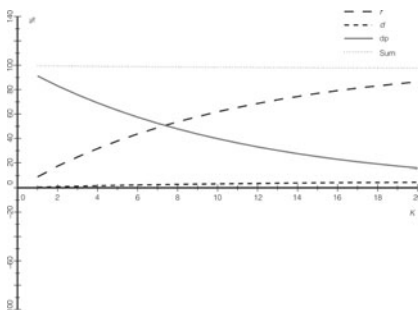

Graph B. Small: t-Statistics
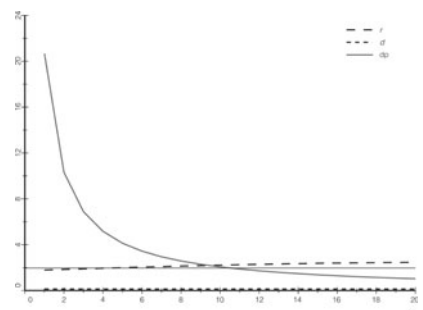

Graph C. Large: Slopes

Graph D. Large: t-Statistics

Graph E. Market: Slopes

Graph F. Market: t-Statistics

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 variation in the dividend-to-price ratio is related to dividend growth predictability.

C. BM Portfolios

The VAR estimation results for the value and growth portfolios are depicted in Panel B of Table 2. Growth stocks have a return predictive slope of 0.05 and a corresponding R^2 of 1.75%. For value stocks, the amount of return predictability is smaller with a coefficient of 0.04 and an 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 clearly 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 R^2 estimate is 4.67%, much higher than the fit for the return equation associated with 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 in Figure 6. In the case of growth stocks, return predictability is the dominant source of dividend yield variance, although the weights are significantly lower than those of 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 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. When compared to the variance decomposition associated with the stock index (Graph E), the growth portfolio has less return predictability and more dividend growth predictability, as in the last section.

For value stocks, most of the variation in the current dividend yield is a result of dividend growth predictability at long horizons (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 forecasting horizons, while the return slopes are not significant at any horizon.
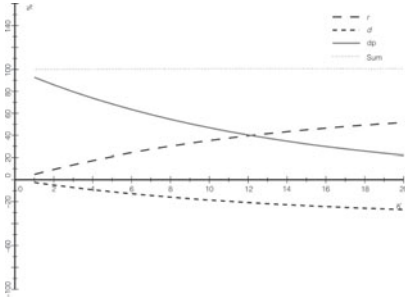
The long-run (infinite horizon) return and dividend growth slopes for growth stocks (Panel B of Table 2) are 0.66 and -0.35, respectively, confirming that long-run return predictability is the main driver of variation in the dividend yield of those stocks. However, due to the large standard errors of the VAR slopes, we cannot reject the null of no return predictability at the 5% level (t-statistic $= 1.50$). In contrast, for value stocks, the estimates for b_r^{LR} and b_d^{LR} are 0.22 and -0.75, respectively, indicating that the key driver of the dividend yield is long-run dividend growth predictability. We do not reject the null of no return predictability by a big margin (t-ratio of 0.65), while the null of no dividend growth 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 major driver of variation in the dividend yield is return predictability (although there is also some relevant predictability
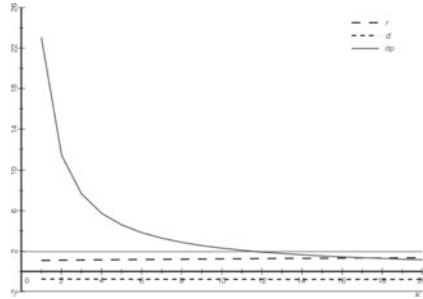
FIGURE 6

VAR-Based Term Structure of Coefficients: BM Portfolios

Figure 6 plots the VAR-based term structure of the long-horizon 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 percent. The long-run coefficients are measured in percent, 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.
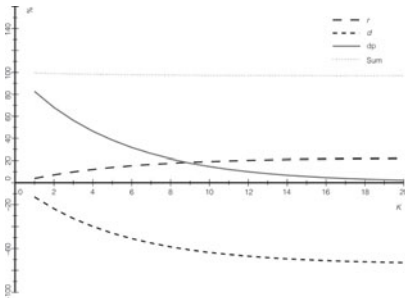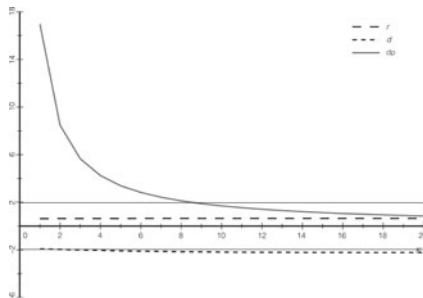
Graph A. Growth: Slopes
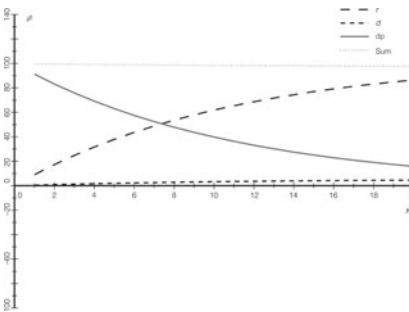
Graph B. Growth: t-Statistics

Graph C. Value: Slopes
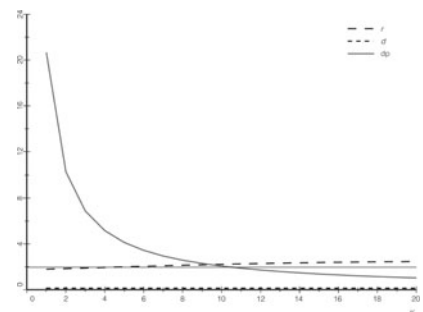
Graph D. Value: t-Statistics

Graph E. Market: Slopes

Graph F. Market: t-Statistics

of dividend growth and future dividend yield), while for value stocks, return predictability is relatively marginal.

V.   Additional Results

In this section, we conduct several robustness checks to the main analysis in Section III. Most of the results are displayed in the Online Appendix.

A.   Bootstrap-Based Inference

As an alternative to the asymptotic statistical inference conducted in Section III, we conduct a bootstrap simulation to assess the finite-sample distribution of the slope estimates from the long-horizon regressions.[11] We generate 10,000 artificial samples in which r, Δd, and dp are drawn with replacement from the original series. The artificial variables are constructed without imposing the predictive regressions; that is, we impose the null of no return/dividend growth predictability.[12] In each replication, we use a common time sequence for all three variables to account for their contemporaneous correlation (see Stambaugh (1999), Lewellen (2004)). Moreover, we use a block bootstrap, with block length of 3, to allow for the serial correlation in each variable, which is especially relevant in the case of the dividend yield. Finally, in each replication, we use the same artificial data for the regressions associated with all forecasting horizons; that is, we do not use a different draw for each horizon. By doing this, we try to capture the fact that the slopes at different forecasting horizons are correlated; for example, in the return regression at $K=2$, we have $b_r^2 = b_r^1 + \rho\beta(r_{t+2}, \mathrm{dp}_t)$ (where $\beta(y, x)$ stands for the regression coefficient of y on x) and so forth (see also Boudoukh, Richardson, and Whitelaw (2008)).[13]

To assess the individual significance of the slopes for future returns and dividend yield, we compute the respective p-values as the fraction of replications in which the pseudoestimates for the coefficients are higher than the original estimates obtained from the actual sample. In the case of the dividend growth slopes, the respective p-values are the percentage of replications in which the pseudoestimates are lower than the corresponding sample estimate.[14]

The results show that the inference based on the bootstrap is not qualitatively very different from the asymptotic inference conducted in Section III. For the small and growth portfolios, all three coefficients are strongly significant at all horizons, as indicated by the p-values around 0, thus showing stronger significance than with the asymptotic t-statistics.

In the case of the large portfolio, there is no statistical significance for the dividend growth coefficients at most horizons (p-values around 1), while the slopes associated with the future dividend yield are not significant at long horizons, in both cases, similarly to the inference based on the asymptotic t-statistics. Regarding the value portfolio, both the return and dividend growth coefficients are clearly significant, as shown by the p-values close to 0, similarly to the asymptotic inference. On the other hand, the slopes for the future dividend-to-price ratio are not significant at long horizons, as indicated by the p-values around 1.

We also use the bootstrap simulation to assess whether the differences in return and dividend growth slopes across the small-large and growth-value portfolios are statistically significant. In this case, the bootstrap simulation is the same as before, except that the artificial samples are common to the size portfolios, on

---

[11]We thank the referee for suggesting this analysis.

[12]An alternative bootstrap, in which only r and dp are simulated, yields similar results.

[13]Full details of the bootstrap algorithm are available from the authors.

[14]An alternative bootstrap, in which we use the Newey and West (1987) t-statistics rather than the slope estimates to compute the p-values, leads to similar results.

the one hand, and the BM portfolios, on the other hand. Untabulated results show that the empirical *p*-values associated with the spreads (small minus large and growth minus value) in the return and dividend growth slopes are very close to 0 at all horizons, thus showing that these spreads are statistically different than 0.

B.   Interaction between Size and BM

In the previous sections, we analyze the predictability pattern for portfolios sorted on either size or BM and observe different patterns across small-large stocks, on the one hand, and across growth-value stocks, on the other hand. An interesting question is whether this "size effect" is robust among both growth and value stocks or if the "BM effect" is robust among both small and large stocks. We can also assess the contribution of both small-growth and small-value stocks for the variance decomposition associated with the "aggregate" small portfolio and similarly for the large portfolio. Likewise, we can evaluate the contribution of both small-value and large-value stocks for the variance decomposition associated with the aggregate value portfolio and similarly for the growth portfolio.[15]

To address these issues, we conduct a dividend yield variance decomposition for four additional portfolios: small-growth, small-value, large-growth, and large-value. These are four of the six size–BM portfolios available from Kenneth French's Web page, which are obtained from the intersection of three BM portfolios (low, medium, and high) and two size portfolios (small and large). The portfolio dividend growth and dividend-to-price series are obtained from the portfolio total return and return without dividends series, as explained in Section II.

The variance decompositions associated with the small-growth and small-value portfolios are presented in Figure 7. The results show that the variance decomposition for the small-value portfolio is qualitatively similar to the corresponding decomposition associated with the small portfolio in Section III, with even more dividend growth predictability (slopes above 1 in absolute value at long horizons) and less return predictability at long horizons. Yet, for the small-growth portfolio, we have quite opposing results relative to small-value stocks: the dividend growth coefficients are around 0 at all forecasting horizons, and the variation in the current dividend yield is driven by predictability of future returns and dividend yields, with similar weights at long horizons. Therefore, the large share of dividend growth predictability observed for small stocks seems concentrated on small-value stocks, since for small-growth stocks, dividend growth predictability plays no role in explaining the current dividend-to-price ratio.

Figure 8 presents the results for the large-growth and large-value portfolios. In the case of the large-value portfolio, the bulk of variation in the dividend yield is driven by dividend growth predictability (coefficients around −100% at long horizons), while the share of return predictability is quite small, especially at long horizons (slopes around or below 20%). This result is in sharp contrast to the variance decomposition obtained for the large portfolio in Section III, in which the driving force of the current dividend yield is return predictability, especially at long horizons. Regarding the large-growth portfolio, it follows that return

---

[15]We thank the referee for suggesting this analysis.

FIGURE 7

Small-Growth and Small-Value Portfolios

Figure 7 plots the term structure of the long-horizon predictive coefficients and respective *t*-statistics for the case of the small-growth and small-value 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 percent. The long-run coefficients are measured in percent, 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.
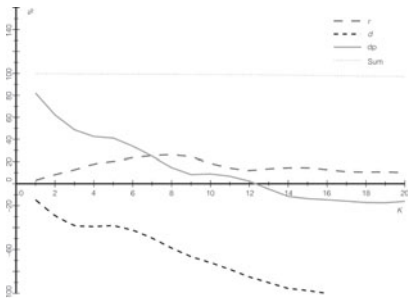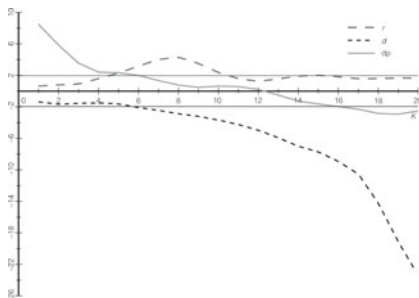
Graph A. Small-Growth: Slopes
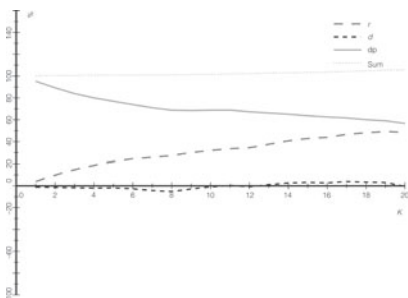
Graph B. Small-Growth: t-Statistics
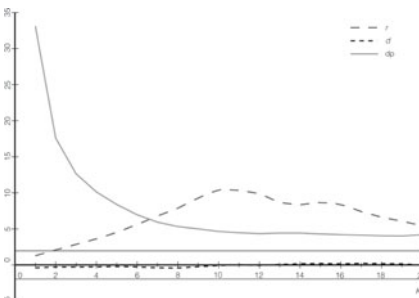
Graph C. Small-Value: Slopes
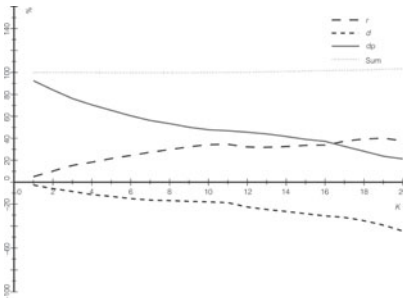
Graph D. Small-Value: t-Statistics

predictability is more important than cash flow predictability at most horizons, although at long horizons ($K \geq 18$), both have similar weights (around 40%). These results seem to indicate that the large share of return predictability observed for large stocks holds only for large-growth (and possibly large-intermediate BM stocks) not for large-value stocks in which cash flow predictability is the key driving force.

Taking Figures 7 and 8 together, we can see that the large share of dividend growth predictability (and small amount of return predictability) verified for the value portfolio in Section III is robust on size; that is, it holds for both small-value and big-value stocks. On the other hand, while there is no cash flow predictability for small-growth stocks, it turns out that for large-growth stocks, dividend growth predictability plays a significant role, particularly at long horizons. Thus, the predictability pattern for large-growth stocks explains why we observe some relevant long-run cash flow predictability for the growth portfolio in Section III. Another way to interpret these results is that the difference in the degree of cash flow predictability between small and large stocks holds only among growth stocks, being absent among value stocks. On the other hand, the spread in dividend growth

FIGURE 8

Large-Growth and Large-Value Portfolios

Figure 8 plots the term structure of the long-horizon predictive coefficients and respective *t*-statistics for the case of the large-growth and large-value 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 percent. The long-run coefficients are measured in percent, 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.
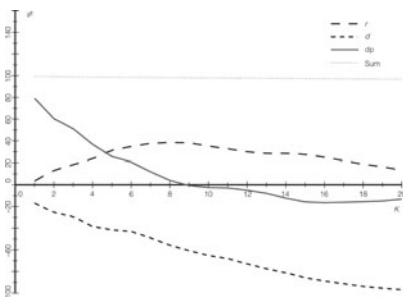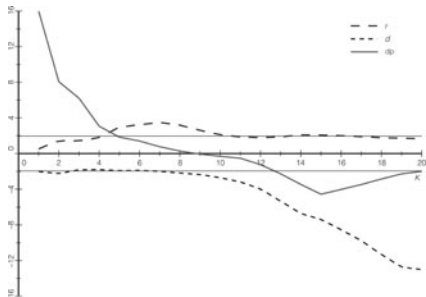
Graph A. Large-Growth: Slopes

Graph B. Large-Growth: t-Statistics
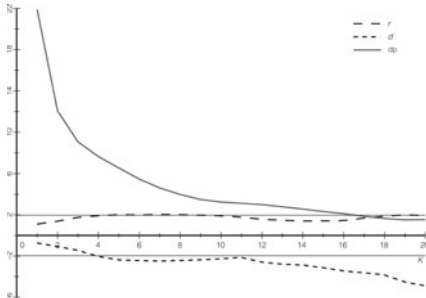
Graph C. Large-Value: Slopes

Graph D. Large-Value: t-Statistics

predictability between value and growth stocks occurs especially among small stocks and less so among large stocks.

C.   Postwar Analysis

We conduct the dividend yield variance decomposition for the postwar sample, 1946–2010, which corresponds roughly to the sample used in Cochrane (2011). Thus, we want to assess whether the predictability pattern observed in the long sample is relatively robust in this shorter period, given the high volatility of dividend growth rates for small and value stocks in the 1930s, as stressed in Section II.

The results indicate that the main difference relative to the full-sample results occurs in the variance decomposition for the small portfolio, since the dividend growth slopes are around −30% at long horizons, compared to values around −70% in the full sample. However, these coefficients are still significant at the 5% level for horizons beyond 7 years. In comparison, the return slopes are around 40% at long horizons, slightly above the estimates obtained in the full sample,

while the coefficients associated with future dp are around 40%, even at long horizons. This result suggests that the dividend yield of small stocks is more persistent in the recent sample.

In the case of the value portfolio, the bulk of variation in the current dividend yield is dividend growth predictability at long horizons, with slope estimates around -100%, and these estimates are statistically significant for $K > 16$. On the other hand, the share of return predictability at long horizons is larger than that in the full sample, with estimates above 40%. Regarding the growth portfolio, as in the full sample, both long-run dividend growth and return predictability have similar weights (around 50%), and both coefficients are significant at most horizons.

Overall, these results indicate that in the postwar period, there is a decline in the share of cash flow predictability for small stocks; still, both return and dividend growth predictability (at long horizons) have similar importance in driving the current dividend yield. On the other hand, for value stocks, the large degree of cash flow predictability remains robust in the most recent period.

D. Do Portfolio Dividend Yields Forecast Equity Premia?

We investigate the predictive role of the dividend yield for future excess stock returns. In fact, the predictability of the aggregate equity premium, rather than the predictability of the nominal or real stock returns, has been the focus of the stock predictability literature (see, e.g., Campbell and Thompson (2008), Goyal and Welch (2008)). We want to assess whether the results from the previous sections hold if we work with excess returns instead of nominal returns.

To analyze the equity premium predictability, we reorganize the Campbell and Shiller (1988a) decomposition in terms of excess returns,

$$(12) \qquad \mathrm{dp}_t = -\frac{c(1-\rho^K)}{1-\rho} + \sum_{j=1}^{K} \rho^{j-1} r_{t+j}^e - \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j}$$

$$+ \sum_{j=1}^{K} \rho^{j-1} r_{f,t+j} + \rho^K \mathrm{dp}_{t+K},$$

where $r_{f,t}$ denotes the log short-term interest rate, and $r_t^e \equiv r_t - r_{f,t}$ is the excess log stock return. Due to the fact that we are now working with excess returns, the current dividend yield should be positively correlated with an additional term: future short-term interest rates, $\sum_{j=1}^{K} \rho^{j-1} r_{f,t+j}$.

We obtain the multiperiod predictive coefficients from the following long-horizon regressions, as in Section III:

$$(13) \qquad \sum_{j=1}^{K} \rho^{j-1} r_{t+j}^e = a_r^K + b_r^K \mathrm{dp}_t + \varepsilon_{t+K}^r,$$

$$(14) \qquad \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} = a_d^K + b_d^K \mathrm{dp}_t + \varepsilon_{t+K}^d,$$

$$(15) \qquad \sum_{j=1}^{K} \rho^{j-1} r_{f,t+j} = a_f^K + b_f^K \mathrm{dp}_t + \varepsilon_{t+K}^f,$$

$$(16) \qquad \rho^K \mathrm{dp}_{t+K} = a_{\mathrm{dp}}^K + b_{\mathrm{dp}}^K \mathrm{dp}_t + \varepsilon_{t+K}^{\mathrm{dp}}.$$

The variance decomposition for the dividend yield is now given by

$$(17) \qquad 1 = b_r^K - b_d^K + b_f^K + b_{\mathrm{dp}}^K.$$

Under this decomposition, some of the variation in dp is positively associated with the predictability of future interest rates, which is captured by the predictive coefficient, b_f^K. Notice that the predictive slopes for future dividend growth and dividend yield are the same as those in the variance decomposition associated with returns. The sole difference relative to the analysis in the previous sections is that the return coefficients at every horizon are decomposed into an equity premium coefficient and an interest rate slope. The full derivation of this variance decomposition is provided in the Online Appendix.[16]

The term structure of predictive coefficients, and respective t-statistics, are displayed in the Online Appendix. The slopes associated with future dividend growth and dividend yield are the same as those in Figure 3, although we present these for comparison purposes. For small stocks, the share of interest rate predictability is quite small in magnitude, achieving values close to 0% at most horizons, and these coefficients are not statistically significant. This implies that the predictability of the excess returns of small stocks is very similar to the already relatively low return predictability found in Figure 3. In the case of the large portfolio, we have a different pattern. The slopes associated with future interest rates are largely negative (i.e., they have the wrong sign), representing at long horizons around 80% of the dividend yield variance. It follows that the equity premium coefficients account for significantly more than 100% of the variation in dp for horizons beyond 12 years. The excess return slopes are significant at all horizons, while the interest rate coefficients are significant for horizons greater than 14 years.

The predictability pattern for the growth portfolio is, to some extent, similar to that observed for the big portfolio: the interest rate coefficients are largely negative, implying that the slopes for the future equity premium are above 100% for horizons beyond 18 years. Moreover, both the equity premium and interest rate slopes are statistically significant at most forecasting horizons. In the case of the value portfolio, we have a different picture. The interest rate coefficients have the correct sign, representing about 20% of the dividend yield variation, at very long horizons. Consequently, the equity premia slopes are smaller than the already quite small return coefficients, and actually assume negative values (around -10%) at very long horizons. However, these excess return slopes are not statistically significant at most horizons, while for the interest rate slopes, there is statistical significance at all horizons.

---

[16]In his analysis with excess returns, Cochrane (2008) defines a composite term, $\Delta d_{t+1} - r_{f,t+1}$. However, there is a priori no reason for the slopes on future dividend growth and short-term interest rate to have the same magnitude, and our results below confirm that.

Thus, according to our results, the dividend yield of growth stocks predicts a decline in short-term interest rates at multiple horizons, while the dividend yield of value stocks is positively correlated with future interest rates (with lower magnitudes). This implies that there is strong equity premium predictability for growth stocks (significantly greater than the corresponding return predictability), while for value stocks, the size of equity premia predictability is even lower than the already small return predictability. Overall, the results of this subsection reinforce the findings from Section III: What drives the variation in the dividend yields of big and growth stocks is the predictability of future (excess) returns, whereas for both small and value stocks, the main driver is cash flow predictability.

VI.   Monte Carlo Simulation

In this section, we conduct a Monte Carlo simulation to analyze the size and power of the asymptotic t-statistics for the return and dividend growth predictive coefficients based on the first-order VAR estimated in Section IV. Since these t-ratios are based on the delta method, their approximation to finite samples might be poor.[17] To save space, the analysis is conducted only for the small and value portfolios, for which there is stronger evidence of dividend growth predictability, as shown in the previous sections.

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:

$$(18) \qquad \begin{pmatrix} r_{t+1} \\ \Delta d_{t+1} \\ \mathrm{dp}_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho\phi - 1 \\ \phi \end{pmatrix} \mathrm{dp}_t + \begin{pmatrix} \varepsilon_{t+1}^d - \rho\varepsilon_{t+1}^{\mathrm{dp}} \\ \varepsilon_{t+1}^d \\ \varepsilon_{t+1}^{\mathrm{dp}} \end{pmatrix}.$$

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:

$$(19) \qquad \begin{pmatrix} r_{t+1} \\ \Delta d_{t+1} \\ \mathrm{dp}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 - \rho\phi \\ 0 \\ \phi \end{pmatrix} \mathrm{dp}_t + \begin{pmatrix} \varepsilon_{t+1}^r \\ \varepsilon_{t+1}^r + \rho\varepsilon_{t+1}^{\mathrm{dp}} \\ \varepsilon_{t+1}^{\mathrm{dp}} \end{pmatrix}.$$

Notice that in both simulations, we impose 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 $\mathrm{dp}_0 \sim N\left[0, \mathrm{var}(\varepsilon_{t+1}^{\mathrm{dp}})/(1 - \phi^2)\right]$. Equipped with the artificial data, we compute the fraction of significant predictive coefficients, that is, the fraction of artificial

---

[17]We thank the referee for suggesting this analysis.

samples in which the *t*-statistics for the return slopes are higher than 1.96 (the 5% critical value), on the one hand, and the fraction of replications in which the *t*-statistics associated with the dividend growth slopes are lower than -1.96, on the other hand.

Figure 9 presents the term structure of the fractions of rejections of the null hypothesis under the two nulls considered above. Under the null of no return predictability, the fraction of replications in which the dividend growth slopes are significant is around 70% at most forecasting horizons in the case of the small portfolio, while these percentages are slightly lower (around 60%) in the case of the value portfolio. On the other hand, the fractions of pseudosamples in which the return slopes are significant are close to 0 (below 5%) for both portfolios. These results indicate that the *t*-statistics associated with the dividend growth slopes exhibit reasonable power, while the size of the *t*-statistics for the return slopes is close to 5%, the significance level associated with the asymptotic *t*-statistics in each pseudosample.

FIGURE 9

Monte Carlo Simulation: Size and Power of the VAR-Based Statistics

Figure 9 plots the size and power results from a Monte Carlo simulation (with 10,000 replications), under the nulls of no return predictability and no dividend growth predictability, for the *t*-statistics in the VAR-based predictability model. The predictive variable is the log dividend yield. The numbers indicate the fraction of pseudosamples under which the *t*-statistic associated with the return (dividend growth) coefficient is higher (lower) than 1.96 (-1.96). *K* represents the number of years ahead. The sample is 1928–2010. The analysis is conducted for small and value portfolios. For details on the Monte Carlo simulation, see Section VI.
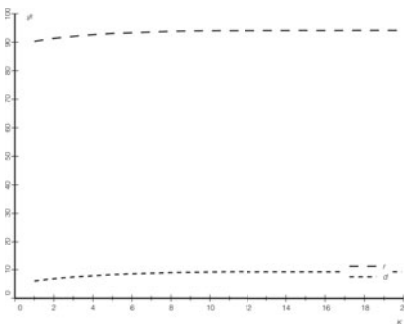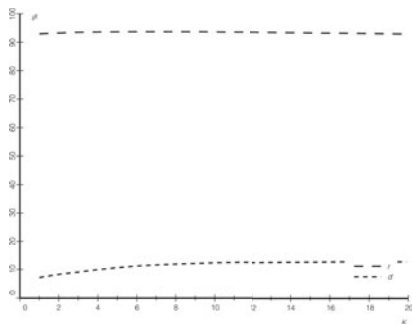
Graph A. Small: No Return Predictability
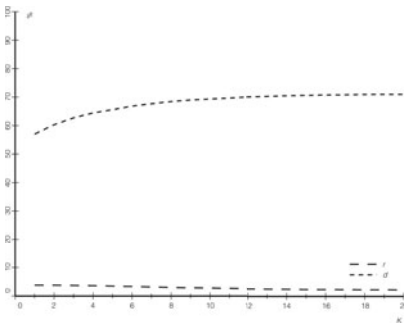
Graph B. Value: No Return Predictability
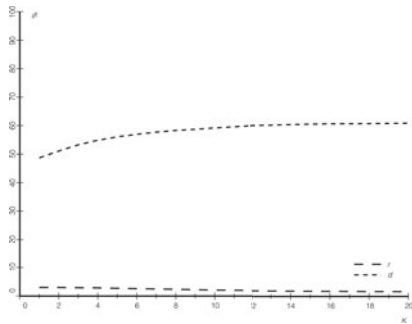
Graph C. Small: No Dividend Predictability

Graph D. Value: No Dividend Predictability

Under the null of no dividend growth predictability, the fractions of replications in which the return slopes are statistically significant are above 90% for both portfolios, thus showing high statistical power associated with the *t*-ratios for the return coefficients. On the other hand, the percentages of pseudosamples in which the dividend growth slopes are significant are around 10% at most horizons, for both portfolios. This means that the asymptotic *t*-statistics reject more often the null hypothesis of no dividend growth predictability than they do the null of no return predictability. Overall, these results show that the asymptotic *t*-statistics associated with the VAR-based predictive slopes exhibit reasonable size and power in the case of small and value portfolios. In other words, we tend not to observe dividend growth/return predictability when there is none; on the other hand, when the data-generating process contains predictability from the dividend yield, this is frequently detected by the asymptotic *t*-ratios.

By using the same two Monte Carlo simulations presented above, we also compute alternative *p*-values to gauge the statistical significance of the VAR-based return and dividend growth coefficients at multiple horizons. These *p*-values represent an alternative to the asymptotic *p*-values computed in Section IV. As in the bootstrap simulation in the last section, we compute the fractions of replications under which the return (dividend growth) coefficients are higher (lower) than the respective estimates found in the data.

Results presented in the Online Appendix show that under the null of no return predictability, the *p*-values for the return slopes are well above 10% at all horizons, for both small and value portfolios. This indicates that according to the marginal distribution, the return coefficients are not statistically significant for these two portfolios. On the other hand, under the null of no dividend growth predictability, for both portfolios, it turns out that the *p*-values associated with the dividend growth slopes are clearly below 10%, and actually lower than 5% at most horizons. In other words, we reject the absence of dividend growth predictability for both small and value stocks.

VII.   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 BM. Our results show that what explains time variation in the dividend-to-price ratio of small stocks is the 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 long-run return and dividend growth predictability drive the variation in the dividend-to-price ratio.

In sum, 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. These 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 related literature. Our results also remain reasonably robust when we conduct several alternative tests such as

computing a bootstrap-based inference, estimating the variance decomposition for the postwar period, and estimating an alternative decomposition based on excess returns and interest rates.

References

Ang, A. "Predicting Dividends in Log-Linear Present Value Models." *Pacific-Basin Finance Journal*, 20 (2012), 151–171.

Ang, A., and G. Bekaert. "Stock Return Predictability: Is It There?" *Review of Financial Studies*, 20 (2007), 651–708.

Boudoukh, J.; M. Richardson; and R. Whitelaw. "The Myth of Long-Horizon Predictability." *Review of Financial Studies*, 21 (2008), 1577–1605.

Brandt, M., and P. Santa-Clara. "Dynamic Portfolio Selection by Augmenting the Asset Space." *Journal of Finance*, 61 (2006), 2187–2217.

Campbell, J. "Understanding Risk and Return." *Journal of Political Economy*, 104 (1996), 298–345.

Campbell, J.; Y. Chan; and L. Viceira. "A Multivariate Model of Strategic Asset Allocation." *Journal of Financial Economics*, 67 (2003), 41–80.

Campbell, J.; C. Polk; and T. Vuolteenaho. "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns." *Review of Financial Studies*, 23 (2010), 305–344.

Campbell, J., and R. Shiller. "The Dividend Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1 (1988a), 195–228.

Campbell, J., and R. Shiller. "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance*, 43 (1988b), 661–676.

Campbell, J., and S. Thompson. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, 21 (2008), 1509–1531.

Campbell, J., and L. Viceira. "Consumption and Portfolio Decisions When Expected Returns Are Time Varying." *Quarterly Journal of Economics*, 114 (1999), 434–495.

Campbell, J., and T. Vuolteenaho. "Bad Beta, Good Beta." *American Economic Review*, 94 (2004), 1249–1275.

Campbell, J., and M. Yogo. "Efficient Tests of Stock Return Predictability." *Journal of Financial Economics*, 81 (2006), 27–60.

Chen, L. "On the Reversal of Return and Dividend Growth Predictability: A Tale of Two Periods." *Journal of Financial Economics*, 92 (2009), 128–151.

Chen, L.; Z. Da; and R. Priestley. "Dividend Smoothing and Predictability." *Management Science*, 58 (2012), 1834–1853.

Cochrane, J. "Explaining the Variance of Price-Dividend Ratios." *Review of Financial Studies*, 5 (1992), 243–280.

Cochrane, J. "The Dog That Did Not Bark: A Defense of Return Predictability." *Review of Financial Studies*, 21 (2008), 1533–1575.

Cochrane, J. "Presidential Address: Discount Rates." *Journal of Finance*, 66 (2011), 1047–1108.

Cohen, R.; C. Polk; and T. Vuolteenaho. "The Value Spread." *Journal of Finance*, 58 (2003), 609–641.

Cornell, B. "Risk, Duration, and Capital Budgeting: New Evidence on Some Old Questions." *Journal of Business*, 72 (1999), 183–200.

Engsted, T., and T. Pedersen. "The Dividend-Price Ratio Does Predict Dividend Growth: International Evidence." *Journal of Empirical Finance*, 17 (2010), 585–605.

Engsted, T.; T. Pedersen; and C. Tanggaard. "The Log-Linear Return Approximation, Bubbles, and Predictability." *Journal of Financial and Quantitative Analysis*, 47 (2012), 643–665.

Fama, E., and K. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 (1988), 3–25.

Fama, E., and K. French. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics*, 25 (1989), 23–49.

Ferson, W., and C. Harvey. "Conditioning Variables and the Cross Section of Stock Returns." *Journal of Finance*, 54 (1999), 1325–1360.

Goetzmann, W., and P. Jorion. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, 48 (1993), 663–679.

Goyal, A., and I. Welch. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies*, 21 (2008), 1455–1508.

Harvey, C. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *Journal of Financial Economics*, 24 (1989), 289–317.

Hodrick, R. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 5 (1992), 357–386.

Koijen, R., and S. Van Nieuwerburgh. "Predictability of Returns and Cash Flows." *Annual Review of Financial Economics*, 3 (2011), 467–491.

Kothari, S., and J. Shanken. "Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis." *Journal of Financial Economics*, 44 (1997), 169–203.

Kelly, B., and S. Pruitt. "Market Expectations in the Cross Section of Present Values." *Journal of Finance*, 68 (2013), 1721–1756.

Lacerda, F., and P. Santa-Clara. "Forecasting Dividend Growth to Better Predict Returns." Working Paper, Nova School of Business and Economics (2010).

Larrain, B., and M. Yogo. "Does Firm Value Move Too Much To Be Justified by Subsequent Changes in Cash Flow?" *Journal of Financial Economics*, 87 (2008), 200–226.

Lettau, M., and S. Van Nieuwerburgh. "Reconciling the Return Predictability Evidence." *Review of Financial Studies*, 21 (2008), 1607–1652.

Lettau, M., and J. Wachter. "Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium." *Journal of Finance*, 62 (2007), 55–92.

Lewellen, J. "Predicting Returns with Financial Ratios." *Journal of Financial Economics*, 74 (2004), 209–235.

Maio, P. "Intertemporal CAPM with Conditioning Variables." *Management Science*, 59 (2013a), 122–141.

Maio, P. "The 'Fed Model' and the Predictability of Stock Returns." *Review of Finance*, 17 (2013b), 1489–1533.

Maio, P., and P. Santa-Clara. "Multifactor Models and Their Consistency with the ICAPM." *Journal of Financial Economics*, 106 (2012), 586–613.

Maio, P., and D. Xu. "Cash-Flow or Return Predictability at Long Horizons? The Case of Earnings Yield." Working Paper, Hanken School of Economics (2014).

Newey, W., and K. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.

Petkova, R. "Do the Fama–French Factors Proxy for Innovations in Predictive Variables?" *Journal of Finance*, 61 (2006), 581–612.

Petkova, R., and L. Zhang. "Is Value Riskier than Growth?" *Journal of Financial Economics*, 78 (2005), 187–202.

Pontiff, J., and L. Schall. "Book-to-Market Ratios as Predictors of Market Returns." *Journal of Financial Economics*, 49 (1998), 141–160.

Rangvid, J.; M. Schmeling; and A. Schrimpf. "Dividend Predictability Around the World." *Journal of Financial and Quantitative Analysis*, 49 (2014), 1255–1277.

Rytchkov, O. "Expected Returns on Value, Growth, and HML." *Journal of Empirical Finance*, 17 (2010), 552–565.

Stambaugh, R. "Predictive Regressions." *Journal of Financial Economics*, 54 (1999), 375–421.

Van Binsbergen, J., and R. Koijen. "Predictive Regressions: A Present-Value Approach." *Journal of Finance*, 65 (2010), 1439–1471.