The long-run relation between returns, earnings, and dividends

Paulo Maio

First version: May 2012
This version: December 2012

1Hanken School of Economics. E-mail: paulofmaio@gmail.com.
2I thank Pedro Santa-Clara and seminar participants at Luxembourg School of Finance for helpful comments. I am grateful to Amit Goyal for providing data on his website. All errors are mine.
Abstract

This paper focuses on the predictive ability of the aggregate earnings yield for market returns and earnings growth by imposing the restrictions associated with a present-value relation. By estimating a variance decomposition for the earnings yield based on weighted long-horizon regressions for the 1872–1925 and 1926–2010 periods, I find a reversal in return/earnings growth predictability: in the earlier period, the bulk of variation in the earnings yield is predictability of earnings growth, while in the modern sample the driving force is return predictability. When the variance decomposition is based on a first-order VAR the results in the modern sample are qualitatively different, i.e., the restrictions imposed by the first-order VAR are not validated by the data. Therefore, in the post-1926 period what drives aggregate financial ratios are expectations about future discount rates rather than future cash flows, irrespective of the financial ratio (dividend yield or earnings yield) being used.

Keywords: asset pricing; predictability of stock returns; earnings-growth predictability; long-horizon regressions; earnings yield; VAR implied predictability; present-value model; dividend payout ratio

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

Using financial ratios to forecast future stock market returns has been a common practice in the empirical finance literature. However, the focus has been on the predictive ability of the dividend-to-price ratio while the earnings-to-price ratio has assumed a secondary role.\(^1\) Moreover, dividend smoothing or the changing payout policy in recent years by U.S. firms (favoring share repurchases over dividends) (see Fama and French (2001), Brav, Graham, Harvey, and Michaely (2005), Leary and Michaely (2011), among others) might make the earnings yield more informative about future equity returns or cash flows than the dividend yield.

This paper focuses on the predictive ability of the aggregate earnings yield for market returns and earnings growth by imposing the restrictions associated with a present-value relation.\(^2\) Under this present-value decomposition it follows that the log earnings yield is positively correlated with future returns and negatively correlated with both future log earnings growth and log dividend payout ratios. I define a term structure of variance decompositions for the earnings-to-price ratio, similarly to the analysis conducted for the dividend-to-price in Cochrane (2008, 2011) or the decomposition for the book-to-market ratio in Cohen, Polk, and Vuolteenaho (2003): at each forecasting horizon (from one to 20 years ahead) the variation in the earnings yield is the result of four types of predictability from this financial ratio: predictability of future returns, earnings growth predictability, predictability of future payout ratios, or the predictability of the earnings yield at some future date.

By estimating weighted long-horizon regressions for the 1872–1925 and 1926–2010 periods I find a reversal in return/earnings growth predictability from the earnings yield: in the earlier period, the bulk of variation in the earnings yield is predictability of earnings growth, while in the modern sample what drives the earnings-to-price ratio is return predictability. These results are consistent with the findings in Cochrane (2008, 2011) and Chen (2009), based on a VAR-based variance decomposition for the market dividend yield, which show that most


of the variation in the dividend yield in the earlier period is attributable to dividend growth predictability, while in the later period the main driver is return predictability.

Furthermore, the finding in this paper is opposite to the result in Chen, Da, and Priestley (2012) that the key driver of the market earnings yield in the modern period is earnings growth, rather than return, predictability. I show that their results only hold when the variance decomposition for the earnings yield is based on a first-order VAR. Thus, this paper makes a contribution to the debate on using long-horizon regressions to estimate predictive coefficients at long horizons versus the alternative approach of obtaining implied estimates from a short-order VAR, which has been widely used in the related literature. I show that, in the 1926–2010 period, the two approaches yield similar results when the predicting variable is the dividend yield, but quite opposing results when the forecasting variable is the earnings yield. In other words, the restrictions imposed by the first-order VAR are not validated by the data when the predictor is the earnings yield.

Therefore, the results in this paper are inconsistent with a dividend smoothing argument behind the changing predictability pattern of dividend yield over the two periods, as argued by Chen, Da, and Priestley (2012). Instead, I argue that this reversal in predictability associated with both the earnings and dividend yield is a consequence of the changing characteristics and risk-return profiles of the average firm in the U.S. stock market over time. One possibility is that in the 1926–2010 period (and especially in the post-war period) the average stock in the market is tilted towards younger firms; growth stocks with higher investment opportunities; stocks with longer duration of cash flows; stocks with higher dependence on external equity finance; and stocks with lower profitability than in the earlier period. Consequently, it is likely that the price (valuation) of the average stock over the 1926–2010 period (and particularly, in the post-war period) is more sensitive to shocks in the respective discount rates than to cash flows.

Furthermore, it is likely that the average stock faces higher limits to arbitrage and higher transaction costs during the earlier period, making their valuation less sensitive to shocks in discount rates and thus more responsive to shocks in cash flows. Hence, this paper provides additional evidence that the changing characteristics of the typical U.S. firm might explain the changing pattern of return/cash flow predictability in the U.S. stock market over the long sample, 1872–2010.

I conduct a Monte-Carlo simulation by imposing the null of no return and no earnings
growth predictability, that is, under this null all the variation of the earnings yield comes from predicting the payout ratio or the future earnings yield. The Monte-Carlo p-values confirm the asymptotic t-statistics associated with the predictive slopes from the long-horizon regressions: in the earlier period, one rejects the null of no earnings growth predictability, while one cannot reject the null of no return predictability. In the modern sample, we have exactly opposite results.

I also derive and estimate a variance decomposition for the earnings yield in terms of excess returns rather than returns. Under this decomposition, part of the variation in the earnings yield is the result of positive predictability from the earnings yield for future excess returns and short-term interest rates. The results are qualitatively similar to the benchmark variance decomposition based on returns since the market earnings yield has little forecasting power for future interest rates.

Overall, the results in this paper show that in the post-1926 period what drives aggregate financial ratios are expectations about future discount rates rather than future cash flows, irrespective of the financial ratio (dividend yield or earnings yield) being used.

The paper is organized as follows. In Section 2, I describe the data and variables. Section 3 presents the benchmark variance decomposition for the earnings yield, based on long-horizon regressions. In Section 4, I present an alternative variance decomposition based on a first-order VAR. Section 5 presents the results from a Monte-Carlo simulation. In Section 6, I analyze the predictability of the earnings yield for excess stock returns. Section 7 concludes.

2 Data and variables

I use annual data on earnings \((E)\), dividends \((D)\), and price level \((P)\) associated with the Standard and Poors (S&P) 500 index. The data are available from Amit Goyal’s webpage.\(^3\) The sample is 1872–2010. The gross annual return is computed as \(\frac{P_{t+1} + D_{t+1}}{P_t}\). The descriptive statistics for the log return \((r)\), log dividend-to-price ratio \((d – p)\), log earnings-to-price ratio \((e – p)\), log dividend payout ratio \((d – e)\), log dividend growth \((\Delta d)\), and log earnings growth \((\Delta e)\) are presented in Table 1.

We can see that the dividend-to-price ratio is much more persistent than the earnings-to-price ratio with an autocorrelation coefficient of 0.88 versus 0.71. The dividend yield is also

\(^3\)For a description of the data, see Goyal and Welch (2008).
slightly more volatile than the earnings yield. On the other hand, the log payout ratio is less persistent than both \( d - p \) and \( e - p \) with an autoregressive coefficient of 0.59, and is also slightly less volatile than the two price multiples. Both dividend growth and earnings growth are not persistent variables as indicated by the autocorrelation coefficients of 0.25 and -0.07, respectively. Earnings growth is significantly more volatile than dividend growth, which should be related to dividend smoothing at the aggregate level.

The correlations between the variables in Panel B show that \( d - p \) and \( e - p \) are positively correlated, but this correlation is far from perfect (0.70). This can be confirmed in Figure 1, which plots the time-series for both price ratios. We can see that both the earnings yield and dividend yield assume higher values in the first half of the sample and that there is a downward trend in both variables over the post-war period. The log earnings growth and log dividend growth are only weakly correlated (0.31), which again should be related to dividend smoothing. Figure 2, Panel A shows periods of high volatility in earnings growth, the great depression, and most notably, in the later recession (2007–2009). In turn, this implies that the dividend payout ratio was especially volatile in the 20–30s period and also in the late 2000’s, as shown in Panel B of Figure 2. From Table 1, we can also see that the log payout ratio is positively correlated with \( d - p \) (0.51) and negatively correlated with \( \Delta e \) (-0.61).

3 Predictability from the earnings yield: long-horizon regressions

3.1 Methodology

In this section, I analyze the predictability of the earnings yield for future returns, earnings growth, and dividend payout ratios. I conduct a variance decomposition for the earnings yield based on direct weighted long-horizon regressions as in Cochrane (2008, 2011). This approach differs from the traditional approach of constructing long-horizon predictive coefficients based on the estimates from a first-order VAR (e.g., Cochrane (2008), Lettau and Van Nieuwerburgh (2008), Chen (2009), Binsbergen and Kojien (2010), among others). The slope estimates from the long-horizon regressions might be different than the implied VAR slopes if the correlation between the log earnings-to-price ratio and future multi-period returns or earnings growth is not fully captured by a first-order VAR. This might occur, for example, due to a gradual reaction of returns/earnings growth to shocks in the current earnings yield. Thus, the long-
horizon regressions provide, in principle, more correct estimates of the long-horizon predictive coefficients in the sense that they do not depend on the restrictive assumptions imposed by the short-run VAR. In the next section, I present the long-horizon estimates implied from a first-order VAR.

I estimate weighted long-horizon regressions of future log returns ($r$), log earnings growth ($\Delta e$), log payout ratio ($d-e$), and log earnings-to-price ratio ($e-p$) on the current earnings-to-price ratio:

\[
\sum_{j=1}^{K} \rho^{j-1} r_{t+j} = a_r^K + b_r^K (e_t - p_t) + \varepsilon_{r,t+K},
\]

\[
\sum_{j=1}^{K} \rho^{j-1} \Delta e_{t+j} = a_e^K + b_e^K (e_t - p_t) + \varepsilon_e^{de,t+K},
\]

\[
(1-\rho) \sum_{j=1}^{K} \rho^{j-1} (d_{t+j} - e_{t+j}) = a_{de}^K + b_{de}^K (e_t - p_t) + \varepsilon_{de,t+K},
\]

\[
\rho^K (e_{t+K} - p_{t+K}) = a_{ep}^K + b_{ep}^K (e_t - p_t) + \varepsilon_{ep,t+K},
\]

where $\rho$ is a (log-linearization) discount coefficient that depends on the mean of $d-p$; the $\varepsilon$s denote forecasting errors; and $K$ is the forecasting horizon. The $t$-statistics associated with the long-horizon predictive slopes are based on Newey and West (1987) standard errors computed with $K$ lags.\(^{4}\)

Following Chen, Da, and Priestley (2012) and Maio (2012b), the dynamic accounting identity for $e-p$ can be represented as

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

where $c$ is a log-linearization constant that is irrelevant for the forthcoming analysis. This decomposition generalizes the Campbell and Shiller (1988a) present-value relation and postulates that the log earnings yield is positively correlated with future returns and the earnings yield at a future date, and negatively correlated with both future log earnings growth and log payout ratios.

By combining the present-value relation in (5) with the predictive regressions above, one obtains an identity involving the predictability coefficients associated with $e-p$, at every horizon

\(^{4}\)Sadka (2007) also analyzes the predictability of earnings growth at multiple horizons, but using the dividend yield, rather than the earnings yield, as the predictor.
which can be interpreted as a variance decomposition for the log earnings yield. The predictive coefficients $b^K_r$, $b^K_{de}$, $b^K_e$, and $b^K_{ep}$ represent the fraction of the variance of current $e - p$ attributable to return, payout ratio, earnings growth, and earnings yield predictability, respectively. Thus, these estimates represent a measure of the relative importance of each variable in driving the variation in the earnings-to-price ratio.

### 3.2 Predictability from dividend yield

I first present the empirical results for the variance decomposition associated with the dividend yield, which serve as a reference point for the main results associated with the earnings yield. Following Cochrane (2008, 2011), the variance decomposition for the log dividend-to-price ratio $(d - p)$ is given by

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

(7)

where $b^K_r$, $b^K_d$, and $b^K_{dp}$ represent the fraction of the variance of current $d - p$ attributable to return, dividend growth, and dividend yield predictability at horizon $K$, respectively. These predictive slopes are obtained from the following weighted long-horizon regressions:

\[ \sum_{j=1}^{K} \rho^{j-1} r_{t+j} = a^K_r + b^K_r (d_t - p_t) + \varepsilon^r_{t+K}, \]

(8)

\[ \sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} = a^K_d + b^K_d (d_t - p_t) + \varepsilon^d_{t+K}, \]

(9)

\[ \rho^K (d_{t+K} - p_{t+K}) = a^K_{dp} + b^K_{dp} (d_t - p_t) + \varepsilon^{dp}_{t+K}. \]

(10)

The term structure of the predictive slopes associated with $d - p$ are presented in Figure 3, while Figure 4 plots the $t$-statistics. In the full sample (1872–2010) the main driver of variation in $d - p$ is dividend growth predictability with a slope close to -70% at $K = 20$. The estimates for $b^K_d$ are statistically significant at all horizons. In comparison, the return slopes do not exceed 30% and are not significant at the 5% level for horizons above 12 years. On the other hand, for very long horizons the share associated with future $d - p$ predictability does not exceed 10%, and the respective estimates are not significant for horizons beyond 12 years.

In the earlier sample (1872–1925), the dividend growth coefficients at long horizons ($K > 12$)
represent more than 100% of the variance of $d - p$, and the respective slopes are statistically significant at all horizons. The reason for these estimates of $b^K_d$, below -100%, is that both $b^K_r$ and $b^K_{dp}$ have the wrong signs for horizons greater than 18 and 11 years, respectively, although most of these estimates are not statistically significant at the 5% level.

For the modern sample (1926–2010) we have quite different results than those observed for the earlier period: the return coefficients achieve values around 80% at long horizons, and the return slopes are significant for all horizons beyond 2 years. On the other hand, the maximum share of dividend growth predictability over the $d - p$ variance is around 20% at $K = 20$, and the estimates for $b^K_d$ are statistically significant only for $K ≥ 19$. The coefficients associated with the future dividend yield have the right sign, but converge to zero faster than in the full sample, with estimates below 10% for horizons greater than 15 years.

In the most recent period (1946–2010) the results show an even stronger return predictability than for the 1926–2010 sample: the estimates for $b^K_r$ achieve values above 100% for horizons beyond 12 years, and the return slopes are statistically significant at all horizons. On the other hand, the dividend growth coefficients have the wrong sign at most forecasting horizons, with the exception of $K ≥ 18$. These estimates are not significant at the 5% level for most horizons (the exception is for horizons beyond 10 and 12 years ahead). Moreover, the slopes for future $d - p$ converge to zero faster than in the 1926–2010 period, and actually turn negative at $K = 14$, although they are not significant.

Overall, these results are consistent with previous work based on a first-order VAR (Cochrane (2008, 2011) and Chen (2009)) that the bulk of the variation of the market dividend yield in the 1926–2010 period is return predictability, although in the earlier period (1872–1925) the main driver of the dividend yield is dividend growth predictability.

### 3.3 Predictability from earnings yield

I next present the benchmark results for the variance decomposition associated with the earnings yield. The term structure of the predictive slopes and associated $t$-statistics are presented in Figures 5 and 6, respectively.

In the long sample, at long horizons (around $K = 20$) the shares associated with return predictability (around 60%) and earnings growth predictability (around 70%) are relatively similar. Moreover, both the return and earnings growth slopes are significant at the 5% level for all horizons. The sum of the magnitudes of $b^K_r$ and $b^K_e$ exceed 100% at long horizons.
because the coefficients associated with future $e - p$ are negative for horizons beyond 11 years (at $K = 20$, we have $b^K_{ep}$ around -20%), and these estimates are significant for $K > 17$. The coefficients associated with $d - e$ have also the wrong sign, but the magnitudes are quite small, showing that the payout ratio has a residual importance.

In the earlier sample (1872–1925) we have stronger earnings predictability: the earnings growth coefficients are around -150% at long horizons, and the estimates are statistically significant at all horizons. On the other hand, the return coefficients are negative for horizons greater than ten years, although most of these estimates are not statistically significant (the exceptions are for $K=19, 20$). The coefficients associated with the future earnings yield also have the wrong sign for horizons beyond 11 years, but most of these estimates are not significant. The slopes for the payout ratio are positive at most horizons (around 10% at $K = 20$), and there is statistical significance for $K > 13$.

In the modern sample (1926–2010) we have a complete different picture than in the earlier period: the return coefficients achieve values very close to 100% for horizons beyond 16 years, and there is statistical significance at all horizons. In contrast, the maximum share associated with earnings growth predictability is around 40% at the 20-year horizon. The dividend growth slopes are significant only for horizons beyond 11 years. As in the full sample, the slopes for future earnings yield turn negative at long horizons (around -30% at $K = 20$), and these negative estimates are significant for horizons beyond 17 years. Thus, contrary to the early sample, in the modern period the bulk of variation in the earnings yield stems from return predictability rather than earnings growth predictability. These results are also in line with the decomposition for the dividend yield discussed above: in the earlier period we have more cash flow (dividends or earnings) predictability while in the later period we have more return predictability.\footnote{By conducting the same analysis for the 1926–2006 period (which excludes the recent years of very large volatility in earnings growth), I obtain a similar variance decomposition for the earnings yield.}

In the most recent sample (1946–2010), the share of return predictability is even greater than for the 1926–2010 period, with predictive slopes above 100% at long horizons. Both $b^K_e$ and $b^K_{ep}$ are about -40% at $K = 20$, and both estimates are statistically significant. The small positive estimates for the payout ratio slopes are also significant for horizons beyond 11 years. These results reinforce the finding that the bulk of variation in $e - p$ in the modern sample is return predictability rather than earnings growth predictability. We can see that for all four samples, the levels of the variance decomposition are very close to 100%, meaning that the
identity in (6) is accurate.

3.4 Discussion

Why is the main driver of variation in the market earnings yield in the pre–1926 period earnings growth predictability while in the post-1926 period it is all about return predictability? Since this reversal occurs also for the predictability from the aggregate dividend yield, as shown above and also in Chen (2009), it seems implausible that this pattern is associated with dividend smoothing, as argued by Chen, Da, and Priestley (2012). Instead, I claim that this reversal in predictability associated with both the earnings and dividend yield is a consequence of the changing characteristics and risk-return profiles of the average firm in the U.S. stock market over time. One possibility is that in the 1926–2010 period (and especially in the post-war period) the average stock in the market is tilted towards younger firms; growth stocks with higher investment opportunities; stocks with longer duration of cash flows; stocks with higher dependence on external equity finance; and stocks with lower profitability than in the earlier period. This trend is likely to be associated with the IPO wave in the 1960s, the emergence of the NASDAQ in the 1970s, the boom in technology stocks in the 1990s, and the decrease in listing requirements over the modern period. Consequently, it is likely that the price (valuation) of the average stock over the 1926–2010 period (and particularly, in the post-war period) is more sensitive to shocks in the respective discount rates than to shocks in cash flows [see Cornell (1999), Campbell and Vuolteenaho (2004), among others]. Moreover, it is likely that the average stock faces higher limits to arbitrage and higher transaction costs during the earlier period, making their valuation less sensitive to shocks in discount rates and thus more responsive to shocks in cash flows [see D’Avolio (2002) and Mitchell, Pulvino, and Stafford (2002), among others].

4 Predictability from the earnings yield: a VAR approach

4.1 Predictability from earnings yield

In this section, I conduct an alternative variance decomposition for the earnings yields based on a first-order VAR, as in Chen, Da, and Priestley (2012). The term-structure of predictive
slopes are based on the following first-order restricted VAR:

\[ r_{t+1}^e = a_r + b_r(e_t - p_t) + \varepsilon_{t+1}^e, \] (11)
\[ \Delta e_{t+1} = a_e + b_e(e_t - p_t) + \varepsilon_{t+1}^e, \] (12)
\[ d_{t+1} - e_{t+1} = a_{de} + b_{de}(e_t - p_t) + \varepsilon_{t+1}^{de}, \] (13)
\[ e_{t+1} - p_{t+1} = a_{ep} + \phi(e_t - p_t) + \varepsilon_{t+1}^{ep}. \] (14)

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 present-value relation in (5), one obtains the VAR-based variance decomposition for \( e - p \), at horizon \( K \):

\[ 1 = b^K_r - b^K_{de} - b^K_e + b^K_{ep}, \] (15)
\[ b^K_r \equiv b_r \left( 1 - \rho^K \phi^K \right) \frac{1}{1 - \rho \phi}, \]
\[ b^K_{de} \equiv (1 - \rho) b_{de} \left( 1 - \rho^K \phi^K \right) \frac{1}{1 - \rho \phi}, \]
\[ b^K_e \equiv b_e \left( 1 - \rho^K \phi^K \right), \]
\[ b^K_{ep} \equiv \rho^K \phi^K. \]

The \( t \)-statistics associated with the predictive coefficients in (15) are computed based on the \( t \)-statistics associated with the VAR slopes above by using the Delta method. Details of these standard errors are presented in the appendix.

I also compute the implied slopes from the one-period variance decomposition:

\[ 1 = b_r - b_{de} - b_e + \rho \phi, \] (16)

and the respective \( t \)-statistics, which are based on the standard errors of the VAR estimates by employing the Delta method. The implied estimates provide an idea of the accuracy of the log-linear approximation, underlying the present-value relation, at the one-year horizon.

The VAR estimation results are presented in Table 2. In the full sample, the earnings yield has relatively weak forecasting power for the market return with a coefficient of determination around 2%. The return slope is 0.07 and this estimate is significant at the 10\% level. The earnings growth predictability from \( e - p \) is much stronger as shown by the \( R^2 \) estimate close to 11%. The earnings growth coefficient is -0.25, which is strongly significant (1\% level). The
earnings yield has no forecasting ability for the dividend payout ratio as shown by the very small forecasting ratio (0.4%) and a largely insignificant slope. We can also see that the implied predictive coefficients, and respective t-statistics, are quite similar to the original estimates, which provides evidence that the log-linear approximation works relatively well.

By comparing the two subsamples (Panels B and C) we can see that there is significantly greater return predictability in the modern sample as indicated by the $R^2$ estimate of 4.7% compared to an estimate around 0% in the earlier sample. The estimate for the return slope in the recent sample (0.10) is significant at the 5% level, compared to a negative estimate, albeit largely insignificant, in the first period. In contrast, earnings growth predictability is much stronger in the first period: the forecasting ratio is as large as 22%, compared to 6.9% in the modern sample. The earnings growth coefficient in the first period is -0.47 and this estimate is significant at the 1% level, while the slope for the second period is significant only at the 10% level. The earnings yield is more persistent in the recent sample with an autoregressive slope of 0.74 (compared to 0.56 for the first sample) and a forecasting ratio in the AR(1) equation of 54% (versus 31%). As in the full sample, $e - p$ does not forecast the payout ratio in any of the subsamples. Thus, these results show that in the modern sample there is more one-year return predictability while in the earlier sample there is more earnings predictability. Moreover, the financial ratio is more persistent in the later sample.

For the most recent sample (1946–2010), we observe the greatest amount of return predictability among all the periods with a forecasting ratio close to 8%. The corresponding predictive slope (0.10) is significant at the 5% level. Furthermore, the evidence of earnings growth predictability is relatively weak as indicated by the $R^2$ estimate of 4.3% and the highly insignificant slope. The earnings-to-price ratio exhibits the larger persistence during the most recent period with an autoregressive slope of 0.78 and a forecasting ratio in the AR(1) equation of 61%.

Next, I discuss the VAR-based predictability at several forecasting horizons. The term structure of predictive coefficients is displayed in Figure 7. In the full sample, the main driver of variation in $e - p$ is future earnings growth with coefficients around -80% at long horizons compared to return coefficients around only 20%. The term structure of t-statistics in Figure 8 shows that the earnings growth coefficients are significant at all horizons while there is no statistical significance (at the 5% level) for the return coefficients. Given the relatively small persistence of the earnings yield it follows that the slopes associated with future $e - p$ converge
relatively fast to zero, achieving values below 10% at the five-year horizon. The respective $t$-statistics point to non-significance for $K > 5$. The long-run implied return and dividend growth slopes (for $K \to \infty$) are 0.21 and -0.79, respectively, which basically correspond to the estimates associated with $K = 20$.

By comparing the two subsamples (Panels B and C) we can see that earnings growth predictability drives almost all the variation of $e - p$ in the first period (slopes around -100%) while the magnitudes of $b^K$ are somewhat smaller in the modern sample (around -70%). In both periods the earnings growth coefficients are statistically significant at all horizons. In contrast, there is no return predictability in the earlier sample (coefficients around zero), while in the modern sample it turns out that return predictability accounts for around 30% of the variance of the earnings yield. However, the return slopes are statistically significant (at the 5% level) only at the near term ($K < 8$).

In the most recent sample, we observe the highest return predictability in the long-run (around 40%) while the earnings growth slopes are about -60% at very long horizons. However, the return slopes are not significant for horizons beyond five years.

4.2 Discussion

By comparing the VAR-based variance decomposition with the benchmark decomposition based on the long-horizon regressions from the previous section we observe a dramatic change in the predictability pattern for the modern samples (1926–2010 and 1946–2010). While in the benchmark case the bulk of variation in $e - p$ is return predictability at long horizons, in the case of the VAR-based decomposition the main driver of the earnings yield is earnings growth predictability. Thus, it seems that the first-order VAR misses relevant information concerning the correlation between the current earnings-to-price ratio and both future returns and earnings growth at longer horizons, for the modern period. These results also show that the evidence in Chen, Da, and Priestley (2012) that earnings growth predictability is the main driver of the earnings yield over this period only holds under the restrictive first-order VAR. In other words, when one uses the less restrictive long-horizon regressions it follows that return predictability is the main driver for both the earnings yield and dividend yield, rather than cash flow predictability, in the modern sample.\(^6\)

To see better why the (weighted) long-horizon regressions and first-order VAR yield different

---

\(^6\)Cochrane (1988) represents another example in which a low-order VAR does not capture long-run dynamics.
predictive coefficients at horizons beyond one-year, consider the simplest case of $K = 2$. In this case, the return coefficient from the long-horizon regression is given by

$$\beta(r_{t+1} + \rho r_{t+2}, e_t - p_t) = \beta(r_{t+1}, e_t - p_t) + \rho \beta(r_{t+2}, e_t - p_t) = b_r + \rho \beta(r_{t+2}, e_t - p_t),$$

(17)

where $b_r$ is the one-year return coefficient, which is the same under the two methodologies. The implied VAR return coefficient is equal to

$$b_r \frac{1 - \rho^2 \phi^2}{1 - \rho \phi} = b_r + \rho b_r \phi.$$  

(18)

In general, the regression coefficient $\beta(r_{t+2}, e_t - p_t)$ will be different than $b_r \phi$, thus explaining the difference in results from the two methodologies. This difference is likely to be more severe at longer horizons because the gap above potentially spreads to several periods instead of occurring only in the second period, as in this simple example. Specifically, in the 1926–2010 sample (for which there is a greater discrepancy between the VAR-based and direct estimates) the VAR-based estimates for the return slopes are 0.29, 0.34, and 0.35 at $K = 5$, $K = 10$, and $K = 20$, respectively, while the corresponding estimates from the long-horizon regressions are 0.41, 0.80, and 0.93, respectively. The VAR implied earnings growth coefficients are -0.53, -0.63, and -0.65 at $K = 5$, $K = 10$, and $K = 20$, respectively, and the corresponding direct estimates are -0.21, -0.05, and -0.38, respectively. Thus, the differences in the magnitudes of the estimates between the two methods tends to increase with horizon, especially, in the case of the return slopes. Thus, there is a misspecification in the first-order VAR in capturing long-horizon dynamics, which is magnified at longer horizons.

The punch line of this simple example is that the first-order VAR imposes a restriction on the true regression coefficients at longer horizons, which might not be confirmed by the data in some cases. One reason why the results associated with the earnings yield and dividend yield are qualitatively different in this respect (for the dividend yield the two methodologies yield relatively similar results while for the earnings yield we have the opposite pattern) is that the dividend yield is significantly more persistent (higher $\phi$) than the earnings yield in this sample, as documented in Section 2. Specifically, in the simple example above ($K = 2$), a high value of $\phi$ may imply that the estimate for $\beta(r_{t+2}, e_t - p_t)$ will not be significantly different than $b_r \phi$. 

13
5 Monte-Carlo simulation

A part of the return predictability literature has focused on the poor small-sample properties of long-horizon regressions (see Valkanov (2003), Torous, Valkanov, and Yan (2004), and Boudoukh, Richardson, and Whitelaw (2008), among others). To address this issue, I conduct a Monte-Carlo simulation of the VAR model presented in the last section.

I impose a null hypothesis in which there is no return and no earnings growth predictability, that is, under this null, all the variation of the earnings yield comes from predicting either the payout ratio or the future earnings yield. Thus, I simulate the first-order VAR by imposing the restrictions (in the predictive slopes and residuals) associated with this null hypothesis:

\[
\begin{pmatrix}
    r_{t+1} \\
    d_{t+1} - e_{t+1} \\
    \Delta e_{t+1} \\
    e_{t+1} - p_{t+1}
\end{pmatrix}
= \begin{pmatrix}
    0 \\
    \frac{\rho \phi - 1}{1 - \rho} \\
    0 \\
    \phi
\end{pmatrix}
\begin{pmatrix}
    e_t - p_t
\end{pmatrix}
+ \begin{pmatrix}
    \epsilon^r_{t+1} \\
    \epsilon^{de}_{t+1} \\
    \epsilon^{ep}_{t+1} - (1 - \rho)\epsilon^{de}_{t+1} + \rho\epsilon^{ep}_{t+1} \\
    \epsilon^{ep}_{t+1}
\end{pmatrix}.
\]  

(19)

Following Cochrane (2008), in drawing the VAR residuals (10,000 times), I assume that they are jointly normally distributed and use their covariances from the original sample. The earnings yield for the base period is simulated as \( e_0 - p_0 \sim N(0, \text{Var}(\epsilon^{ep}_{t+1})/(1 - \phi^2)) \). Armed with the artificial data, one computes the empirical \( p \)-values associated with the return and earnings growth slopes from the long-horizon regressions, that is, the fractions of simulated estimates for the return (earnings) coefficients that are higher (lower) than the estimates found in the data.\footnote{To make the plots clear, and since the focus of the analysis is on return/earnings predictability, I do not present the results for the payout ratio and future earnings yield slopes.}

The term-structure of empirical \( p \)-values is displayed in Figure 9. We can see that the simulated \( p \)-values associated with the return slopes are consistent with the corresponding asymptotic \( t \)-statistics in Figure 6: the return slopes are significant (at the 5\% level) in all samples, except in the 1872–1925 period. In contrast, regarding the earnings growth coefficients it turns out that the Monte-Carlo and asymptotic inferences lead to some differences in comparison to the asymptotical inference: in the modern periods (1926–2010 and 1946–2010) the Monte-Carlo simulation indicates that the earnings coefficients are not significant (with the exception of very short horizons in the 1926–2010 period), while the Newey-West \( t \)-statistics indicate significance at longer horizons. On the other hand, in the earlier sample the simulated \( p \)-values point to statistical significance of the earnings slopes, although lower than shown by the \( t \)-statistics (for
horizons between six and 10 years there is only marginal significance, with simulated $p$-values around 10%). Moreover, in the full sample, the earnings growth coefficients are significant only at short or longer horizons based on the simulated $p$-values, while the asymptotic $t$-statistics show significance at all horizons.

I also present the joint distribution of the return and earnings slopes at the one-year and 20-year horizons in Figures 10 and 11, respectively. To make the plots clear, I only present 1,000 out of the 10,000 pseudo samples. Table 3 presents the fractions of Monte-Carlo replications under which the return and earnings growth coefficients are higher or lower than the respective estimates from the original sample.

The results show that at the one-year horizon, in both the full sample and in the earlier period (1872–1925), we cannot reject (at the 5% level) the joint null of no return or no earnings growth predictability, with probabilities of 6% (6%+0%) and 56% (56%+0%), respectively. This is a consequence of not rejecting the null of no return predictability since the null of no earnings predictability is clearly rejected (probabilities of 0%). In contrast, in the most recent period (1946–2010) the joint null is not rejected with a probability of 18% (16%+2%) because the null of no earnings predictability is not rejected (probability of 16%).

When we consider a long forecasting horizon (20 years), we have a different picture in the full sample: the joint null is again not rejected (11%) mainly due to the non-rejection of the null of no earnings predictability (probability of 9%). On the other hand, the results for the earlier period are qualitatively similar to those obtained for the one-period slopes. In the modern periods (1926–2010 and 1946–2010), the non-rejection of the joint null arises because we cannot reject the null of no cash flow predictability (probabilities of 49% and 64%, respectively).

Overall, this Monte Carlo simulation confirms the findings from Section 3: the predictability of earnings growth drives the earnings yield in the earlier sample, while return predictability represents the bulk of variation of the earnings-to-price ratio in the modern period.

6 Does the earnings yield forecasts the equity premium?

In this section, I analyze the predictability of the earnings yield for future excess stock returns rather than nominal returns. The goal is to check whether the main results obtained in Section 3 for the case of returns, also hold for excess returns, which has been the focus of the predictability literature (see, for example, Campbell and Thompson (2008), Goyal and Welch (2008), and Maio
By adding and subtracting the one-period interest rate, $r_{f,t+j}$, it is straightforward to derive the following present-value relation of the earnings yield in terms of future equity premia, $r_{e,t+j}$:

\[
e_t - p_t = \frac{-c(1 - \rho^K)}{1 - \rho} + \sum_{j=1}^{K} \rho^{j-1} r_{e,t+j} - (1 - \rho) \sum_{j=1}^{K} \rho^{j-1}(d_{t+j} - e_{t+j}) - \sum_{j=1}^{K} \rho^{j-1} \Delta e_{t+j}
\]
\[+ \sum_{j=1}^{K} \rho^{j-1} r_{f,t+j} + \rho^K (e_{t+K} - p_{t+K}),
\tag{20}
\]

which states that conditional on the other variables, the current earnings yield is positively correlated with both future excess returns and short-term interest rates.

To analyze the earnings yield predictability for excess stock returns, I conduct the following weighted long-horizon regressions:

\[
\sum_{j=1}^{K} \rho^{j-1} r_{e,t+j} = a^K_r + b^K_r (e_t - p_t) + \varepsilon^K_{t+K},
\tag{21}
\]
\[
\sum_{j=1}^{K} \rho^{j-1} \Delta e_{t+j} = a^K_e + b^K_e (e_t - p_t) + \varepsilon^K_{t+K},
\tag{22}
\]
\[
(1 - \rho) \sum_{j=1}^{K} \rho^{j-1}(d_{t+j} - e_{t+j}) = a^K_{de} + b^K_{de}(e_t - p_t) + \varepsilon^K_{de},
\tag{23}
\]
\[
\sum_{j=1}^{K} \rho^{j-1} r_{f,t+j} = a^K_f + b^K_f (e_t - p_t) + \varepsilon^K_{f},
\tag{24}
\]
\[
\rho^K (e_{t+K} - p_{t+K}) = a^K_{ep} + b^K_{ep}(e_t - p_t) + \varepsilon^K_{ep},
\tag{25}
\]

where the differences relative to the set of regressions conducted in Section 3 are the inclusion of a regression for future interest rates, and the fact that one now forecasts excess returns rather than returns. By combing these predictive regressions with the dynamic present-value relation in (20) one obtains the following variance decomposition for the earnings yield:

\[
1 = b^K_r - b^K_{de} - b^K_e + b^K_f + b^K_{ep}.
\tag{26}
\]

Essentially, this new variance decomposition is similar to the one presented in (15), and decomposes the return slope into an excess return coefficient, $b^K_r$, and an interest rate effect, $b^K_f$. The proxy for the one-period interest rate is the annual Treasury-bill rate, available from Amit Goyal’s webpage.
The term structure of slope estimates, and respective t-statistics, are presented in Figures 12 and 13, respectively. In the full sample the interest rate coefficients are quite small in magnitude and non-significant at all horizons. This implies that the equity premium slopes are very similar to the return coefficients in Figure 5. Moreover, similarly to the return slopes, the estimates for the excess return coefficients are significant at most horizons.

In the earlier period, the interest rate coefficients achieve values around 10% at very long horizons, and these estimates are significant at all horizons. This implies that the equity premium slopes go more in the wrong direction (negative values) than the corresponding return slopes, at longer horizons (these coefficients are significant for horizons greater than 16 years). In the modern sample, the estimates for \( b^K \) reach their maximum (around 10%) for horizons between 8 and 12 years. However, there is no statistical significance for the interest rate slopes at any horizon. At long horizons, the equity premium coefficients are very close to the corresponding return estimates, and are also statistically significant.

For the most recent sample (1946–2010), the interest rate coefficients become negative for horizons greater than 16 years (close to -20% at \( K = 20 \)), although they are not significant. This implies that the amount of excess return predictability is even stronger than the corresponding return predictability, with estimates above 130% at the longer horizons. These estimates are strongly significant, especially for large \( K \).

Therefore, these results show that using a variance decomposition based on excess returns does not change the main result from Section 3: in the earlier period, the bulk of variation in the earnings yield is earnings growth predictability while in the second period it is all about predictability in market discount rates.

7 Conclusion

This paper focuses on the predictive ability of the aggregate earnings yield for market returns and earnings growth by imposing the restrictions associated with a present-value relation. I define a term structure of variance decompositions for the earnings-to-price ratio: at each forecasting horizon (from one to 20 years ahead) the variation in the earnings yield is the result of four types of predictability from this financial ratio: predictability of future returns, earnings growth predictability, predictability of future payout ratios, or the predictability of the earnings yield at some future date.
By estimating weighted long-horizon regressions for the 1872–1925 and 1926–2010 periods I find a reversal in return/earnings growth predictability from the earnings yield: in the earlier period, the bulk of variation in the earnings yield is predictability of earnings growth, while in the modern sample what drives the earnings-to-price ratio is return predictability. These results are consistent with the findings in Cochrane (2008, 2011) and Chen (2009), based on a VAR-based variance decomposition for the market dividend yield, which show that most of the variation in the dividend yield in the earlier period is attributable to dividend growth predictability, while in the later period the main driver is return predictability.

Furthermore, the finding in this paper is opposite to the result in Chen, Da, and Priestley (2012) that the key driver of the market earnings yield in the modern period is earnings growth, rather than return, predictability. I show that their results only hold when the variance decomposition for the earnings yield is based on a first-order VAR. Thus, this paper makes a contribution to the debate on using long-horizon regressions to estimate predictive coefficients at long horizons versus the alternative approach of obtaining implied estimates from a short-order VAR, which has been widely used in the related literature. I show that, in the 1926–2010 period, the two approaches yield similar results when the predicting variable is the dividend yield, but quite opposing results when the forecasting variable is the earnings yield. In other words, the restrictions imposed by the first-order VAR are not validated by the data when the predictor is the earnings yield.

Therefore, the results in this paper are inconsistent with a dividend smoothing argument behind the changing predictability pattern of dividend yield over the two periods, as argued by Chen, Da, and Priestley (2012). Instead, I argue that this reversal in predictability associated with both the earnings and dividend yield is a consequence of the changing characteristics and risk-return profiles of the average firm in the U.S. stock market over time.

I conduct a Monte-Carlo simulation by imposing the null of no return and no earnings growth predictability, that is, under this null all the variation of the earnings yield comes from predicting the payout ratio or the future earnings yield. The Monte-Carlo p-values confirm the asymptotic t-statistics associated with the predictive slopes from the long-horizon regressions: in the earlier period, one rejects the null of no earnings growth predictability, while one cannot reject the null of no return predictability. In the modern sample, we have exactly opposite results.

I derive and estimate a variance decomposition for the earnings yield in terms of excess
returns rather than returns. The results are qualitatively similar to the benchmark variance decomposition based on returns since the market earnings yield has little forecasting power for future interest rates.
References


A Delta method for standard errors of VAR-based coefficients

To compute the asymptotic standard errors of the VAR-based predictive coefficients, \( b^K \equiv (b^K_r, b^K_{de}, b^K_e, b^K_{ep})' \), I use the delta method:

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

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

\[
\frac{\partial b^K}{\partial b'} \equiv \begin{bmatrix}
\frac{\partial b^K_r}{\partial b_r} & \frac{\partial b^K_r}{\partial b_{de}} & \frac{\partial b^K_r}{\partial b_e} & \frac{\partial b^K_r}{\partial \phi} \\
\frac{\partial b^K_{de}}{\partial b_r} & \frac{\partial b^K_{de}}{\partial b_{de}} & \frac{\partial b^K_{de}}{\partial b_e} & \frac{\partial b^K_{de}}{\partial \phi} \\
\frac{\partial b^K_e}{\partial b_r} & \frac{\partial b^K_e}{\partial b_{de}} & \frac{\partial b^K_e}{\partial b_e} & \frac{\partial b^K_e}{\partial \phi} \\
\frac{\partial b^K_{ep}}{\partial b_r} & \frac{\partial b^K_{ep}}{\partial b_{de}} & \frac{\partial b^K_{ep}}{\partial b_e} & \frac{\partial b^K_{ep}}{\partial \phi}
\end{bmatrix}.
\]

(A.2)
Table 1: Descriptive statistics
This table reports descriptive statistics for the log market return \( r \); log dividend-to-price ratio \( d - p \); log earnings-to-price ratio \( e - p \); log dividend payout ratio \( d - e \); log dividend growth \( \Delta d \); and log earnings growth \( \Delta e \). The sample is 1872–2010. \( \phi \) designates the first-order autocorrelation. The correlations between the variables are presented in Panel B.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Stdev.</th>
<th>Min.</th>
<th>Max.</th>
<th>( \phi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r )</td>
<td>0.084</td>
<td>0.179</td>
<td>-0.540</td>
<td>0.425</td>
<td>0.022</td>
</tr>
<tr>
<td>( d - p )</td>
<td>-3.190</td>
<td>0.427</td>
<td>-4.478</td>
<td>-2.288</td>
<td>0.879</td>
</tr>
<tr>
<td>( e - p )</td>
<td>-2.664</td>
<td>0.380</td>
<td>-4.106</td>
<td>-1.670</td>
<td>0.706</td>
</tr>
<tr>
<td>( d - e )</td>
<td>-0.526</td>
<td>0.313</td>
<td>-1.225</td>
<td>0.646</td>
<td>0.591</td>
</tr>
<tr>
<td>( \Delta d )</td>
<td>0.032</td>
<td>0.124</td>
<td>-0.495</td>
<td>0.427</td>
<td>0.252</td>
</tr>
<tr>
<td>( \Delta e )</td>
<td>0.038</td>
<td>0.297</td>
<td>-1.492</td>
<td>1.231</td>
<td>-0.069</td>
</tr>
</tbody>
</table>

Panel B
\[
\begin{array}{cccccc}
 r & d - p & e - p & d - e & \Delta d & \Delta e \\
 r & 1.00 & -0.30 & -0.08 & -0.31 & 0.05 & 0.34 \\
d - p & 1.00 & 0.70 & 0.51 & -0.07 & -0.19 & \\
e - p & 1.00 & -0.25 & 0.24 & 0.29 \\
d - e & 1.00 & -0.38 & -0.61 & & & \\
\Delta d & 1.00 & 0.31 & & & & \\
\Delta e & & & & 1.00 & & \\
\end{array}
\]
Table 2: VAR estimates

This table reports the one-year restricted VAR estimation results when the predictor is the market earnings yield. The variables in the VAR are the log stock return ($r$), log earnings growth ($\Delta e$), log dividend payout ratio ($d - e$), and log earnings-to-price ratio ($e - p$). $b(\phi)$ denote the VAR slopes associated with lagged $e - p$, while $t$ denotes the respective Newey and West (1987) $t$-statistics (calculated with one lag). $b'(\phi')$ denote the slope estimates implied from the variance decomposition for $e - p$, 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 %. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>$b(\phi)$</th>
<th>$t$</th>
<th>$b'(\phi')$</th>
<th>$t$</th>
<th>$R^2(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A (1872-2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>0.068</td>
<td>1.82</td>
<td>0.071</td>
<td>1.88</td>
<td>2.10</td>
</tr>
<tr>
<td>$d - e$</td>
<td>0.050</td>
<td>0.61</td>
<td>-0.022</td>
<td>-0.31</td>
<td>0.37</td>
</tr>
<tr>
<td>$\Delta e$</td>
<td>-0.253</td>
<td>-2.93</td>
<td>-0.256</td>
<td>-2.98</td>
<td>10.51</td>
</tr>
<tr>
<td>$e - p$</td>
<td>0.706</td>
<td>9.65</td>
<td>0.709</td>
<td>9.66</td>
<td>49.85</td>
</tr>
<tr>
<td>Panel B (1872-1925)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>-0.007</td>
<td>-0.10</td>
<td>-0.009</td>
<td>-0.14</td>
<td>0.02</td>
</tr>
<tr>
<td>$d - e$</td>
<td>-0.084</td>
<td>-0.69</td>
<td>-0.031</td>
<td>-0.27</td>
<td>0.91</td>
</tr>
<tr>
<td>$\Delta e$</td>
<td>-0.469</td>
<td>-2.92</td>
<td>-0.467</td>
<td>-2.90</td>
<td>22.16</td>
</tr>
<tr>
<td>$e - p$</td>
<td>0.563</td>
<td>3.56</td>
<td>0.561</td>
<td>3.56</td>
<td>31.29</td>
</tr>
<tr>
<td>Panel C (1926-2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>0.101</td>
<td>2.44</td>
<td>0.103</td>
<td>2.48</td>
<td>4.74</td>
</tr>
<tr>
<td>$d - e$</td>
<td>0.029</td>
<td>0.35</td>
<td>-0.021</td>
<td>-0.30</td>
<td>0.15</td>
</tr>
<tr>
<td>$\Delta e$</td>
<td>-0.189</td>
<td>-1.86</td>
<td>-0.190</td>
<td>-1.88</td>
<td>6.85</td>
</tr>
<tr>
<td>$e - p$</td>
<td>0.735</td>
<td>8.85</td>
<td>0.737</td>
<td>8.85</td>
<td>54.42</td>
</tr>
<tr>
<td>Panel D (1946-2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>0.100</td>
<td>2.32</td>
<td>0.100</td>
<td>2.31</td>
<td>7.69</td>
</tr>
<tr>
<td>$d - e$</td>
<td>0.012</td>
<td>0.17</td>
<td>0.025</td>
<td>0.34</td>
<td>0.04</td>
</tr>
<tr>
<td>$\Delta e$</td>
<td>-0.141</td>
<td>-1.27</td>
<td>-0.141</td>
<td>-1.26</td>
<td>4.30</td>
</tr>
<tr>
<td>$e - p$</td>
<td>0.783</td>
<td>8.76</td>
<td>0.783</td>
<td>8.76</td>
<td>61.32</td>
</tr>
</tbody>
</table>

Table 3: Probability values for return and earnings growth slopes

This table reports the fractions of Monte-Carlo simulated values of the return ($b^K_r$) and earnings growth coefficients ($b^K_e$) that are higher/lower than the respective estimates from the original sample ($b^K_{r0}$, $b^K_{e0}$). The predictive variable is the earnings yield and the forecasting horizon is one year (Panel A) and 20 years ahead (Panel B). The total sample is 1872–2010 and the results for different subsamples are also provided. For details on the Monte-Carlo simulation see Section 5.

<table>
<thead>
<tr>
<th></th>
<th>$b^K_r &lt; b^K_{r0}$</th>
<th>$b^K_r &gt; b^K_{r0}$</th>
<th>$b^K_e &lt; b^K_{e0}$</th>
<th>$b^K_e &gt; b^K_{e0}$</th>
<th>$b^K_r &lt; b^K_{r0}$, $b^K_e &lt; b^K_{e0}$</th>
<th>$b^K_r &gt; b^K_{r0}$, $b^K_e &gt; b^K_{e0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A ($K = 1$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–2010</td>
<td>0.94</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1872–1925</td>
<td>0.44</td>
<td>0.56</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1926–2010</td>
<td>0.91</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1946–2010</td>
<td>0.82</td>
<td>0.02</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Panel B ($K = 20$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872–2010</td>
<td>0.89</td>
<td>0.02</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1872–1925</td>
<td>0.13</td>
<td>0.86</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1926–2010</td>
<td>0.50</td>
<td>0.01</td>
<td>0.49</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1946–2010</td>
<td>0.36</td>
<td>0.00</td>
<td>0.64</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Figure 1: Dividend yield and Earnings yield

This figure plots the time-series for the market dividend-to-price \((d/p)\) and earnings-to-price \((e/p)\) ratios (original series). The sample is 1872–2010.
Figure 2: Dividend growth, earnings growth, and dividend payout ratio

This figure plots the time-series for the aggregate dividend ($\Delta d$) and earnings growth rates ($\Delta e$) (Panel A), and the dividend payout ratio (Panel B). The sample is 1872–2010.
Figure 3: Term structure of direct long-run coefficients: dividend yield
This figure plots the term structure of direct long-run predictive coefficients for the variance decomposition associated with the dividend yield. 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. “sum” denotes the value of the variance decomposition, in %. The long-run coefficients are measured in %, and $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 4: Term structure of $t$-stats for direct long-run coefficients: dividend yield
This figure plots the term structure of the $t$-statistics associated with the direct long-run predictive coefficients for the dividend yield variance decomposition. 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. The horizontal lines represent the 5% critical values (-1.96, 1.96). $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 5: Term structure of direct long-run coefficients: earnings yield
This figure plots the term structure of direct long-run predictive coefficients for the variance decomposition associated with the earnings yield. The predictive slopes are associated with the log return \((r)\), log earnings growth \((e)\), log dividend payout ratio \((de)\), and log earnings-to-price ratio \((ep)\). The forecasting variable is the log earnings-to-price ratio. “sum” denotes the value of the variance decomposition, in %. The long-run coefficients are measured in %, and \(K\) represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 6: Term structure of $t$-stats for direct long-run coefficients: earnings yield

This figure plots the term structure of the $t$-statistics associated with the direct long-run predictive coefficients for the earnings yield variance decomposition. The predictive slopes are associated with the log return ($r$), log earnings growth ($e$), log dividend payout ratio ($de$), and log earnings-to-price ratio ($ep$). The forecasting variable is the log earnings-to-price ratio. The horizontal lines represent the 5% critical values (-1.96, 1.96). $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 7: Term structure of VAR-based long-run coefficients: earnings yield
This figure plots the term structure of the VAR-based long-run predictive coefficients for the variance decomposition associated with the earnings yield. The predictive slopes are associated with the log return ($r$), log earnings growth ($e$), log dividend payout ratio ($de$), and log earnings-to-price ratio ($ep$). The forecasting variable is the log earnings-to-price ratio. “sum” denotes the value of the variance decomposition, in %. The long-run coefficients are measured in %, and $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 8: Term structure of $t$-stats for VAR-based long-run coefficients: earnings yield

This figure plots the term structure of the $t$-statistics associated with the VAR-based long-run predictive coefficients for the earnings yield variance decomposition. The predictive slopes are associated with the log return ($r$), log earnings growth ($e$), log dividend payout ratio ($de$), and log earnings-to-price ratio ($ep$). The forecasting variable is the log earnings-to-price ratio. The horizontal lines represent the 5% critical values (-1.96, 1.96). $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 9: Term-structure of simulated p-values
This figure plots the simulated p-values for the return ($r$) and earnings growth ($e$) slopes from a Monte-Carlo simulation with 10,000 replications under the null of no return and earnings growth predictability. The predictive variable is the earnings yield. The numbers indicate the fraction of pseudo samples under which the return (earnings growth) coefficient is higher (lower) than the corresponding estimates from the original sample. $K$ represents the number of years ahead. The total sample is 1872–2010 and the results for different subsamples are also provided. For details on the Monte-Carlo simulation see Section 5.
Figure 10: Joint distribution of return and earnings growth coefficients ($K = 1$)
This figure reports the joint distribution of the return ($b_r$) and earnings growth coefficients ($b_e$) from a Monte-Carlo simulation with 10,000 replications under the null of no return and earnings growth predictability. The predictive variable is the earnings yield and the forecasting horizon is one year. The total sample is 1872–2010 and the results for different subsamples are also provided. The lines represent the sample estimates. For details on the Monte-Carlo simulation see Section 5.
Figure 11: Joint distribution of return and earnings growth coefficients ($K = 20$)
This figure reports the joint distribution of the return ($b_r$) and earnings growth coefficients ($b_e$) from a Monte-Carlo simulation with 10,000 replications under the null of no return and earnings growth predictability. The predictive variable is the earnings yield and the forecasting horizon is 20 years. The total sample is 1872–2010 and the results for different subsamples are also provided. The lines represent the sample estimates. For details on the Monte-Carlo simulation see Section 5.
Figure 12: Term structure of direct long-run coefficients: excess returns

This figure plots the term structure of direct long-run predictive coefficients for the variance decomposition associated with the earnings yield. The predictive slopes are associated with the excess log return ($r$), log earnings growth ($e$), log dividend payout ratio ($de$), log short-term interest rate ($f$), and log earnings-to-price ratio ($ep$). The forecasting variable is the log earnings-to-price ratio. “sum” denotes the value of the variance decomposition, in %. The long-run coefficients are measured in %, and $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.
Figure 13: Term structure of $t$-stats for direct long-run coefficients: excess returns

This figure plots the term structure of the $t$-statistics associated with the direct long-run predic-
tive coefficients for the earnings yield variance decomposition. The predictive slopes are associ-
ated with the excess log return ($r$), log earnings growth ($e$), log dividend payout ratio ($de$), log short-term interest rate ($f$), and log earnings-to-price ratio ($ep$). The forecasting variable is the log earnings-to-price ratio. The horizontal lines represent the 5% critical values (-1.96, 1.96). $K$ represents the number of years ahead. The full sample corresponds to annual data for the 1872–
2010 period (Panel A) and Panels B to D present the estimation results for different subsamples.