Reconciling the Return Predictability Evidence *

Martin Lettau
New York University Stern School of Business, CPER and NBER

Stijn Van Nieuwerburgh
New York University Stern School of Business

April 27, 2006

Abstract

Evidence of stock return predictability by financial ratios is still controversial, as documented by inconsistent results for in-sample and out-of-sample regressions and by substantial parameter instability. This paper shows that these seemingly incompatible results can be reconciled if the assumption of a fixed steady-state mean of the economy is relaxed. We find strong empirical evidence in support of shifts in the steady-state and propose simple methods to adjust financial ratios for such shifts. The in-sample forecasting relationship of adjusted price ratios and future returns is statistically significant and stable over time. In real-time, however, changes in the steady-state make the in-sample return forecastability hard to exploit out-of-sample. The uncertainty of estimating the size of steady-state shifts rather than the estimation of their dates is responsible for the difficulty of forecasting stock returns in real-time. Our conclusions hold for a variety of financial ratios and are robust to changes in the econometric technique used to estimate shifts in the steady-state.

*Lettau: Department of Finance, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012; mlettau@stern.nyu.edu; Tel: (212) 998-0378; http://www.stern.nyu.edu/~mlettau. Van Nieuwerburgh: Department of Finance, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012; svnieuwe@stern.nyu.edu; Tel: (212) 998-0673; http://www.stern.nyu.edu/~svnieuwe. We thank an anonymous referee, Matt Spiegel (the editor), Yakov Amihud, John Campbell, Kenneth French, Sydney Ludvigson, Eli Ofek, Matthew Richardson, Ivo Welch, Robert Whitelaw and the seminar participants at Duke, McGill, NYU, UNC and Wharton for comments.

1 Introduction

The question of whether stock returns are predictable has received an enormous amount of attention. This is not surprising because the existence of return predictability is not only of interest to practitioners but also has important implications for financial models of risk and return. One branch of the literature asserts that expected returns contain a time-varying component that implies predictability of future returns. Due to its persistence, the predictive component is stronger over longer horizons than over short horizons. Classic predictive variables are financial ratios, such as the dividend-price ratio, the earnings-price ratio, and the book-to-market ratio (Rozeff (1984), Fama and French (1988), Campbell and Shiller (1988), Cochrane (1991), Goetzman and Jorion (1993), Hodrick (1992), Lewellen (2004), and others), but other variables have also been found to be powerful predictors of long-horizon returns (e.g., Lettau and Ludvigson (2001), Lustig and Van Nieuwerburgh (2005a), Menzly, Santos, and Veronesi (2004), Piazzesi, Schneider, and Tuzel (2006)). Moreover, these studies conclude that growth rates of fundamentals, such as dividends or earnings, are much less forecastable than returns, suggesting that most of the variation of financial ratios is due to variations in expected returns.

These conclusions are controversial because the forecasting relationship of financial ratios and future stock returns exhibits a number of disconcerting features. First, correct inference is problematic because financial ratios are extremely persistent; in fact, standard tests leave the possibility of unit roots open. Nelson and Kim (1993), Stambaugh (1999), Ang and Bekaert (2006), Ferson, Sarkissian, and Simin (2003), and Valkanov (2003) conclude that the statistical evidence of forecastability is weaker once tests are adjusted for high persistence. Second, financial ratios have poor out-of-sample forecasting power, as shown in Bossaerts and Hillion (1999) and Goyal and Welch (2003, 2004), but see Campbell and Thompson (2005) for a different interpretations of the out-of-sample evidence [1]. Third, and related to the poor out-of-sample evidence, the forecasting relationship of returns and financial ratios exhibits significant instability over time. For example, in rolling 30-year regressions of annual log CRSP value-weighted returns on lagged log dividend-price ratios, the ordinary least squares (OLS) regression coefficient varies between zero and 0.5 and the associated R^2 ranges from close to zero to 30% depending on the subsample. Not surprisingly, the hypothesis of a constant regression coefficient is routinely rejected (Viceira (1996), Paye and Timmermann (2005)).

In addition to concerns that return forecastability might be spurious, the benchmark model of time-varying expected returns faces additional challenges. The extreme persistence of price ratios implies that expected returns have to be extremely persistent as well. But if shocks to expected returns have a half-life of many years or even decades, as implied by the high persistence

[1] There is some ambiguity about the use of the term "forecast" in this literature. Most papers use "forecast" to refer to in-sample regressions using the entire sample. In contrast, predictions using only currently available data are referred to as "out-of-sample forecasts." We follow this convention.

of financial ratios, they are unlikely to be linked to many plausible economic risk factors, such as those linked to business cycles. Instead, researchers have to identify slow-moving factors that are primary determinants of equity risk. In addition, the extraordinary valuation ratios in the late 1990s represent a significant challenge for the benchmark model. Given the historical record of returns, fundamentals, and prices, it is exceedingly unlikely that persistent stationary shocks to expected returns are capable of explaining price multiples like those seen in 1999 or 2000.

In summary, the return predictability literature has yet to provide convincing answers to the following four questions: What is the source of parameter instability? Why is the out-of-sample evidence so much weaker than the in-sample evidence? Why has even the in-sample evidence disappeared in the late 90s? Why are price ratios extremely persistent? In this paper, we show that these puzzling empirical patterns can be explained if the steady-state mean of financial ratios has changed over the course of the sample period. Such changes could be due to changes in the steady-state growth rate of economic fundamentals resulting from permanent technological innovations and/or changes in the expected return of equity caused by, for example, improved risk sharing, changes in stock market participation, changes in the tax code, or lower macroeconomic volatility.

Using standard econometric techniques, we show that the hypothesis of permanent changes in the mean of various price ratios is supported by the data. We then ask how such changes affect the forecasting relationship of returns and lagged price ratios. Standard econometric techniques that assume that the regressor is stationary will lead to biased estimates and incorrect inference. However, since deviations of price ratios from their steady-state values are stationary, it is straightforward to correct for the non-stationarity if the timing and magnitudes of shifts in steady states can be estimated. We conduct tests that incorporate such adjustments from the perspective of an econometrician with access to the entire historical sample (in-sample tests), as well as from the perspective of an investor who forecasts returns in real-time (out-of-sample tests).

Our in-sample results conclude that "adjusted" price ratios have favorable properties compared to unadjusted price ratios. In the full sample, the slope coefficient in regressions of annual log returns on the lagged log dividend-price ratio increases from 0.094 for the unadjusted ratio to 0.235 and 0.455 for the adjusted ratio with one and two steady-state shifts, respectively. While the statistical significance of the coefficient on the unadjusted dividend-price ratio is marginal, coefficients on the adjusted dividend-price ratios are strongly significant. Finally, the regression coefficients using adjusted price ratios as regressors are more stable over time. We find similar differences for other price ratios such as the earnings-price ratio and the book-to-market ratio.

In real time, however, the changes in the steady-state are not only difficult to detect but also estimated with significant uncertainty, making the in-sample return forecastability hard to exploit. Results for out-of-sample forecasting tests reflect this difficulty. While adjusted price ratios have superior out-of-sample forecasting power relative to their unadjusted counterparts, they do not

outperform the benchmark random walk model. Why does the real time prediction fail to beat the random walk model? In real time an investor faces two challenges. First, she has to estimate the timing of a break. Second, if she detects a new break, she has to estimate the new mean after the break occurred. If the new break occurred towards the end of the sample that the investor has access to, the new mean has to be estimated using a small number of observations and is subject to significant uncertainty. We perform additional test to evaluate the relative difficulty of estimating the break dates versus estimating the means relative to the pure out-of-sample forecasts and the ex-post adjusted dividend price ratio. We find that *(i)* the estimation of the break dates in real time is not crucial and the resulting prediction errors are smaller than for the random walk model, and *(ii)* that the estimation of the magnitude of the break in the mean dividend-price ratio entails substantial uncertainty, and is ultimately responsible for the failure of the real time out-of-sample predictions to beat the random walk. These findings can explain the lack of out-of-sample predictability documented by Goyal and Welch (2004).

Several papers have explored the impact of structural breaks on return predictability. For example, Viceira (1996) and Paye and Timmermann (2005) reported evidence in favor of breaks in the OLS coefficient in the forecasting regression of returns on the lagged dividend-price ratio. Our focus is instead on shifts in the mean of financial ratios, which in turn render the forecasting relationship unstable if such shifts are not taken into account. In other words, in contrast to Viceira (1996) and Paye and Timmermann (2005), we focus on the behavior of the mean of price ratios instead of the behavior of the slope coefficient. Pastor and Stambaugh (2001) use a Bayesian framework to estimate breaks in the equity premium. They found several shifts in the equity premium since 1834 and identified the sharpest drop in the 1990s, which is consistent with the timing of the shift in price ratios identified in this paper. This paper is also related to the recent literature on inference in forecasting regressions with persistent regressors (see e.g., Amihud and Hurwich (2004), Ang and Bekaert (2006), Campbell and Yogo (2002), Lewellen (2004), Torous, Volkanov, and Yan (2004), and Eliasz (2005)). In these papers, asymptotic distributions for OLS regressions are derived under the assumption that the forecasting variable is a close-to unit, yet stationary, root process. In contrast, we allow for the presence of a small but econometrically important non-stationary component in forecasting variables.

The rest of the paper is organized as follows. In Section 2 we establish that the standard dividend-price ratio does not significantly forecast stock returns or dividend growth. In contrast, we find much stronger evidence for return predictability in various subsamples. The slope coefficient in the return equation is much smaller in the full sample than in any of the constituent subsamples, which confirms the instability of the forecasting relationship over time. In sections 3 and 4, we show how changes in the steady-state affect the dividend-price ratio and other price ratios. For the log dividend-price ratio, we find evidence for either one break in the early 1990s or two breaks around 1954 and 1994. Other valuation ratios such as the earnings-price ratio and the book-to-market

value ratio exhibit similar breaks. We show that filtering out this non-stationary component yields adjusted price ratios that have strong and stable in-sample return predictability. In Section 5 we study out-of-sample predictability. We use a recursive Perron procedure that estimates both the break dates and the means of the regimes in real-time. We show that using the break-adjusted dividend-price series produces superior one-step ahead return forecasts compared to using the unadjusted dividend-price series, but does slightly worse than the naive random walk model. Using a Hamilton (1989) regime-switching model, we show that if the investor did not have to estimate regime means in real time, but only the regime switching dates, her out-of-sample forecast would improve substantially, and beat the random walk. The Hamilton procedure leads to slightly later break dates but predictability results that are virtually as good as those when the (ex-post) break dates were known and used. In sum, the hardest part of real-time out-of-sample prediction in the presence of regimes is the estimation uncertainty about the mean of the new regime. In Section 6 we consider a vector error correction model that includes the return and dividend growth predictability equations and imposes a joint present value restriction on the slope parameters from both equations. We find that this restriction is satisfied when we use the adjusted dividend-price ratio as an independent variable, but not when we use the unadjusted series. We use this framework to estimate long-horizon regressions. Finally, in Section 7, we find that our simple model serves as a plausible data generating process. It is able to replicate both the findings of no predictability when the unadjusted dividend-price ratio is used and the findings of in-sample and out-of-sample predictability when the adjusted series is used.

2 Instability of Forecasting Relationships

In this section we document the instability of the forecasting relationship between returns, dividend growth, and the lagged dividend-price ratio. The forecasting relationship of returns and other financial ratios (such as the earnings-price ratio and the book-to-market ratio) and alternative measures of dividends (such as accounting for repurchases or considering only dividend-paying firms) are similar and will be presented later. The data are based on annual CRSP value-weighted returns from 1927 to 2004 and are described in Appendix A. The top panel of Figure 1 shows the estimation results for the forecasting regression of demeaned returns on the demeaned lagged dividend-price ratio using 30-year rolling windows:

$$r_{t+1} - \overline{r} = \kappa_r(dp_t - \overline{dp}) + \tau^r_{t+1}, \tag{1}$$

where r_t denotes the log return, dp_t denotes the log dividend-price ratio $d_t - p_t$, and $\bar{r}$ and $\overline{dp}$ denote the sample means of returns and the log dividend-price ratio in each of the subsamples, respectively. The top panel plots the slope coefficient κ_r along with two standard error bands.

The instability of the forecasting relationship is strikingly illustrated by the variation of the return predictability coefficient over time. The estimates of κ_r are around 0.5 in the subsamples ending in the late 1950s and in the samples ending in the early 1980s to the mid 1990s. In contrast, κ_r is much smaller for the samples ending in the mid 1960s and is close to zero and statistically insignificant in samples ending in the late 1990s and early 2000s. Similarly, the R^2 of the forecasting regression displays instability with values ranging from 34% in 1982 to 0% at the end of the 1990s (not shown). This evidence has led some researchers to conclude that the dividend-price ratio does not forecast stock returns, or at least not robustly so. Not surprisingly, the hypothesis of a constant regression coefficient is routinely rejected.

We also estimate a predictability regression for demeaned dividend growth rates:

$$\Delta d_{t+1} - \overline{d} = \kappa_d(dp_t - \overline{dp}) + \tau^d_{t+1}, \tag{2}$$

where d_t denotes log dividends and $\overline{d}$ denotes the sample mean of dividend growth. Dividend growth rates are even less forecastable than returns. For most of the sample, the point estimate is not statistically significantly different from zero, and the regression R^2 never exceeds 16% (not shown). Interestingly, the dividend-price ratio at the end of the 1990s seems to forecast neither stock returns nor dividend growth. This is a conundrum from the perspective of any present value model (see Section 3.1), as also pointed out by Cochrane (2006).

[Figure 1 about here.]

The left two columns of Table 1, denoted "No Break," report the coefficients κ_r and κ_d from equations (1) and (2) and their asymptotic standard errors for the entire 1927-2004 sample, as well as for various subsamples. The first row shows that the dividend-price ratio marginally predicts stock returns (first column); the coefficient is significant at the 5% level if asymptotic standard errors are used for inference. However, small sample standard errors computed from a bootstrap simulation suggest that the coefficient κ_r is not statistically different from zero for the entire sample.[2] The dividend-price ratio does not forecast dividend growth at conventional significance

[2]Asymptotic standard errors may be a poor indicator of the estimation uncertainty in small samples, and the p-values for the null of no predictability may be inaccurate. The asymptotic corrections advocated by Hansen and Hodrick (1980) have poor small sample properties. Ang and Bekaert (2006) find that use of those standard errors leads to over-rejection of the no-predictability null. The bootstrap exercise imposes the null of no predictability and asks how likely it is to observe the estimated κ_r coefficients reported in the first column of Table 1. We find that the small sample p-value for κ^r is 6.8% compared to an asymptotic p-value of 4.1%. We also conduct a second bootstrap exercise to find the small-sample bias in the return coefficient. Consistent with Stambaugh (1999), we find an upward bias. If the true value is .094, the bootstrap exercise estimates a coefficient of .115. Detailed results are available upon request. The empirical size of tests based asymptotic and bootstrapped standard errors tend to be larger than their nominal size if the regressor is highly persistent (e.g. Amihud, Hurvich, and Wang (2005)). Alternative tests with better size properties weaken the evidence for forecastability with the dividend price ratio further.

levels (third column). Thus, we cannot reject the hypothesis that the dividend-price ratio forecasts neither dividend growth nor returns.

Rows 2 and 3 report the results for two non-overlapping samples that span the entire period: 1927-1991 and 1992-2004. We will justify this particular choice of subsamples in Section 3. The estimates of κ_r display a remarkable pattern across subsamples: In both subsamples κ_r is much larger than its estimate in the whole sample. In fact, the estimates are almost identical in the two subsamples: .2353 in the 1927-1991 subsample compared to .2351 in the later 1992-2004 subsample. Yet, when we join the two subsamples, the point estimate drops to .094. In addition, κ_r is strongly statistically significant in both subsamples but only marginally significant in the whole sample. Confirming the instability of κ_r estimates, row 4 reports the results of a Chow test, which rejects the null hypothesis of no structural break in 1991 at the 4% level. Finally, the dividend growth forecasting relationship displays less instability, and the coefficient remains insignificant in both subsamples.

The pattern of κ_r is not unique to the specific subsamples chosen. We obtain very similar results when we use three non-overlapping subsamples: 1927-1954, 1955-1994, and 1995-2004 (bottom half of Table 1). Again, we find that the return predictability coefficient κ_r is estimated to be much higher in each of the three subsamples than in the entire subsample. In row 5, the predictability coefficient is .09, whereas it is .51, .38, and .53 in rows 6, 7, and 8 respectively. Moreover, it is statistically significant in each subsample. Row 9 shows that we strongly reject the joint null hypothesis of parameter stability in 1954 and 1994. For dividend growth, the evidence is more mixed. We fail to reject the same null hypothesis of no breaks in 1954 and 1994, but the κ_d coefficient is marginally statistically different from zero in rows 7 and 8.

Finally, the last two columns repeat the analysis using returns in excess of a 90-day Treasury bill rate instead of gross returns. The exact same findings hold for excess returns. In the rest of the analysis, we proceed with gross returns only.

[Table 1 about here.]

We conclude that the forecasting relationship between returns and the dividend-price ratio is unstable over time. Coefficient estimates of κ_r are almost identical in non-overlapping subsamples, but the point estimate for the whole sample is much lower than it is in each of the subsamples. Next, we investigate what might explain this intriguing pattern of the regression coefficients that links returns to past dividend-price ratios.

3 Steady-State Shifts and Forecasting

The macroeconomics literature has recently turned to models with persistent changes in fundamentals to explain the dramatic change in valuation ratios in the bull market of the 1990s. Most such

models imply a persistent decline in expected returns or an increase in the steady-state growth rate of the economy. Lettau, Ludvigson, and Wachter (2004) argued that a persistent decline in the volatility of aggregate consumption growth leads to a decline in the equity premium. Another class of models argues for persistent improvements in the degree of risk-sharing among households or regions, either due to developments in the market for unsecured debt or the market for housing-collateralized debt (Krueger and Perri (2005) and Lustig and Van Nieuwerburgh (2005b)). In the model of Lustig and Van Nieuwerburgh (2005c), the improvement in risk sharing implies a persistent decline in the equity premium. McGrattan and Prescott (2005) argued that persistent changes in the tax code can explain the persistent decline in the equity premium. Lastly, models of limited stock market participation argue that the gradual entry of new participants has persistently depressed equity premia (Vissing-Jorgensen (2002), Calvet, Gonzalez-Eiras, and Sadini (2003), and Guvenen (2003)). Other models argue that there was a persistent increase in the long-run growth rate of the economy in the 1990s (Quadrini and Jermann (2003) and Jovanovic and Rousseau (2003)). The first set of models lower the long-run required return of equity ($\overline{r}$), the last set of models raise the long-term growth rate of the economy ($\overline{d}$). A intuition based on the Gordon growth model implies that either effect lowers the steady-state level of the dividend-price ratio $\overline{dp}$. In this section, we augment the Campbell-Shiller framework for such changes in $\overline{dp}$, we estimate these shifts in the data, and explore their implications for return predictability.

3.1 Changes in the Mean of Price Ratios

The standard specification of stock returns and forecasting variables assumes that all processes are stationary around a constant mean. For example, Stambaugh (1986, 1999), Mankiw and Shapiro (1986), Nelson and Kim (1993), and Lewellen (1999) considered the following model:

$$r_{t+1} = \bar{r} + \kappa_r y_t + \tau^r_{t+1} \tag{3}$$

$$y_t = \bar{y} + v_t. \tag{4}$$

The mean of the forecasting variable y_t, $\bar{y}$, is constant and the stochastic component v_t is assumed to be stationary, often specified as an AR(1) process. Means of financial ratios are determined by properties of the steady-state of the economy. For example, the mean of the log dividend-price ratio $\overline{dp}$ is a function of the growth rate $\overline{d}$ of log dividends and expected log return $\overline{r}$ in steady-state:

$$\overline{dp} = \log(\exp(\overline{r}) - \exp(\overline{d})) - \overline{d}, \tag{5}$$

whereas the stochastic component depends on expected future deviations of returns and dividend growth from their steady-state values (Campbell and Shiller (1988)):

$$dp_t = \overline{dp} + E_t \sum_{j=1}^{\infty} \rho^{j-1} \big[(r_{t+j} - \overline{r}) - (\Delta d_{t+j} - \overline{d}) \big], \tag{6}$$

where $\rho = (1 + \exp(\overline{dp}))^{-1}$ is a constant. Similar equations can be derived for other financial ratios (e.g., Vuolteenaho (2000)). Berk, Green, and Naik (1999) show how stock returns and book-to-market ratios are related in a general equilibrium model.

A crucial assumption is that the steady-state of the economy is constant over time: The average long-run growth rate of the economy as well as the average long-run return of equity are fixed and not allowed to change. However, if either the steady-state growth rate or expected return were to change, the effects on financial ratios and their stochastic relationships with returns would be profound. Even relatively small changes in long-run growth and/or expected return have large effects on the mean of the dividend-price ratio, as can be seen from (5). The effects of steady-state shifts on other valuation ratios, such as the earnings-price ratio and the book-to-market ratio, are similar. In this paper, we entertain the possibility that the steady change of the U.S. economy has indeed changed since 1926, and we study the effect of these changes on the forecasting relationship of returns and price ratios.

A steady-state is characterized by long-run growth and expected return. Any short-term deviation from steady-state is expected to be only temporary and the economy is expected to return to its steady-state eventually. Thus, steady-state growth and expected return must be constant in expectations, but the steady-state might shift unexpectedly. Correspondingly, we assume that $E_t \overline{r}_{t+j} = \overline{r}_t, E_t \overline{d}_{t+j} = \overline{d}_t, E_t \overline{dp}_{t+j} = \overline{dp}_t$.[3]

The log linear framework introduced above illustrates the effect of time-varying steady-states, though none of our results depend on the accuracy of the approximation. Just as in the case with constant steady-state, the log dividend-price ratio is the sum of the steady-state dividend-price ratio and the discounted sum of expected returns minus expected dividend growth in excess of

[3]Although the log dividend-price ratio is a nonlinear function of steady-state returns and growth, we assume that the steady-state log dividend-price ratio is also (approximately) a martingale: $E_t \overline{dp}_{t+j} = \overline{dp}_t$. This assumption is justified for the specific processes for steady-state returns and growth that we will consider below. Appendix C spells out a simple asset pricing model where the price-dividend ratio in levels follows a (bounded) martingale. It shows that $\overline{dp}$ and $\overline{d}$ are approximate (bounded) martingales.

steady-state growth and returns [4]:

$$dp_t = \overline{dp}_t + E_t \sum_{j=1}^{\infty} \rho_t^{j-1} \big[(r_{t+j} - \overline{r}_t) - (\Delta d_{t+j} - \overline{d}_t) \big], \tag{7}$$

where $\rho_t = (1 + \exp(\overline{dp}_t))^{-1}$. The important difference of (7) compared to (6) is that the mean of the log dividend-price ratio is no longer constant. In fact, it not only varies over time but it is non-stationary. If, for example, the steady-state growth rate increases permanently, the steady-state dividend-price ratio decreases and the current log dividend-price ratio declines permanently. While the log dividend-price ratio contains a non-stationary component it is important to note that *deviations* of dp_t from steady-states are stationary as long as deviations of dividend growth and returns from their respective steady-states are stationary, an assumption we maintain throughout the paper.[5] In other words, the dividend-price ratio dp_t itself contains a non-stationary component $\overline{dp}_t$ but the appropriately demeaned dividend-price ratio $dp_t - \overline{dp}_t$ is stationary. The implications for forecasting regressions with the dividend-price ratio are immediate. First, in the presence of steady-state shifts, a non-stationary dividend-price ratio is not a well-defined predictor and this non-stationarity could be the cause for the empirical patterns described in the previous section. Second, the dividend-price ratio must be adjusted to remove the non-stationary component $\overline{dp}_t$ to render a stationary process.

While we emphasized the effect of steady-state shifts on the dividend-price ratio, the intuition carries through to other financial ratios. Changes in the steady-state have similar effects on the earnings-price ratio and the book-to-market ratio. However, other permanent changes in the economy, such as changes in payout policies, could affect different ratios differently. In the following section, we provide evidence that steady-state shifts have occurred in our sample and propose simple methods to adjust financial ratios for such shifts.

[4] Appendix B presents a detailed derivation. Under our assumption, the log approximation in a model with time-varying steady-states is as accurate as the approximation for the corresponding model with constant steady-state. In fact, the ex-ante expressions of the approximate log dividend-price ratio (6) and (7) are exactly the same. Only their ex-post values are different in periods when the steady-state shifts.

[5] Of course, in a finite sample it is impossible to conclusively distinguish a truly permanent change from an extremely persistent one. Thus, our insistence of non-stationarity might seem misguided. However, the important insight is that the dividend-price ratio is not only a function of (less) persistent changes in expected growth rates and expected returns that could potentially have cyclical sources but is also affected by either extremely persistent or permanent structural changes in the economy. This distinction turns out be very useful, as we will show in the remainder of the paper. In this sense our assumption of true non-stationarity can be regarded to include "extremely persistent but stationary." In a finite sample, the conclusions will be the same in either setting. The distinction of "permanent" versus "extremely persistent" is important, however, for structural asset pricing models because permanent shocks might have much larger impact on prices than very persistent ones.

3.2 Steady-state Shifts in the Dividend-Price Ratio

Has the steady-state relationship of growth rates and expected returns shifted since the beginning of our sample in 1926? If so, have these shifts affected the stochastic relationship between returns and price ratios? In this section we use econometric techniques that exploit the entire sample to detect changes in the steady-state. In section 5, we study how investors in real-time might have assessed the possibility of shifts in the steady-states without the benefit of knowing the whole sample. In both cases, there is strong empirical evidence in favor of changes in the steady-state and we find that such changes have dramatic effects on the forecasting relationship of returns and price ratios. We suggest a simple adjustment to the dividend-price ratio and revisit the forecasting equations from Section 2. We first study shifts in the dividend-price ratios in detail and consider alternative ratios in section 4.

Our econometric specification is directly motivated by the framework that allows for changes in the steady-state laid out in the previous section. Equation (7) implies that the log dividend-price ratio is the sum of a non-stationary component and a stationary component. In this section, we model the non-stationary component as a constant that is subject to rare structural breaks as in Perron (1989).[6]

The full line in each of the panels of Figure 2 shows the log dividend-price ratio from 1927 to 2004. Visually, the series displays evidence of non-stationarity. Especially the bull market of the 1990s seems hard to reconcile with a stationary model. The dividend-price ratio has risen since, but at the end of our sample in 2004, prices would have to fall an additional 46% for the dividend-price ratio to return to its historical mean. A first explanation we entertain is that the bull market of the 1990s represents a sequence of extreme realizations from a stationary distribution.

[Figure 2 about here.]

The solid line in Figure 3 shows the smoothed empirical distribution of the log dividend-price ratio dp_t. This distribution has a fat left tail, mainly due to the observations in the last 15 years. To investigate whether this is a typical plot from a stationary distribution, we conduct two exercises. Following Campbell, Lo, and MacKinlay (1997), Stambaugh (1999), Campbell and Yogo (2002), Ang and Bekaert (2006), and many others we estimate an AR(1) process for the log dividend-price ratio. First, in a bootstrap exercise, we draw from the empirical distribution with replacement. The smoothed bootstrap distribution is the dash-dotted line in the figure. Second, we compute the density of dp_t using Monte Carlo simulations from an estimated AR(1) model with normal innovation. This density is plotted as the dashed line. The graph shows that neither the bootstrap nor the Monte Carlo can replicate the fat left tail that we observe in the data. Interestingly, the

[6] As an alternative, we have also studied in-sample predictability in a Hamilton regime-switching model. Because the estimated regimes are so persistent, the predictability coefficients are very close to the ones we report here. In section 5 we revisit the Hamilton model in the context of out-of-sample predictability.

stationary model also cannot generate the right tail of the empirical distribution. In summary, it is unlikely that the dp_t data sample from 1927 to 2004 was generated by a stationary distribution.

[Figure 3 about here.]

An alternative explanation is that the long-run mean of the log dividend-price ratio is subject to structural breaks. To investigate this possibility, we test the null hypothesis of no break against the alternative hypotheses of one or two breaks with unknown break dates. Table 2 reports sup-F test statistics suggested by Perron (1989). The null hypothesis of no break is strongly rejected (the p-value is less than 1%) in favor of a break in 1991 or two breaks in 1954 and 1994. While the evidence against no breaks is very strong, the question whether the dividend-price ratio is subject to one or two breaks does not have a clear answer. The sup-F test of the null of a single break against the alternative of two breaks is rejected at the 10% level but not at the 5% level. The null of two breaks against the alternative of three breaks is not rejected (not shown). Alternatively, one can use an information criterion to select the number of breaks. Both the Bayesian Information Criterion (BIC) and the modified Schwartz criterion proposed by Liu, Wu, and Zidek (1997) (LWZ) favor two breaks. In summary, the data seem to strongly favor one or two breaks, rather than zero or three, but the relative evidence for one or two breaks is not as strong and only slightly in favor of two breaks.

[Table 2 about here.]

The table also reports the estimated change in the log dividend-price ratio before and after the break. In the one-break case, the change in $\overline{dp}$ is -.86, whereas in the two-break case, the first change in 1954 is -.37 and the second change is -.78. The two plots in the left column of Figure 2 overlay the long-run mean $\overline{dp}$ on the raw dp series. For now, we are agnostic as to whether the break(s) is (are) due to a change in the long-run mean of dividend growth or expected returns, or a combination of the two. We return to this question later. It is worth emphasizing, however, that the date(s) of the shift in the dividend-price ratio is (are) consistent with the breaks in the equity premium identified by Pastor and Stambaugh (2001).

This result motivates us to construct two adjusted dividend-price series, one for the one-break case and one for the two-break case. For each, we simply subtract the mean in the relative subsample(s). In the one-break case with break date τ, the adjusted ratio is defined as

$$\widetilde{dp}_t = \begin{cases} dp_t - \overline{dp}_1 & \text{for } t = 1, ..., \tau \\ dp_t - \overline{dp}_2 & \text{for } t = \tau + 1, ..., T, \end{cases} \tag{8}$$

where $\overline{dp}_1$ is the sample mean for 1927-1991 and $\overline{dp}_2$ is the sample mean for 1992-2004. The adjusted dp ratio in the two-break case is defined analogously. The right column of Figure 2 illustrates this procedure graphically.

The bottom half of Table 2 compares the autocorrelation properties of the unadjusted and adjusted dp series. As is well known, the raw dp series is very persistent. The first and second order autocorrelations are .91 and .81. The null hypothesis of a unit root cannot be rejected, according to an Augmented Dickey Fuller (ADF) test (third column). In contrast, the two adjusted $\widetilde{dp}$ series are much less persistent; the first order autocorrelation drops to .77 and .61, respectively. The null of a unit root in the adjusted series is rejected at the 4% and 1% levels. Interestingly, the volatility of the adjusted series is only half as large as for the adjusted series (last column). This substantially alleviates the burden on standard asset pricing models to match the volatility of the price-dividend ratio, once the non-stationary nature of the mean dp ratio has been taken into account.

3.3 Forecasting with the Adjusted Dividend-Price Ratio

We now revisit the return and dividend growth predictability equations (1) and (2), but use the adjusted dividend-price ratios instead of the raw series as predictor variable. The second and fourth columns of Table 1 show the estimation results of the return and dividend growth predictability regressions using $\widetilde{dp}$, respectively. Rows 1-4 are for the one-break case; rows 5-9 are for the two-break case. Starting with the one-break case, because the adjusted dividend-price ratio is the same as the raw series with each subsample, the results in rows 2 and 3 are unchanged. But now in row 1, we find that the adjusted dividend-price ratio significantly predicts stock returns. The coefficient for the entire sample is .235, which is almost identical to the estimates in the two subsamples. Thus the low point estimate for κ_r in the first column was due to averaging across regimes. Not taking the non-stationarity of the dp ratio into account severely biases the point estimate for κ_r downwards. Furthermore, row 4 shows that the evidence for a break in the forecasting relationship between returns and the dividend-price ratio has disappeared. The null hypothesis of parameter stability can no longer be rejected when using $\widetilde{dp}$. The full sample regression R^2 is 10%, more than twice the value of the first column. The results for dividend growth predictability remain largely unchanged. This is not surprising given that we did not detect much instability in the relationship between Δd_{t+1} and dp_t to begin with.

The rolling window estimates confirm this result.[7] The middle panel of Figure 1 shows that the coefficient κ_r is much more stable in the one-break case than in the no-break case (top panel). In particular, its value in the 1990s hovers around .3, compared to 0 without the adjustment. Likewise, the regression R^2 is also more stable and does not drop off in the 1990s. The same exercise shows that the dividend growth relationship is stable and that κ_d never moves far from zero (not shown). The evidence for dividend growth predictability is weak at best.[8]

[7] In the rolling window estimation we assume that the break in $\overline{dp}$ is caused by a break in mean expected returns $\bar{r}$. The alternative assumption that the break is in the long-run growth rate of the economy $\bar{g}$ gives identical results.

[8] The lack of predictive power of the dividend-price ratio for dividend growth does not imply that dividend growth

The bottom panel of Table 1 uses $\widetilde{dp}$, adjusted for breaks in 1954 and 1994. The full sample estimate for κ_r is now .455 (row 5) and highly significant.[9] The full sample regression R^2 is 22%. In contrast, dividend growth is not predictable. The bottom panel of Figure 1 shows that the rolling estimates for κ_r are very stable when we use $\widetilde{dp}$ adjusted for two breaks. The point estimate hovers around .4 and the return regression R^2 goes up as high as 40%. Moreover, the Chow test in row 9 finds no evidence for instability in either forecasting equation.

We conclude that taking changes in the long-run mean of the dividend-price ratio into account is crucial for forecasts of stock returns. Forecasting with the unadjusted dividend-price ratio series results in coefficient instability in the forecasting regression and unreliable inference (insignificance in small samples, and results depending on the subsample). These disconcerting properties are due to a non-stationary component that shifts the mean of the dividend-price ratio. In Section 3.1 we extended the model to allow for such non-stationarity in $\overline{dp}$. In this section we examined a simple form of non-stationarity, a structural break. Appropriately adjusting the dividend-price ratio for the structural break strengthens the evidence for return predictability, but not dividend growth predictability. The predictability coefficient is stable over time and least squares coefficient estimates are highly significant. Finally, the in-sample return predictability evidence stands up to the usual problem of persistent regressor bias (Nelson and Kim (1993), Stambaugh (1999), Ang and Bekaert (2006) and Valkanov (2003)) because the adjusted dividend-price ratio is much less persistent.

4 Other Financial Ratios

While the dividend-price ratio has been the classic prediction variable at least in the academic literature, it is useful to investigate to what extent our results are robust to a different measure of payouts. Lamont (1998) finds that the log earnings-price ratio *ep* forecasts returns. We find very much the same patterns for the earnings-price ratio as for the dividend-price ratio. The earnings data start in 1946 and are described in Appendix A. The book-to-market ratio is computed from the same earnings and dividend data using the clean-surplus method (Vuolteenaho (2000)).

[Table 3 about here.]

Table 3 shows that the null hypothesis of no structural break in the *ep* ratio is strongly rejected in favor of one or two breaks (first row). The Perron test estimates a 1990 break date in the one-

is not forecastable because any correlated movement in expected returns and expected dividend growth cancels in $d - p$, as shown in Lettau and Ludvigson (2005).

[9] A bootstrap analysis confirms that the small sample p-value (asymptotic p-value) is 1.11% (0.00%) in the one-break case and 0.00% (0.00%) in the two-break case. A second bootstrap exercise shows that the small sample bias in the coefficients is small relative to their magnitude. In the one-break case, the bias is .019 (we estimate .254 when the true coefficient is .235). In the two-break case, the bias is .013 (we estimate .468 when the true value is .455). Detailed results are available upon request.

break case and 1953 and 1994 break dates in the two-break case. These line up almost perfectly with the dp break dates in Table 2. One other often used valuation ratio, the log book-to-market ratio (bm) also displays strong evidence of two breaks with similar break dates in 1953 and 1990 (row 2). Clearly, there is evidence for a permanent or strongly persistent component in all valuation ratios.

Some researchers have argued that there were persistent changes in firms' payout policies in the 1990s and have argued to adjust dividend-price ratios for repurchases (Fama and French (2001), Grullon and Michaely (2002), and Boudoukh, Michaely, Richardson, and Roberts (2004)). First, we find no evidence for a break in the payout ratio $de = d - e$ at the 10% level (row 3 of Table 2). This is consistent with the view that both $\overline{dp}$ and $\overline{ep}$ contain structural breaks. Second, even if there was a break in the $\overline{de}$ ratio and we took the point estimate for $\overline{de}$ in the subsamples, we would find that more than three-fourths of the change in the mean dividend-price ratio comes from a change in the mean earnings-price ratio and less than one-fourth from a change in $\overline{de}$ ($\overline{dp} = \overline{ep} + \overline{de}$). In particular, for our S&P 500 sample from 1946-2004 with a 1991 break date, we find a change in $\overline{dp}$ of -.81, a change in $\overline{ep}$ of -.63, and a change in $\overline{de}$ of -.18.

To further investigate the role of repurchases and the role of a changing composition in CRSP, we consider two additional valuation ratios. First, we consider the CRSP universe without NASDAQ stocks. Arguably, removing NASDAQ stocks goes a long way towards eliminating new economy and non-dividend paying companies that became more prevalent in the 1990s.[10] We compute the dividend-price ratio, dp^{nas}, and the dividend growth for this group. This time series has properties very similar to those of the series with the NASDAQ. The Perron tests in row 5 of Table 3 show a break of -75% in 1992, close to the -86% change in the full sample series in 1991. For the two-break case, the break dates and magnitudes are also very similar: 1954 and 1995 and -35% and -70%.

Second, we use the Boudoukh, Michaely, Richardson, and Roberts (2004) repurchase yield data, available from 1971 onwards, and construct a corrected dividend-price ratio and dividend growth rate series. We label this repurchase-adjusted dividend price series dp^{rep}.[11] The case favored by the data is a three-break case with break dates in 1957, 1973, and 1990 (see last row of Table 3). We show below that these two adjustments do not materially affect our predictability results with the standard dividend-price ratio presented earlier. This leads us to conclude that structural changes in payout policies and/or the composition of firms in the 1990s can only explain a small part of

[10]Fama and French (2001) document that the fraction of non-financial, non-utility firms that paid dividends declined by almost 45% between 1978 and 1999. However, most of that decline is attributable to new firms and to small firms. They write: "The characteristics of dividend payers (large profitable firms) do not change much after 1978." We take this group to be the value-weighted CRSP index without NASDAQ stocks. This series starts to deviate from the full sample series in 1973. We verified that the dividend growth rate of this set of firms did not change in the 1990s. Average dividend growth from 1927-1991 was 5.45%. Average dividend growth from 1992-2004 was 5.46%.

[11]We note that this is just one possible adjustment. The correct adjustment depends on the investor under consideration. Here, an investor's cash flows are adjusted for aggregate repurchases, but not for seasoned equity offerings nor initial public offerings.

the change in the dividend-price ratio.

We first turn to the forecasting regressions with the earnings-price ratio. When we use the earnings-price ratio as a return predictor, we obtain similar results to what we reported for the dividend-price ratio in Table 1. The first row of Table 4 shows that when the unadjusted earnings-price ratio is the independent variable, the slope coefficient is .119. Just as in the dividend-price ratio regressions, this coefficient displays parameter instability among subsamples: The full sample point estimate is lower than the estimates in all subsamples, and the Chow test of no break has a p-value of only .15 (not shown). The next two columns show that this bias is due to averaging over subsamples. Once we use the adjusted $\widetilde{ep}$ ratio, the instability disappears and the full sample point estimate increases to .215 in both the one-break and two-break case. These coefficients are twice the size of the ones obtained with the unadjusted ep ratio and are measured precisely. The regression R^2 almost doubles. The slope coefficient is very similar to the one we found in the first panel of Table 1: $\kappa_r = .235$. One difference from the results in Table 1 is that the adjusted earnings-price ratio also significantly forecasts earnings growth, with a negative sign (not shown).

[Table 4 about here.]

For the unadjusted lagged log book-to-market ratio $bm = b - m$, we find that the predictability coefficient κ_r is only marginally significant. The point estimate is .07, lower than the point estimates in the subsamples 1927-1952 (.26), 1953-1990 (.44), and 1991-2004 (.72), all of which are strongly significant. Again, this downward bias is due to averaging over the break(s). The full sample point estimate increases to .255 with the one-break adjusted $\widetilde{bm}$ series as regressor and to .308 with the two-break adjusted $\widetilde{bm}$ series. The regression R^2 increases from 3% in the first column to 19% in the third column.

The return predictability findings for dp^{nas} and dp^{rep} are also similar to the benchmark dp results. First, using dp^{nas} without break adjustment, we find a point estimate for κ_r of .11. This point estimate is lower than in either subsample (.24 for 1927-1992 and .30 in 1993-2004). Once we use the break-adjusted series, the point estimate more than doubles to .250. Just as for the standard dp ratio, the downward bias comes from averaging over the break. The break-adjusted point estimate is close to the .235 we found for the sample that includes the NASDAQ. We obtain further increases in the point estimate and the R^2 in the two-break case. Second, using dp^{rep}, the full sample return predictability coefficient is .19, higher than the .09 for the standard dp series, but again lower than in either subsample (.25 in 1927-1990 and .53 in 1991-2004). Clearly, adjusting for repurchases improves the forecasting power of the dp ratio. However, adjusting for the breaks is important and further strengthens the case for predictability. In the preferred case of three breaks, the predictability coefficient is .58, three times its unadjusted value. The regression R^2 is also three times higher.

We conclude that the other financial ratios indicate a predictability pattern similar to that of the dividend-price ratio. Without an adjustment for the change in their long-run mean and the relationship between one-year ahead returns and the financial ratios is unstable over time. However, once we filter out the non-stationary component, we find a stable forecasting relationship and a large predictability coefficient. The fact that the results are so similar for earnings and dividend data suggests that an explanation that exclusively rests on changing payout policies misses the most important structural changes in the economy: changes in long-run growth rate or long-run expected returns.

5 Out-Of-Sample Predictability

The in-sample predictability results presented so far used a break adjustment constructed using the entire data sample. In this section we investigate how an investor who forms an adjusted dividend-price ratio in real time fares in predicting out-of-sample returns. We compare the out-of-sample forecasting properties of adjusted dividend-price ratios to the unadjusted series and a random walk model. We find that a real time dividend-price ratio adjustment yields uniformly smaller prediction errors compared to the unadjusted series but slightly larger forecast errors than the random walk model.

Why does the real time prediction errors fail to beat the random walk model? In real time an investor faces two challenges. First, she has to estimate the timing of a break. Second, if she detects a new break, she has to estimate the new mean after the break occurred. If the new break occurred towards the end of the sample that the investor has access to, the new mean can only be estimated using a small number of observations and is subject to significant estimation uncertainty. To investigate which issue is responsible for the deterioration of the out of sample forecasting power, we consider two additional exercises. In the first exercise, the investor predicts out-of-sample using the ex-post adjusted dividend-price ratio series considered in section 3.1. In other words, we endow the investor with information about estimated break dates and means from the entire sample, thus this case is not a pure out-of-sample test. However, it sets an informative benchmark for the analysis of the pure out-of-sample forecasts. In a second exercise, we consider a Hamilton (1989) regime-switching model for the mean of the dividend-price ratio. In this model, the regime means are estimated using data from the entire data sample. The Hamilton model computes two estimates for the break dates (regime switching probabilities): *(i)* "unsmoothed" probabilities estimated using only currently available data and *(ii)* "smoothed" probabilities based on the entire sample. While this case is not a pure out-of-sample forecast, it allows us to study the relative difficulty of estimating the break dates versus estimating the means relative to the pure out-of-sample forecasts and the ex-post adjusted dividend price ratio.

The conclusions from these two exercises are that *(i)* the estimation of the break dates in

real time is not crucial and the resulting prediction errors are smaller than for the random walk model, and *(ii)* that the estimation of the magnitude of the break in the mean dividend-price ratio entails substantial uncertainty, and is ultimately responsible for the failure of the real time out-of-sample predictions to beat the random walk. These findings can explain the lack of out-of-sample predictability documented by Goyal and Welch (2004).

5.1 Real-Time Dividend Price Adjustments

Before presenting the estimation results, we confirm our earlier conclusion (based on ex-post data, see Figure 3) that it is extremely unlikely that the dividend-price ratio sample is drawn from a stationary distribution, based on real-time data only. Figure 4 shows a recursive (i.e., real-time) estimation of the empirical distribution of the log dividend-price ratio. In each year, the investor estimates an AR(1) model for dp, using data up to the current year. She then bootstraps from the available sample to compute the empirical distribution of the log dividend-price ratio. Each year she recomputes the 2.5, 5, 95, and 97.5 percentiles of the bootstrapped distribution (dashed and dotted lines in the figure). The figure also plots the realized dividend-price ratio, in deviation from its recursive sample mean (full line). By 1958, the investor is quite confident that the realized dividend-price ratio is far below the mean; it hits the 2.5 percentile of the empirical distribution. Likewise, in 1994, the observed dividend-price ratio falls in the 2.5% tail of the distribution. Between 1995 and 1999, the investor is almost certain that the observed dividend-price ratio has *not* been drawn from a stationary distribution. Interestingly, these 'crossing' dates are almost identical to the break dates estimated in the previous sections. This shows that the permanent changes in the dividend-price ratio that were identified by the ex-post break tests do not rely on having the benefit of the entire sample through 2004. Even an investor in real-time would have concluded that extreme observations of dividend-price ratios are unlikely to be generated by a stationary process with constant parameters.

[Figure 4 about here.]

Next we construct a real-time adjustment of the dividend-price ratio that can be used in out-of-sample forecasting tests. In each year $T' \leq T$, the investor estimates the Perron structural break test using data available up to year T' using one of the three tests: the sequential sup-F test with a 10% critical value, the BIC criterion, and the LWZ criterion. Given the break dates and corresponding means, the real-time adjustment of the dividend-price ratio is analogous to the adjustment using the entire sample in (8), with the exception that only data up to date T' instead of T are used in the estimation. Denote this corrected ratio $\widehat{dp}^P_{t,T'}, t = 1, ..., T'$. The out-of-sample return forecast for period $T' + 1$ is then computed from a regression of returns r_{t+1} on $\widehat{dp}^P_{t,T'}$ for

$t = 1, ..., T' - 1$ and the currently observed adjusted dividend-price ratio $\widehat{dp}^{P}_{t,T'}$.[12]

To gauge the relative difficulties of estimating the timing of a break versus the estimation of the means, we also use a "pseudo" real-time adjustment based on a Hamilton (1989) regime-switching model in real-time. The dividend-price ratio is assumed to follow an AR(1) with different means in either two regimes (one-break case) or three regimes (two-break case). The top panel of Figure 5 shows the smoothed (based on the entire sample) and unsmoothed (real-time) estimates of the probability that the dividend-price ratio is drawn from the low regime, when two regimes are considered. Using real-time data, the investor puts non-zero probability on a shift to the low $d-p$ regime starting in early 1990. By 1995, she is more than 50% certain that the shift occurred. The difference between smoothed probabilities and unsmoothed probabilities is that the investor is assigning a higher likelihood to the new, low dividend-price regime earlier if she has access to the entire data set. Note that the $\overline{dp}$ estimates for the two regimes coincide with our ex-post estimates because they use the entire data sample. When she considers three regimes instead, the investor increases the real-time probability of a switch from the high to the middle dp regime in 1954. By 1960 she is more than 50% certain that the first shift occurred (see middle panel of Figure 5). In 1990, she starts to attribute probability mass to the low-p regime, and by 1996, she is more than 50% confident that the economy left the middle-dp regime for the low-dp regime. As in the two-regime case, the smoothed probabilities indicate that knowing the entire sample yields a faster transition to the new regime.

Not surprisingly, the regimes in the Hamilton model are estimated to be very persistent. The probability of remaining in the current regime is above 0.9 in all cases. In other words, the regimes identified by the Hamilton estimation are close to permanent structural breaks as specified in the Perron-model.

[Figure 5 about here.]

5.2 Out-Of-Sample Forecast Errors

We follow the approach taken by Goyal and Welch (2003) and predict one-year ahead returns with the lagged dividend-price ratio. The first forecasting regression uses 20 years of data, so that the first forecasted return is the one in 1946. For all future years we use expanding windows and compare three different forecasters. The first is the current sample mean return implied by the "naive" random walk model. The second is the standard unadjusted dividend-price ratio dp. The third is the break-adjusted dividend-price ratio. To explain the results, we use several alternatives for the break-adjusted series.

[12] This procedure implies that there are no breaks between T' and $T' + 1$, consistent with our assumption that the break is unpredictable.

[Table 5 about here.]

Panel 1 of Table 5 reports the mean absolute forecast error (MAE) and the root mean-squared forecasting error (RMSE). Comparing the second to the first row, we confirm the result of Goyal and Welch (2003): The random walk model has superior out-of-sample properties compared to the standard dividend-price ratio specification. The latter's prediction errors are almost 1% per year higher.

The third through fifth rows report the real-time adjusted dividend-price ratio $\widehat{dp}^P$ obtained from the Perron procedure. For all three criteria for dating the breaks, sup-F, BIC and LWZ, the prediction errors of the adjusted series are lower than those obtained from using the unadjusted series (row 2) but the out-of-sample prediction errors are larger than for the random walk model. The sequential sup-F test has the lowest mean absolute error while the BIC criterion delivers the lowest root mean squared error.

We contrast these results with three "pseudo" out-of-sample cases to investigate why the out-of-sample forecast errors are higher than those for the random walk model. First, we use the dividend-price ratio adjustment based on the Perron break test estimated in the entire sample $\widetilde{dp}$. Second, we construct adjusted dividend-price ratios based on the Hamilton model $\widehat{dp}_t^H$ by subtracting the average of the means of each regime weighted by the time-t probability of each regime from dp_t. Figure 6 plots the mean dividend-price ratios that result from probability-weighting the regime mean estimates. The weights are either the unsmoothed (solid line) or the smoothed probabilities (dashed line). It also plots the ex-post mean (dotted line). Naturally, the mean dividend-price ratio based on smoothed probabilities tracks the ex-post breaks somewhat faster than the unsmoothed mean.

[Figure 6 about here.]

Rows six and seven in Panel 1 of Table 5 show that the use of the ex-post adjusted dividend-price ratios $\widetilde{dp}$ substantially reduces the forecasting error. The MAE and RMSE of the dividend-price ratio adjusted for a single break in 1991 are lower than those for the random walk model. The out-of-sample forecasting power of the dividend-price ratio adjusted for two breaks in 1954 and 1994 is dramatically improved compared to the unadjusted ratio and to the random walk model. The RMSE and MAE are reduced by 12-15% compared to the random walk model.

The pseudo out-of-sample forecast errors based on the Hamilton correction with smoothed probabilities are similar but slightly lower than those in the ex-post Perron correction. Since both cases use the entire sample to estimate the regime means and probabilities it is not surprising that the out-of-sample results are comparable. The Hamilton model allows for a smoother adjustment to a new regime compared to the Perron model, which might account for the lower forecast error.

For the Hamilton regime switching model we also compute pseudo out-of-sample forecast errors based on unsmoothed probabilities. Recall that these probabilities are estimated using only currently available information but estimates of regime means are based on information of the entire sample. Thus, this case enables us to assess the relative difficulty of real-time estimation of the timing of breaks versus real-time estimation of the level of regimes. Interestingly, the out-of-sample prediction errors of the case with unsmoothed probabilities are only slightly higher than those with ex-post Perron adjustment and smoothed probabilities. In the one-break example, the RMSE for the unsmoothed Hamilton adjustment is 15.90% compared to 15.58% for the ex-post Perron adjustment and 15.43% for the smoothed Hamilton adjustment. In contrast, the forecast errors for the pure OOS adjustments are significantly higher; between 16.46% and 16.80% depending on the selection criterion.

Where does the difference in forecast errors come from? In the Hamilton adjustment with unsmoothed probabilities, the investor estimates the probability of each regime in real time but has access to the entire time series to estimate the mean dp ratio in the different regimes.[13] In the ex-post Perron and smoothed Hamilton procedures, the investor not only estimates break dates in real-time, but also the long-run mean in the current regime. Our results suggest that the poor real-time out-of-sample forecasts are to a larger extent due to the uncertainty about the magnitudes of breaks rather than uncertainty about the exact timing of a break. Estimating a new long-run mean based on a few data points incurs a lot of measurement error. For example, if the Perron investor in 2001 detects a break in 1995, she only has six data points to estimate the new $\overline{dp}^P$. The difficulty in estimating this mean is what accounts for the increase in prediction errors.

6 Long-Horizon Predictability

An important component of the empirical work on return predictability uses long-horizon regressions. In this section we provide a framework for analyzing long-horizon predictability. We use the ex-post dividend-price ratio as the predictor and derive theoretical restrictions that link return and dividend growth to lagged dividend-price ratios at different horizons. Rather than estimating regressions for various horizons separately, the advantage of this approach is that fewer parameters have to be estimated. Moreover, the restrictions impose that the estimates across horizons are consistent with each other.

[13] A satisfactory resolution that avoids any look-ahead bias in the Hamilton model would be to conduct a full Bayesian analysis in which the investor estimates the number of regimes, the regime switching dates, and their associated long-run means based on real-time available data and prior information. Pettenuzzo and Timmermann (2005) have worked out such an estimation in the context of an asset allocation problem.

Recall the return and dividend growth predictability equations (1) and (2)

$$\tilde{r}_{t+1} = \kappa_r \widetilde{dp}_t + \tau^r_{t+1} \tag{9}$$

$$\Delta\tilde{d}_{t+1} = \kappa_d \widetilde{dp}_t + \tau^d_{t+1}, \tag{10}$$

where variables with a tilde are appropriately demeaned and stationary. Subtracting (10) from (9) and using the log-linear approximation for log returns $\tilde{r}_{t+1} = \widetilde{dp}_t - \rho\widetilde{dp}_{t+1} + \Delta\tilde{d}_{t+1}$ yields the implied AR(1) process for the dividend-price ratio in (11):

$$\widetilde{dp}_{t+1} = \phi\widetilde{dp}_t + \tau^{dp}_{t+1} \text{ , where} \tag{11}$$

$$1 - \rho\phi = \kappa_r - \kappa_d \tag{12}$$

where the innovations are linked by $\rho\tau^{dp}_{t+1} = \tau^d_{t+1} - \tau^r_{t+1}$. The model imposes a non-linear present value restriction (12) on the predictability coefficients κ_r and κ_d. Because $\rho < 1$ and stationarity implies $|\phi| < 1$, $\kappa_r - \kappa_d$ must be positive. This is another way of saying that either returns ($\kappa_r \neq 0$) or dividend growth ($\kappa_d \neq 0$) have to be forecastable (or both). Most researchers work with equations (9) and (11); we work with (9) and (10) instead because long-horizon restrictions are more easily derived in this case. Iterating forward on equations (9) and (10), we obtain the (annualized) H-period dividend growth and return forecasting equations

$$\frac{1}{H}\sum_{j=1}^{H} \tilde{r}_{t+j} = \kappa_r(H)\, \widetilde{dp}_t + \tau^r_{t,t+H} \tag{13}$$

$$\frac{1}{H}\sum_{j=1}^{H} \Delta\tilde{d}_{t+j} = \kappa_d(H)\, \widetilde{dp}_t + \tau^d_{t,t+H}, \tag{14}$$

where

$$\kappa_d(H) = \kappa_d \frac{1}{H}\left(\frac{1-\phi^H}{1-\phi}\right) \tag{15}$$

$$\kappa_r(H) = \kappa_r \frac{1}{H}\left(\frac{1-\phi^H}{1-\phi}\right). \tag{16}$$

Let N be the number of horizons $H > 1$. Then the joint system of one-year ahead and H-year ahead predictability regressions for returns and dividend growth contains $2 + 2N$ equations but only two free parameters (κ_d, κ_r). The parameter ϕ is implied by the present value constraint (12). The typical approach in the literature is to estimate univariate long-horizon return predictability equations without imposing these restrictions. Instead, we estimate the entire system of long-horizon return and dividend growth regressions jointly. This estimation procedure not only takes the high correlation of return regression coefficients at different horizons, pointed out by

Boudoukh, Richardson, and Whitelaw (2005), explicitly into account. It also takes the present value relationship (12) explicitly into account.

The estimation routine we describe below finds parameters to match the entire 'term structure' of univariate predictability coefficients. Figure 7 illustrates this procedure for $H = \{1, 3, 5, 7, 10\}$. The top panel plots the univariate predictability coefficients obtained from standard OLS regressions of H-year ahead returns and dividend growth on the unadjusted dividend-price ratio dp, as well the predictability coefficients $(\kappa_d(H), \kappa_r(H))$ implied by the joint estimation of the system of equations. To match the pattern of the ten OLS predictability coefficients, the optimization routine chooses a value for κ_r above the value from the one-year ahead univariate regression.

[Figure 7 about here.]

We start by estimating the one-period ahead equations for returns and dividend growth (9-10). These are the same equations we estimated in Section 3.3, but the additional restriction provides two new insights. First, the estimation delivers a value for the autocorrelation coefficient of the dividend-price ratio ϕ because we impose the present value constraint (12). Second, we use the break adjusted series $\widetilde{dp}$ from Section 3 in the estimation.

Because return and dividend growth series were demeaned by their sample averages, the previous sections implicitly assumed a break in $\overline{dp}$ without associated break in $\overline{r}$ or $\overline{d}$. The model tells us that a break in $\overline{dp}$ must be associated with a break in either $\overline{r}$ or $\overline{d}$, or both. First, we assume that $\overline{d}$ is constant and focus on changes in the expected returns, consistent with the evidence on breaks in the equity premium in Pastor and Stambaugh (2001). The change in $\bar{r}$ implied by the change in $\overline{dp}$ can be inferred from $\bar{r}_t = (1 + \overline{d}) \exp(\overline{dp}_t) + \overline{d}$. The top panel of Table 6 shows how large the change in $\overline{r}$ is (row 2) corresponding to the change in the mean *dividend-price* ratio in the data (row 1). The left panel is for the one-break case, the right panel for the two-break case. The observed change in $\overline{dp}_t$ implies a decline in mean expected returns of 2.6% in 1991 or a dual decline of 1.7% in 1954 and 2% in 1994, assuming long-run dividend growth did not change. Alternatively, it can stem from an increase in long-run dividend growth of 2.5% in 1991 or a dual increase of 1.6% in 1954 and 2% in 1994, when mean expected returns are held constant. In the results reported below, we choose to correct $\bar{r}_t$, but this choice turns out to be unimportant for the point estimates.[14] The second panel reports the change in the mean earnings-price ratio in the various subsamples, as well as the implied change in the long-run mean return or long-run mean dividend growth rate. Appendix A describes how the latter two are computed. We find that the change in the mean return of -2.47% that accounts for the change in the earnings-price ratio $\overline{ep}$ (bottom panel) is very similar to the -2.60% change that accounted for the change in the

[14]The reason is that returns and dividend growth are very volatile, compared to the change in their mean implied by the change in $\overline{dp}$. For the same reason, the results reported below are virtually identical if we assume that the break takes place in $\overline{d}$ instead. This validates the results in Section 3.3.

dividend-price ratio $\overline{dp}$ (top panel). The same is true when the long-run growth rate does all the adjustment.

[Table 6 about here.]

Panel A of Table 7 reports the estimation of the one-year ahead system. Row 1 uses the unadjusted dp series, whereas rows 2 and 3 use the adjusted series $\widetilde{dp}$ for the one-break and two-break case respectively. The GMM estimation uses the OLS normal conditions to estimate κ_d and κ_r. Therefore, the point estimates are identical to the ones reported in Table 1. Three differences are worth pointing out. First, the adjustment in $\bar{r}$ delivers slightly lower standard errors for κ_r in rows 2 and 3. Second, as foreshadowed by Table 2, the point estimates for ϕ are substantially lower when we use the adjusted dividend-price ratio: .81 and .69 in rows 2 and 3 compared to .95 in row 1. Third, the first number in the last column reports the violation of the present value constraint (12) by the univariate (OLS) coefficient estimates $\kappa_{d,ols}$ and $\kappa_{r,ols}$, expressed in the same units as κ_d and κ_r. Row 1 shows that using the unadjusted dp ratio leads to violations of the present value constraint. They are half the size of the estimated κ_r. Yet, when we use the adjusted dp ratio, constraint (12) is satisfied.

[Table 7 about here.]

Next, we estimate the one-period ahead equations for returns and dividend growth (9-10) jointly with the long-horizon regressions (13-14). We select a small number ($N = 2$) of long-horizon moments, corresponding to $H = \{1, 3, 5\}$. The joint system of one-year, three-year, and five-year ahead predictability regressions for returns and dividend growth contains $2 + 2N = 6$ equations and $2N = 4$ restrictions. Panel B of Table 7 reports the results.

The point estimates for (κ_d, κ_d) are similar to those obtained from the one-year ahead system in panel A. Row 4, which uses the unadjusted dp ratio, fails to find evidence for return predictability or dividend growth predictability at the 5% level. The point estimate falls from .094 in panel A to .068 in panel B. Using long-horizon information makes the case for return predictability weaker when the unadjusted dp series is used. Furthermore, the estimate for $\phi = .99$ and its standard error indicate that we cannot reject the null hypothesis of a unit root in the dividend-price ratio.[15]

Results using adjusted series reported in rows 5 and 6 are quite different. Once we use the adjusted ratio $\widetilde{dp}$, we find strong evidence for return predictability. The point estimates remain large: .210 in the one-break case and .409 in the two-break case. Moreover, the asymptotic standard errors on κ_r are reduced. The reason is that we use restrictions of the term structure of predictability coefficients that cannot be uncovered by estimating the long-horizon moments

[15]The standard error of ϕ is implied by the estimates for κ_d and κ_r through (12) and computed using the delta method.

in isolation; imposing these constraints improves the inference on κ_r and κ_d. Put differently, the univariate OLS long-horizon coefficients violate the present value constraint. The first number in the last column reports the average violation across the three constraints (the RMSE); it is 19% in row 1. This violation is lower in rows 5 and 6. The second number in the last column reports the average moment violation (RMSE) and measures the degree to which the imposed restrictions are satisfied. In row 5, the average moment violation is only 8.5%, less than half as big as in row 4.

The middle and bottom panels of Figure 7 graphically illustrate the long-horizon estimation results with the adjusted dividend-price ratio for the larger $\{1, 3, 5, 7, 10\}$-year system. First, the one-period coefficients are very similar to the ones we reported for the $\{1, 3, 5\}$-year system. Second, the two panels show that when the adjusted dividend-price ratio is used, the pattern of GMM long-horizon regression slopes generated by our model lines up almost perfectly with the univariate OLS regression slopes. Put differently, the normal conditions for the long-horizon moments are satisfied at the model-implied long-horizon predictability coefficients. This contrasts with the first panel, which is for the unadjusted *dp* ratio. There, the OLS coefficients are quite different from the GMM estimates. The correction supports both our specification and our main argument.

The results for the $\{1, 3, 5\}$-year and $\{1, 3, 5, 10\}$-year systems are representative of the results we found for different numbers of long-horizon moments and choices of horizons. Imposing long-horizon information confirms the results of the earlier sections: Returns are predictable by the dividend-price ratio, once its non-stationary component is removed.

7 Monte Carlo Simulations

In this section, we provide further evidence that the model in (9-11) captures the moments of the data well. We use a Monte Carlo exercise to show *(i)* that this model replicates the failures that are found in the in-sample and out-of-sample predictability literature using the unadjusted *dp* ratio, *(ii)* that it matches the moments once the dividend-price ratio is properly adjusted for and after taking into account small sample inference issues.

For the Monte Carlo exercise, we specify a structural model for the joint behavior of expected returns, expected dividend growth, and dividend growth innovations. In Appendix D, we derive the regression residuals $\tau = (\tau^d, \tau^r, \tau^{dp})$ as functions of the structural innovations and show how to identify the structural parameters from the parameters of the vector error correction model (VECM) in equations (9-11). We back out the structural parameters from the previously reported estimates of the VECM parameters and simulate the structural model generating 10,000 time series for returns, dividend growth and the dividend-price ratio of length $T = 78$, the same length as the data. We then run univariate predictability equations on the model-generated data and compare the parameter estimates to the true predictability coefficients and to the data. The Monte Carlo exercise also serves as a way to investigate the small sample properties of the regression coefficients.

Appendix E describes the algorithm in detail.

We start by studying the properties of univariate return predictability regressions. The first row of Table 8 reports the 'true' predictability coefficients for the one-year, three-year, and five-year horizon equations, as well as the theoretical regression R^2.[16] In the top panel, the true parameters come from an estimation of the VECM that assumes a break in *dp* in 1991; in the bottom panel we specify two breaks in 1954 and 1994.

The second and fourth rows report the same coefficients and R^2 estimated on simulated data. This simulation ignores the presence of the break(s). Row 2 uses the unadjusted *dp* ratio, and row 4 the adjusted $\widetilde{dp}$ to predict returns. The second row shows that the model with the unadjusted *dp* ratio fails to detect the return predictability that is present in the data. The simulation-based estimate of κ_r and the regression R^2 are too small. The regression R^2 does not increase enough with the horizon. The one-year ahead coefficient is not significant when we use the small-sample standard error (fourth column). Moreover, this failure of the model matches the failure in the data. Row 3 shows the results of the same univariate regressions in the data using the unadjusted *dp* ratio. The slope coefficients, standard errors, and R^2 line up closely with the results from the Monte Carlo simulation without break adjustment.

In row 4, we adjust each Monte Carlo series for a break in 1991 (top panel) or two breaks in 1954 and 1994 (bottom panel). The predictability coefficients at all horizons now line up closely with their true values. The model with adjusted *dp* ratio recovers the true predictability pattern of row 1. Moreover, the predictability coefficients, their standard errors, and the regression R^2 from the model in row 4 match the ones from regressions of observed returns on observed adjusted dividend-price ratios (row 5). Results for dividend growth regressions are not reported but the simulations recover the lack of predictability in the true κ_d coefficients implied by the VECM.

[Table 8 about here.]

Comparing row 4 to row 1, we notice that there is some small sample bias. In line with the findings of Stambaugh (1999), the estimate for κ_r is upward biased. In the first panel, the true value of the slope coefficient in the one-year ahead return regression is .222 versus .264 in simulation, a bias of .042. At the five-year horizon, the upward bias is only .016. Likewise, the R^2 of the regression is slightly upward biased in the simulation: 10.5% versus 8.9% at one-year horizon and 28% versus 25% at the five-year horizon. In the two-break case reported in the second panel, the bias is smaller. The one-year ahead coefficient estimate is .483 versus the true value of .456. The upward bias disappears at the five-year horizon. Overall, the bias is small relative to the magnitude of the coefficients, and therefore does not affect our conclusions. Finally, the small sample standard

[16] As Campbell, Lo, and MacKinlay (1997) point out, there is no closed form solution for the long-horizon R^2. We approximate it by simulating the structural model for 100,000 periods.

errors, averaged across Monte Carlo simulations of the same length of the data, are very similar to the asymptotic standard errors from the data (middle columns).

Panels 2 and 3 of Table 5 report the out-of-sample prediction errors from a second Monte Carlo exercise. We simulate the structural model under the null hypothesis that the data generating process has one break in 1991 (panel 2) or has two breaks in 1954 and 1994 (panel 3). We compare the same three out-of-sample forecasting exercises as in the data (panel 1). In addition, we look at the forecast errors when we only correct for the second break and not the first one. The 'true' parameters in panels 2 and 3 are the same and were obtained from the VECM parameters estimated under the assumption of two breaks in 1954 and 1994 for the period 1947-2004, the same forecasting period as in panel 1. The details are shown in Appendix E. The Monte Carlo exercise regenerates the pattern we found in the data. When the unadjusted dividend-price ratio is used as a forecasting variable, the out-of-sample prediction errors are large and close to the random walk errors. On the other hand, when we implement the one-break or two-break adjustment in the model, the simulated data generate substantially lower prediction errors, mimicking the improvement in the data.

The simple model (9-11) replicates the patterns of univariate one-year ahead and long-horizon regression results in-sample as well as one-year ahead out-of-sample prediction errors found in the data. In particular, it regenerates *(i)* the failures of using the unadjusted dp ratio as a predictor, and *(ii)* the successes of using the adjusted $\widetilde{dp}$ ratio.

8 Conclusion

The macroeconomics literature has recently turned to models with persistent changes in fundamentals to explain the dramatic change in valuation ratios in the bull market of the 1990s. Most such models imply a persistent decline in expected returns or a persistent increase in the dividend growth rate. In this paper we argue that either of such changes leads to a persistent decline in the mean of financial ratios. Such changes in the mean of valuation ratios have important effects on estimation and inference of return forecasting regressions. We consider various econometric techniques to detect shifts in the mean of price ratios and suggest a simple procedure to extract their stationary component. The adjusted price ratios robustly forecast returns in sample. At the same time, we show that shifts in the steady-state expected returns and growth rate of fundamentals are responsible for the instability of the return forecasting relation. Out-of-sample return predictions based on the adjusted price ratios improve relative to those based on the unadjusted ratios. However, estimation uncertainty about the magnitude of the break, rather than the break date, is ultimately what prevents a real-time investor to profit.

References

Amihud, Y., C. M. Hurvich, and Y. Wang (2005): "Hypothesis Testing in Predictive Regressions," Unpublished paper.

Amihud, Y., and C. M. Hurwich (2004): "Predictive Regressions: A Reduced-Bias Estimation Method," *Financial and Quantitative Analysis*, 39, 813–841.

Ang, A., and G. Bekaert (2006): "Stock Return Predictability: Is it There?," *Forthcoming Review of Financial Studies*.

Berk, J. B., R. C. Green, and V. Naik (1999): "Optimal Investment, Growth Options and Security Returns," *Journal of Finance*, 54, 1153–1607.

Bossaerts, P., and P. Hillion (1999): "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?," *Review of Financial Studies*, 12, 405–428.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts (2004): "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing," NBER Working Paper No. 10651.

Boudoukh, J., M. Richardson, and R. F. Whitelaw (2005): "The Myth of Long-Horizon Predictability," New York University Working Paper.

Calvet, L., M. Gonzalez-Eiras, and P. Sadini (2003): "Financial Innovation, Market Participation and Asset Prices," NBER Working Paper No. 9840.

Campbell, J. Y. (1991): "A Variance Decomposition for Stock Returns," *Economic Journal*, 101, 157–179.

Campbell, J. Y., A. W. Lo, and C. MacKinlay (1997): *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Campbell, J. Y., and R. J. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

Campbell, J. Y., and S. Thompson (2005): "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," NBER Working Paper No. 11468.

Campbell, J. Y., and M. Yogo (2002): "Efficient Tests of Stock Return Predictability," Unpublished paper, Harvard University.

Cochrane, J. H. (1991): "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies*, 5(2), 243–280.

——— (2006): "The Dog That Did Not Bark: A Defense of Return Predictability," Unpublished paper, University of Chicago Graduate School of Business.

Eliasz, P. (2005): "Optimal Median Unbiased Estimation of Coefficients on Highly Persistent Regressors," Unpublished paper, Department of Economcis, Princeton University.

Fama, E. F., and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–27.

——— (2001): "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay," *Journal of Financial Economics*, 60(1), 3–43.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003): "Spurious regressions in financial economics?," *Journal of Finance*, 58(4), 1393–1413.

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

Goyal, A., and I. Welch (2003): "Predicting the Equity Premium with Dividend Ratios," *Management Science*, 49(5), 639–654.

——— (2004): "A Comprehensive Look at the Empirical Performance of the Equity Premium Prediction," Working Paper Brown University.

Grullon, G., and R. Michaely (2002): "Dividends, Share Repurchases, and the Substitution Hypothesis," *Journal of Finance*, 57(4), 1649–1684.

Guvenen, M. F. (2003): "A Parsimonious Macroeconomic Model for Asset Pricing: Habit Formation or Cross-sectional Heterogeneity?," Rochester Center for Economic Research Working Paper No. 499.

Hamilton, J. D. (1989): "A New Approach to the Analysis of Nonstationary Returns and the Business Cycle," *Econometrica*, 57, 357–384.

Hansen, L. P., and R. J. Hodrick (1980): "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88(5), 829–853.

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

Jovanovic, B., and P. L. Rousseau (2003): "Specific Capital and the Division of Rents," Unpublished paper, New York University.

Krueger, D., and F. Perri (2005): "Does Income Inequality Lead To Consumption Inequality? Some Empirical Findings and a Theoretical Explanation," *Review of Economic Studies,* forthcoming.

Lamont, O. (1998): "Earnings and Expected Returns," *Journal of Finance*, 53, 1563–87.

Lettau, M., and S. C. Ludvigson (2001): "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance*, 56(3), 815–849.

——— (2005): "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76, 583–626.

Lettau, M., S. C. Ludvigson, and J. A. Wachter (2004): "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," NBER Working Paper No. 10270.

Lewellen, J. W. (1999): "The Time Series Relations Among Expected Return, Risk, and book-to-market," *Journal of Financial Economics*, 54, 5–53.

——— (2004): "Predicting Returns With Financial Ratios," *Journal of Financial Economics*, 74(2), 209–235.

Liu, J., S. Wu, and J. Zidek (1997): "On Segmented Multivariate Regressions," *Statistica Sinica*, 7, 497–525.

Lustig, H., and S. Van Nieuwerburgh (2005a): "Housing Collateral, Consumption Insurance and Risk Premia: An Empirical Perspective," *Journal of Finance*, 60(3), 1167–1219.

——— (2005b): "How Much Does Household Collateral Constrain Regional Risk Sharing," Unpublished manuscript, University of California at Los Angeles and New York University.

——— (2005c): "Quantitative Asset Pricing Implications of Housing Collateral Constraints," Unpublished manuscript, University of California at Los Angeles and New York University.

Mankiw, N. G., and M. Shapiro (1986): "Risk and Return: Consumption Beta versus Market Beta," *Review of Economics and Statistics*, 68, 452–59.

McGrattan, E. R., and E. C. Prescott (2005): "Taxes, Regulations, and the Value of U.S. and U.K. Corporations," *Review of Economic Studies*, 72.

Menzly, L., T. Santos, and P. Veronesi (2004): "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

Nelson, C. C., and M. J. Kim (1993): "Predictable Stock Returns: The Role of Small Sample Bias," *Journal of Finance*, 43, 641–661.

Pastor, L., and R. F. Stambaugh (2001): "The Equity Premium and Structural Breaks," *Journal of Finance*, 56(4), 1207–1239.

Paye, B. S., and A. Timmermann (2005): "Instability of Return Prediction Models," Forthcoming *Journal of Empirical Finance*.

Perron, P. (1989): "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57(6), 1361–1401.

Pettenuzzo, D., and A. Timmermann (2005): "Predictability of Stock Returns and Asset Allocation under Structural Breaks," Unpublished paper, Bocconi University and University of California San Diego.

Piazzesi, M., M. Schneider, and S. Tuzel (2006): "Housing, Consumption, and Asset Pricing," Forthcoming *Journal of Financial Economics*.

Quadrini, V., and U. Jermann (2003): "Stock Market Boom and the Productivity Gains of the 1990s," Unpublished paper, University of Southern California.

Rozeff, M. S. (1984): "Dividend Yields are Equity Risk Premia," *Journal of Portfolio Managment*, 49, 141–160.

Stambaugh, R. F. (1986): "Bias in Regressions with Lagged Stochastic Regressors," Working Paper University of Chicago.

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

Timmermann, A. (2001): "Structural Breaks, Incomplete Information, and Stock Prices," *Journal of Business and Economic Statistics*, 19, 299–314.

Torous, W., R. Volkanov, and S. Yan (2004): "On Predicting Returns with nearly Integrated Explanatory Variables," *Journal of Business*, 77, 937–966.

Valkanov, R. (2003): "Long-Horizon Regressions: Theoretical Results and Applications," *Journal of Financial Economics*, 68, 201–232.

Viceira, L. (1996): "Testing For Structural Change in the Predictability of Asset Returns," Unpublished manuscript, Harvard University.

Vissing-Jorgensen, A. (2002): "Limited Asset Market Participation and Intertemporal Substitution," *Journal of Political Economy*, 110(4), 825–853.

Vuolteenaho, T. (2000): "Understanding the Aggregate Book-Market Ratio and its Implications to Current Equity-Premium Expectations," Unpublished paper, Harvard University.

A Data Description

We use annual end-of-year data from 1926-2004 from CRSP on the value weighted market return (NASDAQ, NYSE, AMEX), with and without dividend capitalization. Net return including dividends are denoted R_t and returns excluding dividends are R^{ex}. The dividend yield D_{t+1}/P_t as the difference of capitalized and uncapitalized series, the dividend level is the product of the dividend yield D_{t+1}/P_t and the price P_t of a portfolio that does not reinvest dividends. The dividend-price ratio is defined as D_{t+1}/P_{t+1} and dividend growth as the change in the dividend level $(D_{t+1} - D_t)/D_t$ We then define log returns $r_{t+1} = \log(1 + R_{t+1})$, log dividend growth $\Delta d_{t+1} = \log(D_{t+1}/D_t)$, and the log dividend-price ratio $dp_t \equiv d_t - p_t = \log(D_t/P_t)$.

Data used in Section 4 is derived from monthly S&P 500 dividend-price and price-earnings ratios and end-of-month composite price index data from Haver. The data are from January 1946-December 2004. Book-to-market ratios are from Vuolteenaho (2000) for 1927-1999. The observations for 1999-2004 are constructed using the clean-surplus ratio method with earnings and dividend data from Haver.

The second panel of Table 6 reports the subsample and full sample means of the log earnings-price ratio ($\overline{ep}$), dividend-earnings ratio (payout rate), and the implied mean changes in long-run expected returns $\overline{r}$ or expected dividend growth $\overline{d}$. The implied change in $\overline{r}$, holding $\overline{d} = 0.0709$ fixed, is computed as $\overline{r} = \overline{d} + (1+\overline{d})(\overline{D/E})(\overline{E/P})$. Likewise, the implied change in $\overline{d}$, holding $\overline{r} = .1098$ fixed, is computed as: $\overline{d} = \left(\overline{r} - (\overline{D/E})(\overline{E/P})\right) / \left(1 + (\overline{D/E})(\overline{E/P})\right)$.

B A Modified Log-Approximation

In this section we extend the Campbell-Shiller log-linear approximation to allow for time-varying steady-state growth rates and returns. This framework is a useful organizing principle for the empirical analysis in the main text, but most of our specifications do not impose the approximation. Our results do not depend on the accuracy of the approximation, but the framework helps to understand the intuition and implications.

The gross return of an asset is defined as

$$R_{t+1} \equiv \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_{t+1}}{D_t} \frac{1 + P_{t+1}/D_{t+1}}{P_t/D_t} \tag{17}$$

As of period t, the steady-state (gross) growth rate of dividends is $\overline{D}_t$ and steady-state expected (gross) returns are $\overline{R}_t$, implying a steady-state for the price-dividend ratio in levels ($\overline{PD}_t$) and logs ($\overline{pd}_t$):

$$\overline{PD}_t = \frac{\overline{D}_t}{\overline{R}_t - \overline{D}_t} \implies \overline{pd}_t = \overline{d}_t - \log(\exp(\overline{r}_t) - \exp(\overline{d}_t)). \tag{18}$$

Instead of presuming that the steady-state growth rates and expected returns are constant, we allow for the possibility that the steady-state may change over time. The only requirement that we impose on the steady-state log returns and log growth rates is that they are martingales; $E_t[\overline{r}_{t+j}] = \overline{r}_t$, $E_t[\overline{d}_{t+j}] = \overline{d}_t$. In other words, the steady-state is constant in expectation only. We also assume that the steady-state log P/D ratio is a martingale; $E_t[\overline{pd}_{t+j}] = \overline{pd}_t$. Although in general these assumption can be inconsistent with each other since the log P/D ratio is a nonlinear function of $\overline{r}_t$ and $\overline{d}_t$, we show in appendix C that the martingale assumptions are satisfied to a very good approximation for reasonable break processes. Log-linearizing (17) around the steady-state in $t+1$ and expressing the variables in deviations from steady-state yields

$$\begin{aligned} pd_t - \overline{pd}_t = (\Delta d_{t+1} - \overline{d}_t) - (r_{t+1} - \overline{r}_t) + \rho_{t+1}(pd_{t+1} - \overline{pd}_{t+1}) \\ + \Delta\overline{pd}_{t+1} + \Delta\overline{r}_{t+1} - \Delta\overline{d}_{t+1} \end{aligned} \tag{19}$$

where $pd_t = p_t - d_t$ and $\rho_{t+1} = \frac{\exp(\overline{pd}_{t+1})}{1+\exp(\overline{pd}_{t+1})}$. The last two assumptions state that ρ_t is a martingale, and that deviations from the mean price-dividend ratio are uncorrelated with ρ; $E_t[\rho_{t+j}] = \rho_t$ and $E_t\left[\rho_{t+j}(pd_{t+j} - \overline{pd}_t)\right] = 0$. Given our assumptions, we can take conditional expectations and solve the expectational difference equation for pd_t:

$$pd_t - \overline{pd}_t = E_t[\Delta d_{t+1} - \overline{d}_t] - E_t[r_{t+1} - \overline{r}_t] + \rho_t E_t[pd_{t+1} - \overline{pd}_{t+1}] \tag{20}$$

$$= \sum_{j=1}^{\infty} \rho_t^{j-1} E_t[\Delta d_{t+j} - \overline{d}_t] - E_t[r_{t+j} - \overline{r}_t] \tag{21}$$

The log price dividend ratio is the sum of the steady-state price dividend ratio and the discounted sum of expected dividend growth minus expected returns in excess of steady-state growth and returns:

$$pd_t = \overline{pd}_t + \sum_{j=1}^{\infty} \rho_t^{j-1} E_t[\Delta \tilde{d}_{t+j}] - E_t[\tilde{r}_{t+j}], \tag{22}$$

where $\Delta\tilde{d}_{t+j} = \Delta d_{t+j} - \overline{d}_t$ and $\tilde{r}_{t+j} = r_{t+j} - \overline{r}_t$. The expression in the main text for dp_t follows from (22) and $dp_t = -pd_t$.

C An Asset Pricing Model with Steady-State Shifts

This appendix solves a fully specified asset pricing model where the steady-state log dividend-price ratio follows an approximate, bounded martingale. Breaks in the steady-state dividend-price ratio originate in breaks in the steady-state mean dividend growth rate. Evidence for such breaks in US data is provided in Timmermann (2001).

Log dividend growth is $\Delta d_{t+1} = g + q_t + \varepsilon_{t+1}$. The unconditional average consumption growth rate is g, q_t is the stochastic component of the long-run mean, and ε_{t+1} is a temporary consumption growth shock, $\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Let $Q_t = (1 - A\exp(q_t))^{-1}$ for a constant A given below. We assume that Q_t is a martingale and show below that this implies that log dividend growth is also a martingale.

$$Q_{t+1} = Q_t + \eta_{t+1}$$

The innovation to the random walk is independent from ε_{t+1} and has the following distribution:

$$\eta_{t+1} = \begin{cases} 0 & \text{w.p. } \pi \\ w_{t+1} & \text{w.p. } 1-\pi \text{ and } w_{t+1} \sim \mathcal{N}(0, \sigma_{w,t+1}^2). \end{cases}$$

With probability π the innovation w_{t+1} is 0, where π is close to 1 to capture that a break is a rare occurrence. With the complementary probability $1-\pi$ there is a break with normally distributed break size. The innovation standard deviation of w_{t+1} is time-varying. It takes on the following form

$$\sigma_{w,t+1} = \bar{\sigma}_w(1 - \exp(\kappa(1 - Q_t))). \tag{23}$$

This implies that for $Q_t \gg 1$, $\sigma_{w,t+1} \approx \bar{\sigma}_w$ and as $Q_t \to 1$, $\sigma_{w,t+1} \to 0$. The parameter κ governs how fast the standard deviation of w shrinks as Q_t approaches 1 from above. This specification guarantees that $Q_t > 1$ (at least in continuous time). In other words, Q_t is a bounded martingale.

If investors are risk neutral with a time discount factor of β, asset returns have to satisfy the Euler equation

$$1 = E_t \left[\beta \frac{D_{t+1}}{D_t} \frac{1 + \frac{P_{t+1}}{D_{t+1}}}{\frac{P_t}{D_t}} \right].$$

Conjecture that the price-dividend ratio takes the following form

$$\frac{P_t}{D_t} = Q_t - 1.$$

Using this conjecture, the Euler equation becomes

$$1 = E_t \left[\beta e^{g + \varepsilon_{t+1}} A^{-1} \frac{Q_{t+1}}{Q_t} \right].$$

Because the innovation to Q_{t+1} is independent from the innovation ε_{t+1}, the conjecture is verified for

$$A = \beta e^{g + \sigma_\varepsilon^2/2} E_t \left[\frac{Q_{t+1}}{Q_t} \right] = \beta e^{g + \sigma_\varepsilon^2/2},$$

because Q_t is a martingale. The log dividend-price ratio is also an (approximate) martingale. Consider a Taylor approximation of $d_{t+1} - p_{t+1}$ around Q_t:

$$d_{t+1} - p_{t+1} = d_t - p_t + \frac{1}{1 - Q_t}(Q_{t+1} - Q_t) + O_{t+1}.$$

Because Q_t is a martingale, $E_t[d_{t+1} - p_{t+1}] = d_t - p_t + E_t[O_{t+1}]$. As long as the conditional mean of the higher order terms O_{t+1} is small, the log dividend-price ratio is approximately a martingale. For reasonable break processes the higher order terms turn out to be neglible. For example, if $\pi = .99, \bar{\sigma}_w = 5$ and $\kappa =$, the conditional mean of the higher order terms is 0.00006 for a price-dividend ratio level of 40, 0.00017 for a price-dividend ratio level of 20, and 0.00048 for a price-dividend ratio of 2, which tiny compared to the value of $d - p$ which is -3.10 on average in a simulation with the same parameters. By the same token, $E_t[q_{t+1}] \approx q_t$. The difference between the conditional expectation of q_{t+1} and q_t is -0.000005 for a price-dividend ratio level of 40, -0.0000068 for P/D=20 and -0.0003114 for a price-dividend ratio of 2. Again this is small relative to the simulation average for q of -0.0094.

The above specification imposes that $Q_t > 1$, or equivalently $q_t > -\infty$. It would be easy to impose tighter restrictions on the steady-state log dividend growth rate process q_t, based on a priori 'reasonable range' for log dividend growth by modifying (23). For example, one could impose that the long-run dividend growth rate should not be smaller than -5% per year. This amounts to price-dividend ratios in excess of 11.5 ($\tilde{Q} > 12.5$). The same bounded martingale approach would work with such a restriction.

D A Structural Model

We propose a model for log dividend growth Δd and log returns r, where expected dividend growth z and expected returns x follow an AR(1) with autoregressive coefficient ϕ:

$$\Delta d_{t+1} - \overline{d} = z_t + \epsilon_{t+1} \qquad z_{t+1} = \phi z_t + \zeta_{t+1} \tag{24}$$

$$r_{t+1} - \bar{r} = x_t + \eta_{t+1} \qquad x_{t+1} = \phi x_t + \xi_{t+1} \tag{25}$$

where $\overline{d}$ is the long-run mean log dividend growth and $\bar{r}$ is the long-run mean return. The model has three fundamental shocks: a dividend innovation ϵ_{t+1}, an innovation in expected dividends ζ_{t+1}, and an innovation in

expected returns ξ_{t+1}. Campbell (1991)'s return decomposition implies that

$$\eta_{t+1} = -\frac{\rho}{1-\rho\phi}\xi_{t+1} + \frac{\rho}{1-\rho\phi}\zeta_{t+1} + \epsilon_{t+1}. \tag{26}$$

We assume that all three errors are serially uncorrelated and have zero cross-covariance at all leads and lags: $Cov(\epsilon_{t+1}, \zeta_{t+j}) = 0$, $\forall j \neq 1$, and $Cov(\epsilon_{t+1}, \xi_{t+j}) = 0$, $\forall j \neq 1$, except $Cov(\zeta_t, \xi_t) = \chi$ and $Cov(\zeta_t, \epsilon_t) = \lambda$.

In steady-state, the log dividend-price ratio is $\overline{dp} = \log\left(\frac{\bar{r}-\bar{d}}{1+\bar{d}}\right)$, hence $\overline{DP} \approx \frac{\bar{r}-\bar{d}}{1+\bar{d}}$ and $\rho \approx \frac{1+\bar{d}}{1+\bar{r}}$. The log dividend-price ratio can be written as

$$dp_t = d_t - p_t = \overline{dp} + \frac{x_t - z_t}{1-\rho\phi}. \tag{27}$$

This equation clearly shows that the demeaned dividend-price ratio is an imperfect forecaster of returns. Returns are predicted by x_t (25) which not only contains the demeaned dp ratio, but also expected dividend growth z_t: $x_t = (1-\rho\phi)(dp_t - \overline{dp}) + z_t$.

This structural model implies a reduced form model that recovers the two predictability equations from Section (2):

$$(\Delta d_{t+1} - \overline{d}) = \kappa_d(dp_t - \overline{dp}) + \tau^d_{t+1} \tag{28}$$

$$(r_{t+1} - \bar{r}) = \kappa_r(dp_t - \overline{dp}) + \tau^r_{t+1} \tag{29}$$

$$(dp_{t+1} - \overline{dp}) = \phi(dp_t - \overline{dp}) + \tau^{dp}_{t+1}. \tag{30}$$

The third equation of this Vector Error Correction Model is an AR(1) process for the dividend-price ratio. Because of equation (6), the dividend-price ratio is the difference of two AR(1) processes with the same root ϕ, which is again an AR(1) process. We also considered a model where expected dividend growth has a different autoregressive coefficient $\psi \neq \phi$: $z_{t+1} = \psi z_t + \zeta_{t+1}$. In that case the dividend-price ratio is an ARMA(1,1) with roots $\phi + \psi$ and $-\phi\psi$. Since the dp ratio is well described by an AR(1) model in the data, we set $\phi = \psi$.

The slope coefficients are related to the structural parameters:

$$\kappa_d = \frac{Cov(\Delta d_{t+1}, d_t - p_t)}{Var(d_t - p_t)} = \frac{-(1-\rho\phi)(\sigma^2_\zeta - \chi)}{\sigma^2_\xi + \sigma^2_\zeta - 2\chi} \tag{31}$$

$$\kappa_r = \frac{Cov(r_{t+1}, d_t - p_t)}{Var(d_t - p_t)} = \frac{(1-\rho\phi)(\sigma^2_\xi - \chi)}{\sigma^2_\xi + \sigma^2_\zeta - 2\chi} \tag{32}$$

The innovations to the VECM, $\tau = (\tau^d, \tau^r, \tau^{dp})$, are given by:

$$\tau^d_{t+1} = \epsilon_{t+1} + x_t\left(\frac{-\kappa_d}{1-\rho\phi}\right) + z_t\left(\frac{\kappa_r}{1-\rho\phi}\right) \tag{33}$$

$$\tau^r_{t+1} = \epsilon_{t+1} + x_t\left(\frac{-\kappa_d}{1-\rho\phi}\right) + z_t\left(\frac{\kappa_r}{1-\rho\phi}\right) - \rho\frac{\xi_{t+1} - \zeta_{t+1}}{1-\rho\phi} \tag{34}$$

$$\tau^{dp}_{t+1} = \frac{\xi_{t+1} - \zeta_{t+1}}{1-\rho\phi}. \tag{35}$$

The structural model imposes a restriction on the innovation vector: $\rho\tau^{dp}_{t+1} = \left(\tau^d_{t+1} - \tau^r_{t+1}\right)$. Another way to write this restriction is as a restriction on a weighted sum of κ_r and κ_d:

$$\kappa_r - \kappa_d = 1 - \rho\phi.$$

We call this restriction the *present value constraint*.

Leaving aside the mean parameters $(\overline{d}, \bar{r}, \overline{dp})$ which play no role in the demeaned system, the structural parameter

vector is $\Theta = (\phi, \sigma_\zeta, \sigma_\xi, \sigma_\epsilon, \chi, \lambda)$. The variance σ_η^2 is implied by (26). The vector of parameters from the reduced form model (VECM) is $b = (\kappa_d, \kappa_r, \phi, \Sigma_\tau)$, where Σ_τ is the variance-covariance matrix of τ. There are six unique elements in this covariance matrix and the present value constraint imposes three restrictions on these six elements. In addition, the present value constraint imposes a restriction on one of the elements in $(\kappa_r, \kappa_d, \phi)$. Hence, there are five unique elements in $b = (\kappa_r, \kappa_d, \sigma_{\tau^r}, \sigma_{\tau^d}, \sigma_{\tau^r,\tau^d})$, where σ_{τ^r} and σ_{τ^d} are the standard deviations of τ^r and τ^d, and σ_{τ^r,τ^d} is the covariance between the two. When we estimate the non-singular system of equations (28) and (29), we can use these five coefficients to identify five out of six structural parameters in Θ. Therefore, whenever we back out structural parameters from the reduced form estimates, we find it convenient to tabulate results for a range of values for V_z, which measures the contribution of expected dividend growth to the variance of the dividend-price ratio:

$$V_z = \frac{\sigma_z^2}{(1-\rho\phi)^2\sigma_{dp}^2} = \frac{\sigma_\zeta^2}{(\sigma_\xi^2 + \sigma_\zeta^2 - 2\chi)}. \tag{36}$$

E Monte Carlo Simulations

The Monte Carlo exercise simulates the model under the null hypothesis of 1 break or 2 breaks in the dividend-price ratio. It then asks what the univariate long-horizon slope coefficients and R^2 are in a small sample (of the same length as the data) when we use the unadjusted dp as regressor versus the adjusted $\widetilde{dp}$. For the one-break case reported in the top panel of Table 8, we use the following algorithm:

step 1 To find the true parameters, we estimate the (1, 3, 5)-year VECM under the assumption of one break in 1991 for the full sample. This delivers reduced form estimates: $b = (.0121, .2098, .1410, .1868, .0176)$. Throughout, $\rho = .9616$, the value implied by the mean price-dividend ratio in the sample.

step 2 We invert these reduced form parameters to obtain structural parameters based on the identification scheme described above. We need to take a stance on the fraction of the variance in the dividend-price ratio attributable to expected dividend growth (equation 36). We set $V_z = 0.3$, but the results are not sensitive to this choice. The implied structural parameters are: $\Theta = (.8344, .0182, .0355, .1371, .0004, -.0005)$.

step 3 In each Monte Carlo iteration, we draw a new 178×3 vector of i.i.d. standard normal variables. The structural shocks $(\varepsilon_t, \zeta_t, \xi_t)$ are these standard normal variables pre-multiplied by $\Sigma_u^{\frac{1}{2}}$, where Σ_u is the covariance matrix of the structural innovations:

$$\Sigma_u = \begin{bmatrix} \sigma_\varepsilon^2 & \lambda & 0 \\ \lambda & \sigma_\zeta^2 & \chi \\ 0 & \chi & \sigma_\xi^2 \end{bmatrix}$$

step 4 We recursively build up time-series for x, $r - \bar{r}$, z, and $\Delta d - \overline{d}$ according to (25) and (24). We form a time series for the demeaned dividend-price ratio from (27).

step 5a Under the null hypothesis of 1 break in 1991, the *break-adjusted* series for returns, dividend growth and the dividend-price ratio are obtained by adding *constant* long-run means $\bar{r}$, $\overline{d}$, and $\overline{dp}$ to the demeaned series. I.e. the adjusted dp series is the demeaned series plus the mean over the *entire sample* 1927-2004, which is $\overline{dp} = -3.272$. Likewise, the adjusted dividend growth series Δd is formed by adding to the demeaned series $\Delta d - \overline{d}$ the sample mean $\overline{d} = .0432$. We then back out $\bar{r} = \overline{d} + (1 + \overline{d})\exp(\overline{dp}) = .0827$, and form the adjusted return series as the demeaned series plus this $\bar{r}$.

step 5b On the other hand, the *unadjusted* series still displays a break in 1991. To obtain the unadjusted series we need to add in a *different mean* before and after 1991. As before, we assume $\overline{d}$ did not change, so that

the change in $\overline{dp}$ entirely comes from a change in $\bar{r}$. The unadjusted dp series is obtained by adding in the 1927-1991 mean (-3.133) before 1991 and the 1992-2004 mean (-3.968) after 1991. Likewise for returns we add in .0886 before 1991 and .0629 after 1991. These are the subsample means $\overline{dp}$ and $\bar{r}$ that were reported in lines 1 and 2 of Table 6 (left panel).

step 6 We form annualized, cumulative long-horizon returns and dividend growth rates $\frac{1}{H}\sum_{j=1}^{H}\Delta r_{t+j}$ and $\frac{1}{H}\sum_{j=1}^{H}\Delta d_{t+j}$ in the same way. There is one set of unadjusted series and one set of adjusted series, corresponding to each horizon H.

step 7 After the formation, we discard the first 100 observations (burn-in), and are left with the same number of observations as in the data: 78- longest horizon+1. For the 1-, 3-, and 5-year system we report on, the longest horizon is 5, so all statistics are computed with 74 observations.

step 8 We then run univariate 1-, 3-, and 5-year ahead predictability regressions of *adjusted* returns and dividend growth on both *unadjusted* and *adjusted* dividend-price ratios. We keep track of the predictability coefficients and regression R^2.

step 9 We repeat this procedure 10,000 times and report the average predictability coefficients and R^2 across Monte Carlo iterations.

The Monte Carlo exercise in the two-break case is exactly analogous (bottom panel of Table 8). In step 1, we estimate the 1-, 3-, and 5-year ahead VECM system under the null of 2-breaks in 1954 and 1994. The VECM coefficient vector is $b = (0.0795, 0.4089, 0.1391, 0.1743, 0.0157)$. The implied structural parameter vector for $V_z = 0.3$ is $\Theta = (0.6974, 0.0293, 0.0635, 0.1339, 0.0014, -0.0012)$. To construct the unadjusted dp and return series we use the means reported in the right panel of Table 6.

We conduct a separate Monte Carlo exercise for the out-of-sample predictability (panels 2 and 3 in Table 5). We simulate the model under the null hypothesis of two breaks in 1954 and 1994. For simplicity we use the same true parameters in both panels 2 and 3. They are obtained from estimating the 1-year ahead VECM under the assumption of two breaks in 1954 and 1994, but for the period 1947-2004, which is the same sample period as we use in the data (panel 1 of Table 5). We then construct adjusted and unadjusted dp ratios as in the above algorithm, and run the same out-of-sample predictions as in the data.

Table 1: Forecasting Returns and Dividend Growth with the Dividend-Price Ratio.

This table reports estimation results for the equations $r_{t+1} - \bar{r} = \kappa_r(dp_t - \overline{dp}) + \tau^r_{t+1}$ and $\Delta d_{t+1} - \overline{d} = \kappa_d(dp_t - \overline{dp}) + \tau^d_{t+1}$. The first two columns report the equation for returns. The next two columns report the predictability equation for dividend growth. The last two columns are for excess returns instead of gross returns. The table reports point estimates and standard errors in parentheses of κ_r and κ_d, as well as regression R^2 in square brackets. The parameters $(\bar{r}, \overline{d}, \overline{dp})$ are the sample means of log returns r (log excess returns in the last two columns), log dividend growth Δd and the log dividend-price ratio dp. The top panel compares the case of no break in the log dividend-price ratio ($\overline{dp}$ is fixed) with the case where there is a break in the log dividend-price ratio: $\overline{dp}_1$ is the sample mean log dividend-price ratio for 1927-1991 and $\overline{dp}_2$ is the mean for 1992-2004. The estimation is by GMM, where the moments are the OLS normal conditions. Standard errors are by Newey-West with four lags. Row 1 reports results for the full sample; rows 2 and 3 report results for two subsamples. Row 4 reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1991 in the forecasting equations. The bottom panel compares the case of no break in the log dividend-price ratio ($\overline{dp}$ is fixed) with the case where there are two breaks in the log dividend-price ratio: ($\overline{dp}_1$ is the sample mean log dividend-price ratio for 1927-1954 (row 6), $\overline{dp}_2$ is the mean for 1955-1994 (row 7), and $\overline{dp}_3$ is the mean for 1995-2004 (row 8). Row 9 reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural breaks in 1954 and 1994 in the forecasting equations.

	Returns		**Dividend Growth**		**Excess Returns**	
Sample	No Break	1 Break	No Break	1 Break	No Break	1 Break
1927-2004	.094	.235	.005	.019	.113	.282
	(.046)	(.058)	(.037)	(.047)	(.049)	(.059)
	[.038]	[.100]	[.000]	[.001]	[.050]	[.132]
1927-1991	.235	.235	.014	0.014	.295	.295
	(.065)	(.065)	(.053)	(.053)	(.071)	(.071)
	[.087]	[.087]	[.001]	[.001]	[.125]	[.125]
1992-2004	.235	.235	.035	.035	.241	.241
	(.134)	(.134)	(.103)	(.103)	(.139)	(.139)
	[.199]	[.199]	[.006]	[.006]	[.198]	[.198]
Chow F-stat	3.408	.134	.114	.024	4.383	.230
p-val	[.038]	[.875]	[.892]	[.977]	[.016]	[.795]

Sample	No Break	2 Breaks	No Break	2 Breaks	No Break	2 Breaks
1927-2004	.094	.455	.005	.124	.113	.441
	(.046)	(.081)	(.037)	(.073)	(.049)	(.101)
	[.038]	[.223]	[.000]	[.032]	[.050]	[.193]
1927-1954	.510	.510	.037	.037	.529	.529
	(.175)	(.175)	(.182)	(.182)	(.192)	(.192)
	[.163]	[.163]	[.002]	[.002]	[.170]	[.170]
1955-1994	.383	.383	.142	.142	.336	.336
	(.106)	(.106)	(.077)	(.077)	(.144)	(.144)
	[.240]	[.240]	[.064]	[.064]	[.151]	[.151]
1995-2004	.532	.532	.226	.226	.539	.539
	(.129)	(.129)	(.097)	(.097)	(.145)	(.145)
	[.546]	[.546]	[.126]	[.126]	[.533]	[.533]
Chow F-stat	4.390	.235	.998	.500	3.261	.186
p-val	[.003]	[.918]	[.414]	[.736]	[.016]	[.945]

Table 2: Tests for Change in Mean of Log Dividend-Price Ratio

The first panel reports dates of structural breaks in the mean of the log dividend price ratio estimated by the Perron procedure as well as the changes in the mean before and after the breaks. The second panel reports sup-F(i,j) statistics where i is the number of breaks under the null hypothesis and j is the number of breaks under the alternative. '*', '**', '***' denote significance at the 1%, 5% and 10% level, respectively. The third panel reports the number of breaks chosen according to the Bayesian Information criterion (BIC) and the modified Schwartz' criterion proposed by Liu, Wu, and Zidek (1997) (LWZ). The tests allow for autocorrelation in the residuals and the trimming value is set to 5% of the sample. The bottom panel reports first and second order autocorrelation coefficients, an Augmented Dickey Fuller test, testing the null hypothesis of a unit root (and associated p-value), and the time-series standard deviation for the unadjusted log dividend-price ratio, the log price ratio adjusted for a change in its mean in 1991, and the log dividend-price ratio adjusted for a change in its mean in 1954 and 1994.

# of Breaks	**Date(s)**	$\Delta\overline{dp}$
1	1991	-.86
2	1954, 1994	-.37, -.78
Test (H_0, H_1)	**Statistic**	
sup-F (0,1)	13.7***	
sup-F (0,2)	23.9***	
sup-F (1,2)	9.64*	
Information Criterion	**# of Breaks**	
LWZ	2	
BIC	2	

Persistence Properties of Adjusted Dividend-Price Ratio

	AC(1)	AC(2)	ADF Test	p-val	s.d.
dp, unadjusted	.91	.81	-1.383	.586	.42
$\widetilde{dp}$, adjusted, 1 break	.77	.55	-3.016	.038	.26
$\widetilde{dp}$, adjusted, 2 breaks	.61	.23	-4.731	.010	.20

Table 3: Tests for Change in Mean of Financial Ratios

The top half of the table reports the test-statistic of a sup-F Perron structural break tests of the null hypothesis of no break against the alternative hypothesis of one (first row) or two (second row) breaks with unknown break date. It reports the p-value of the test statistic, as well as the resulting break date. The last column reports the estimated change in means before and after the break(s). These tests are performed for the log earnings-price ratio $ep = e - p$, the log book value -to-market value of equity ratio $bm = b - m$, the log dividend-earnings ratio $de = d - e$, the log dividend-price ratio adjusted for repurchases dp^{rep}, and the log dividend-price ratio of the universe of CRSP firms that *excludes* the NASDAQ firms dp^{nas}.

	Structural Break Tests					
	H_0	H_1	$sup-$F-Test	p-value	Date(s)	$\Delta mean$
ep	0 break	1 break	15.5	$< 1\%$	1990	-.67
	0 break	2 breaks	18.0	$< 1\%$	1953, 1994	-.50, -.62
bm	0 break	1 break	9.3	$< 10\%$	1953	-.80
	0 break	2 breaks	17.9	$< 1\%$	1953, 1990	-.71, -.33
de	0 break	1 break	5.6	$> 10\%$	1993	-.24
	0 break	2 breaks	5.3	$> 10\%$	1990, 1993	+.27, -.46
dp^{nas}	0 break	1 break	10.2	$< 5\%$	1992	-.75
	0 break	2 breaks	18.1	$< 1\%$	1954, 1995	-.35, -.70
dp^{rep}	0 break	1 break	3.7	$> 10\%$	1990	-.43
	0 break	2 breaks	4.5	$> 10\%$	1954, 1991	-.23, -.34
	0 break	3 breaks	20.1	$< 1\%$	1957, 1973, 1990	-.46, +.50, -.57

Table 4: Forecasting Returns With Other Financial Ratios.

This table reports estimation results for the equation $r_{t+1} - \bar{r} = \kappa_r(y_t - \overline{y}) + \tau^r_{t+1}$, where y is the log earnings-price ratio $ep = e - p$ in the first row, the log book-to-market value ratio $bm = b - m$ in the second row, the log dividend-price ratio without the NASDAQ firms dp^{nas} in the third row, and the repurchase adjusted log dividend-price ratio dp^{rep} in the fourth row. The table reports point estimates and standard errors in parentheses of κ_r, and the regression R^2 in brackets. The regressor is the unadjusted valuation ratio in the first column, the one-break adjusted valuation ratio in the second column, and the two-break adjusted valuation ratio in the third column. For the predictor dp^{rep}, we report the three-break case as well. The break dates for all regressors are reported in Table 3. The sample is 1946-2004 in row 1, and 1927-2004 in all other rows.

Predictor y	No Break	1 Break	2 Breaks	3 Breaks
ep	.119	.214	.216	
	(.030)	(.039)	(.045)	
	[.104]	[.190]	[.185]	
bm	.070	.255	.308	
	(.036)	(.063)	(.064)	
	[.030]	[.154]	[.188]	
dp^{nas}	.110	.250	.417	
	(.048)	(.056)	(.090)	
	[.043]	[.105]	[.182]	
dp^{rep}	.191	.282	.361	.576
	(.054)	(.065)	(.084)	(.097)
	[.079]	[.126]	[.161]	[.250]

Table 5: Out-of-Sample Predictability.

The table reports one-period-ahead return forecast errors based on the Random Walk model (row 1) and based on the forecasting equation $r_{t+1} - \bar{r} = \kappa_r(dp_t - \overline{dp}) + \tau^r_{t+1}$ with fixed $\overline{dp}$ (row 2). Rows 3 through 5 use the real time Perron procedure to estimate $\widehat{dp}^P$. We report results for three different methods of selecting the number of break: the sequential sup-F test with 10% critical value, and the LWZ and BIC information criteria. Rows 6 and 7 use the ex-post break adjusted dividend price ratios with a change in $\overline{dp}$ in 1991 (row 6), and two changes in the mean $\overline{dp}$ in 1954 and 1994 (row 7). Rows 8 to 11 use the Hamilton approach to construct adjusted dividend-price ratios $\widehat{dp}^H$. We consider the cases of two regimes and three regimes. All numbers denote returns per annum. The second and third panels report results from a Monte Carlo exercise. We simulate the structural model under the null hypothesis that the data generating process has one break in 1991 (panel 2) or has two breaks in 1954 and 1994 (panel 3). We compare the same three out-of sample forecasting exercises as in the data (panel 1). Except in the last panel, we also look at the forecast errors when we only correct for the second break and not the first one. The structural parameters in panels 2 and 3 are the same and were obtained from the Vector Error Correction Model (VECM) parameters estimated under the assumption of two breaks in 1954 and 1994 for the period 1947-2004, the same forecasting period as in panel 1.

	Mean absolute error	Root mean squared error
	Panel 1: Data	
Benchmarks		
Random Walk	.1338	.1605
Unadjusted dp	.1411	.1685
Pure OOS		
$\widehat{dp}^P$ - Perron sequential sup-F	.1350	.1661
$\widehat{dp}^P$ - Perron LWZ criterion	.1391	.1680
$\widehat{dp}^P$ - Perron BIC criterion	.1370	.1646
Pseudo OOS		
$\widetilde{dp}$ - Perron, ex-post, 1 break	.1309	.1558
$\widetilde{dp}$ - Perron, ex-post, 2 breaks	.1158	.1421
$\widehat{dp}^H$ - Hamilton, 1 break, smoothed	.1286	.1543
$\widehat{dp}^H$ - Hamilton, 2 breaks, smoothed	.1100	.1340
$\widehat{dp}^H$ - Hamilton, 1 break, unsmoothed	.1330	.1590
$\widehat{dp}^H$ - Hamilton, 2 breaks, unsmoothed	.1243	.1512
	Panel 2: Monte Carlo - 1 break	
Random Walk	.1212	.1525
Unadjusted dp	.1210	.1555
$\widetilde{dp}$ - ex-post, 1 break	.1058	.1334
	Panel 3: Monte Carlo - 2 breaks	
Random Walk	.1222	.1533
Unadjusted dp	.1203	.1516
$\widetilde{dp}$ - ex-post, 1 break	.1165	.1493
$\widetilde{dp}$ - ex-post, 2 breaks	.1064	.1336

Table 6: Implied Changes in Steady-state Expected Returns and Dividend Growth

The top panel (bottom panel) of the table reports the mean log dividend-price ratio (log earnings-price ratio) in the subsamples, as well as the difference between the two. On the second row is reports the mean return $\overline{r}$ in the subsamples if all of the changes in $\overline{dp}$ ($\overline{ep}$) in the first row were attributable to changes in mean returns. The third row reports mean dividend growth rates $\overline{d}$ in the subsamples if all of the changes in $\overline{dp}$ ($\overline{ep}$) in the first row were attributable to changes in mean dividend growth. The left panel reports the case of one break; the right panel the case of two breaks. For the log dividend-price ratio the break date is estimated to be 1991 for the one break case (1954 and 1994 for the two break case). For the log earnings-price ratio, the break is estimated to be 1990 for the one break case (1953 and 1994 for the two break case). In the bottom panel, we also report the change in the payout rate, the log dividend-earnings ratio de. The sample for the top panel is 1927-2004, the sample for the bottom panel is 1946-2004 (see Appendix A).

			Dividend-Price Ratio					
	27-91	92-04	Change	27-54	55-94	95-04	Changes	
$\overline{dp}$	-3.133	-3.968	-.835	-2.940	-3.301	-4.086	-.362,	-.785
$\overline{r}\downarrow$, $\overline{d}$ constant	8.86%	6.29%	-2.60%	9.83%	8.16%	6.07%	-1.67%,	-2.09%
$\overline{d}\uparrow$, $\overline{r}$ constant	5.00%	7.54%	2.54%	4.07%	5.68%	7.76%	1.61% ,	2.08%
			Earnings-Price Ratio					
	46-90	91-04	Change	46-53	54-94	95-04		
$\overline{ep}$	-2.540	-3.202	-.662	-2.217	-2.652	-3.267	-.435,	-.616
Payout rate $\overline{de}$	-.665	-.779	-.114	-.606	-.661	-.889	-.056,	-.227
$\overline{r}\downarrow$, $\overline{d}$ constant	11.72%	9.24%	-2.47%	13.70%	11.25%	8.87%	-2.46%,	-2.38%
$\overline{d}\uparrow$, $\overline{r}$ constant	6.38%	8.79%	2.41%	4.52%	6.83%	9.16%	2.31%,	2.33%

Table 7: Estimation with Long-Horizon Moments

This table reports GMM estimates for the parameters $(\kappa_d, \kappa_r, \phi)$, their asymptotic standard errors and p-values. The results in panel A are for the system with one-year ahead equations for dividend growth and returns ($H = 1$, $N = 0$). The results in panel B are for the system with one-year, three-year and five-year ahead equations for dividend growth and returns ($H = \{1, 3, 5\}$, $N = 2$). The estimation is by GMM. The first stage weighting matrix is the identity matrix. The asymptotic standard errors and p-values are computed using the Newey-West HAC procedure (second stage weighting matrix) with four lags in panel A and $H = 5$ lags in panel B. The first number in the last column denotes the present-value constraint violation of the univariate OLS slope estimators: $(1 - \rho\phi^{ols})^{-1}(\kappa_r^{ols} - \kappa_d^{ols})$. It is expressed in the same units as κ_d and κ_r. In panel B this number is the average violation of the three constraints, one constraint at each horizon. The second number in the last column reports the average moment violation. In panel A that number is not available (N/A) because the system is exactly identified. The dividend-price ratio in rows 1 and 4 is the unadjusted one. In rows 2 and 5, the dividend-price ratio is adjusted for one break in 1991 (see equation 8), and in rows 3 and 6, it is the series adjusted for two breaks in 1954 and 1994. All estimation results are for the full sample 1927-2004.

	κ_d	κ_r	ϕ	PV violation	moment violation
Panel A: No Long-Horizon Moments $H = \{1\}$					
No Break	.005	.094	.945	-.046	N/A
	(.037)	(.046)	(.052)		
1 Break ('91)	.019	.2353	.813	.004	N/A
	(.047)	(.0554)	(.052)		
2 Breaks ('54, '94)	.124	.4553	.694	-.001	N/A
	(.073)	(.0792)	(.070)		
Panel B: Long-Horizon Moments $H = \{1, 3, 5\}$					
No Break	.021	.068	.990	.189	.205
	(.018)	(.038)	(.032)		
1 Break ('91)	.012	.210	.834	.076	.085
	(.019)	(.043)	(.042)		
2 Breaks ('54, '94)	.080	.409	.697	.100	.144
	(.065)	(.078)	(.060)		

Table 8: Long-Horizon Predictability in Data and Monte Carlo Exercise.

The table reports results from univariate regressions of cumulative long-horizon returns on the log dividend-price ratio. The left columns denote slope coefficients for one-year, three-year, and five-year horizon regressions; the middle columns report standard errors for the slope coefficients; the right columns report the corresponding regression R^2. The rows labeled 'Data, *dp* unadj.' denote regressions run with real data using the unadjusted log dividend-price ratio as independent variable. The rows labeled 'Data, *dp* adj.' denote regressions run with real data using the log dividend-price ratio, adjusted for a break in the mean $\overline{dp}$. In the top panel there is one break in this mean in 1991; in the bottom panel there are two breaks in 1954 and 1994. The results for the data are contrasted with a the results from a Monte carlo exercise. The return and dividend growth system is estimated until 1991 in the top panel (1954 in the bottom panel) on real data. The estimated parameters imply 'true' structural parameters. The theoretical long-horizon slope coefficients and regression R^2 are reported in the row with label 'True Values in Sim.' For these parameters, the structural model is then simulated 10,000 times for 78 periods. The row 'Sim., *dp* unadj.' denotes the Monte carlo average slope coefficient and R^2 statistic using the unadjusted log dividend-price ratio as independent variable. The row 'Sim., *dp* adj.' also reports regression coefficients and statistics of regressions on artificial data, but now the independent variable comes from a model where the mean $\overline{r}$ is adjusted to equal the change in $\overline{dp}$ in the data.

	κ_r			*s.e.*			R^2 (%)		
Horizon	1	3	5	1	3	5	1	3	5
				One Break in 1991					
True Values in Sim.	.222	.187	.159	N/A	N/A	N/A	8.9	20.0	25.2
Sim., *dp* unadj.	.115	.092	.075	.064	.034	.025	5.1	11.5	14.6
Data, *dp* unadj.	.087	.095	.071	.065	.037	.026	2.4	8.1	9.2
Sim., *dp* adj.	.264	.214	.175	.093	.048	.034	10.5	22.7	28.0
Data, *dp* adj.	.220	.200	.150	.084	.047	.033	8.6	20.0	22.0
				Two Breaks in 1954 and 1994					
True Values in Sim.	.456	.330	.249	N/A	N/A	N/A	20.7	37.0	38.8
Sim., *dp* unadj.	.128	.089	.064	.063	.032	.023	6.2	11.5	12.8
Data, *dp* unadj.	.087	.095	.071	.065	.037	.026	2.4	8.1	9.2
Sim., *dp* adj.	.483	.340	.249	.110	.052	.037	21.4	37.5	38.6
Data, *dp* adj.	.442	.355	.231	.102	.054	.041	20.7	37.4	30.8

Figure 1: Forecasting Returns - Rolling Regressions

The top panel plots estimation results for the equation $r_{t+1} - \bar{r} = \kappa_r(dp_t - \overline{dp}) + \tau^r_{t+1}$. It shows the estimates for κ_r using 30-year rolling windows. The dashed line in the left panels denote the point estimate plus or minus one standard deviation. The parameters $\bar{r}$ and $\overline{dp}$) are the sample means of log returns r and the log dividend-price ratio dp. The data are annual for 1927-2004. The middle panel gives the slope coefficient κ_r from a regression where the right-hand side variable is $\widetilde{dp}$, adjusted for 1 break in 1991 (see Section 3.3). The bottom panel gives the slope coefficient κ_r from a regression where the right-hand side variable is $\widetilde{dp}$, adjusted for 2 breaks in 1954 and 1994 (see Section 3.3). The standard errors are asymptotic.

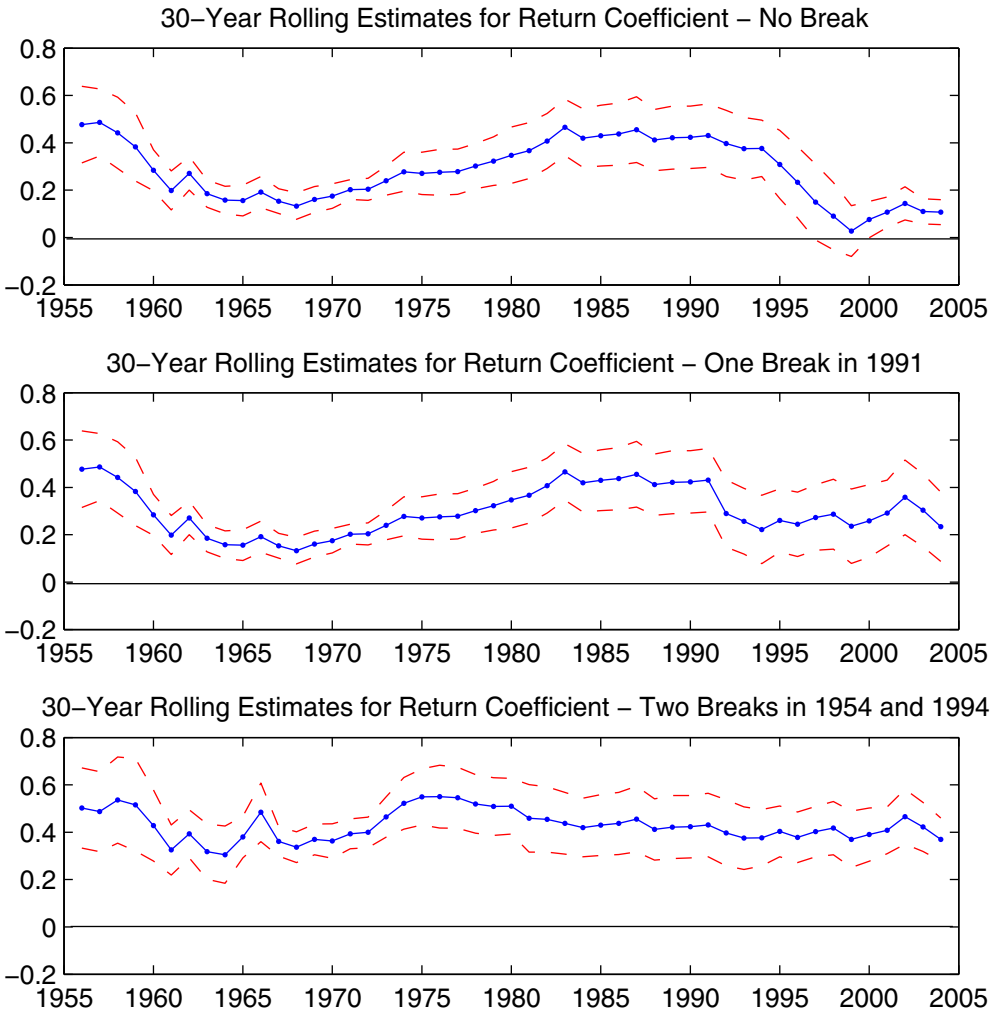

Figure 2: Change in the Mean of the Dividend-Price Ratio.

The top left panel plots the log dividend-price ratio $dp_t = d_t - p_t$ (solid line) as well as its sample means $\overline{dp}_1$ in the subsample 1927-1991 and $\overline{dp}_2$ in the subsample 1992-2004 (dashed line). The bottom left panel overlays the subsample means $\overline{dp}_1$ in 1927-1954, $\overline{dp}_2$ in 1955-1994, and $\overline{dp}_3$ in 1995-2004. The top right panel plots the adjusted dividend-price ratio $\widetilde{dp}_t = dp_t - \overline{dp}_1, t = 1, ..., \tau$ and $dp_t - \overline{dp}_2, t = \tau, ..., T$. The bottom right panel plots the adjusted dividend-price ratio in the two-break case. In the two bottom panels, the adjusted series is rescaled so that it coincides with the adjusted series for the first subsample.

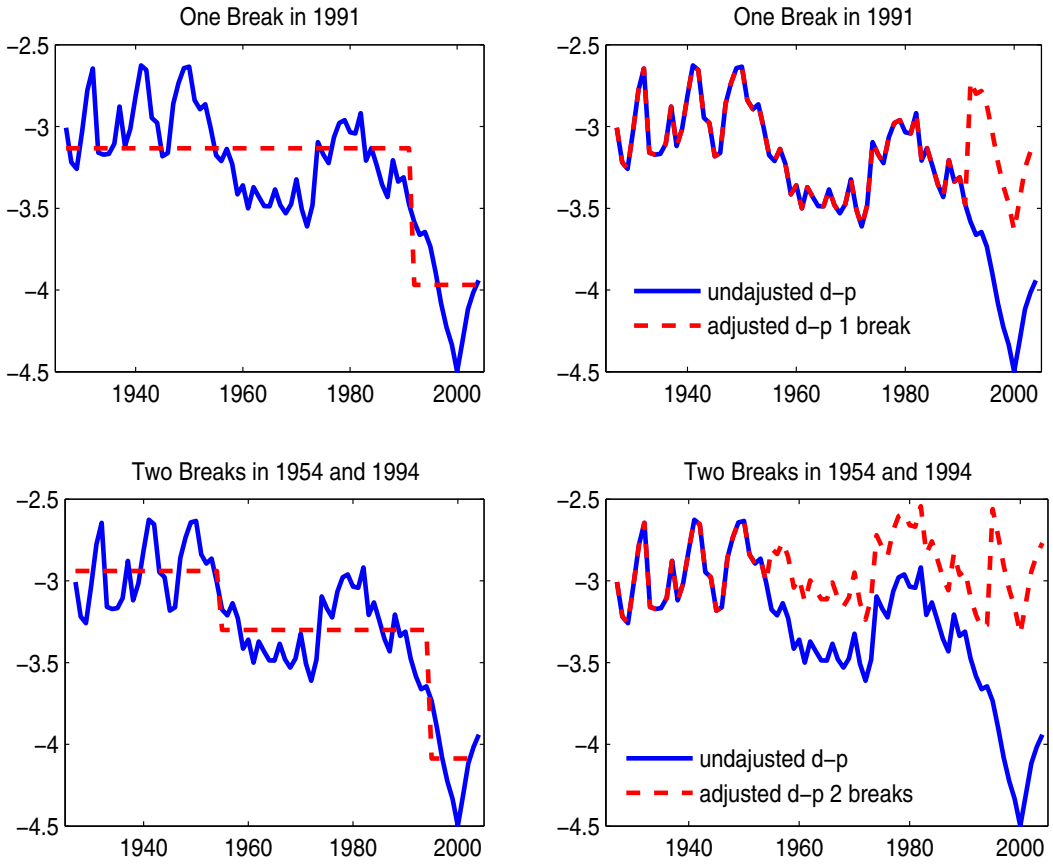

Figure 3: The Empirical Distribution of the Dividend-Price Ratio.

The figure plots the smoothed empirical distribution of the log dividend-price ratio dp (solid line), alongside the smoothed density obtained from drawing from the empirical distribution with replacement (bootstrap, dash-dotted line), and the smoothed density from a Monte-Carlo exercise (dashed line).

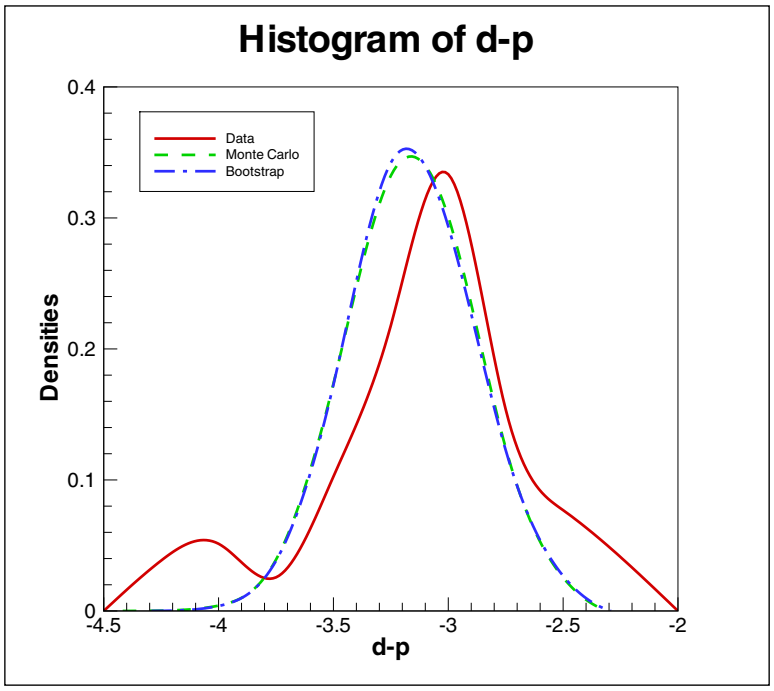

Figure 4: Recursive Estimation of the Empirical Distribution of the Dividend-Price Ratio.

We recursively estimate an AR(1) for the log dividend-price ratio dp, $dp_{t+1} = c + \phi dp_t + \tau_{t+1}^{dp}$ using data up to time $t+1$, and bootstrap percentiles of the empirical distribution by drawing with replacement from the residuals $\{\tau_1^{dp}, \cdots, \tau_{t+1}^{dp}\}$. The dashed lines represent the 2.5, 5, 95, and 97.5 percentiles of the bootstrapped distribution. The initial sample is 1927-1951. Each successive exercise adds one year of data. The solid line represents the observed log-dividend-price ratio in deviation from its recursive sample mean.

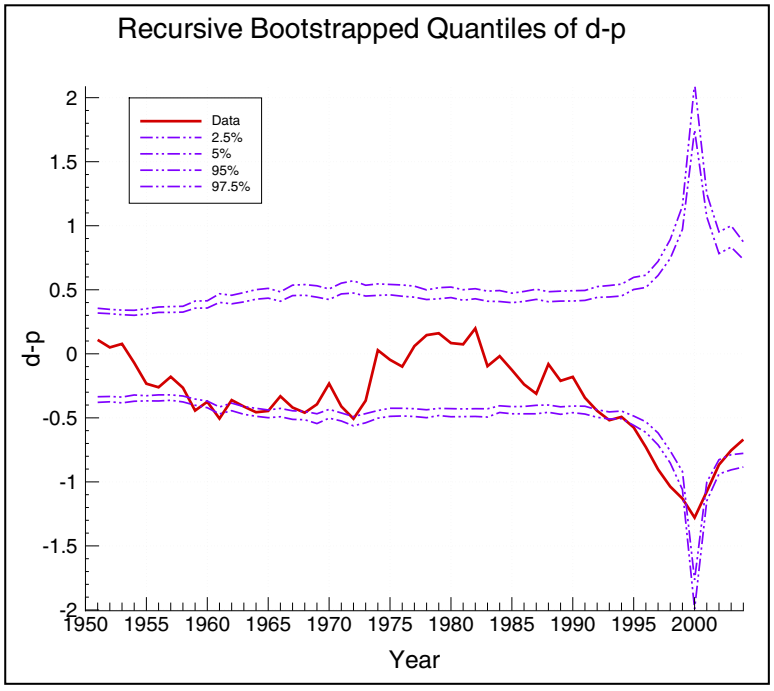

Figure 5: Regime Switching Model - State probabilities

This figure reports estimation results from a Hamilton regime-switching model for the log dividend-price ratio $d-p$. The top panel is for the two-regime (one-break) case; the bottom panel is for the three-regime (two-break) case. The solid lines represent unsmoothed probabilities, based on real-time information. The dashed lines represent smoothed probabilities. The top panel represents the probability of being in the low $\overline{dp}$ regime; the bottom panel plots the probabilities of being in the low and in the middle $d-p$ regimes. The latter are denoted by squares and diamonds.

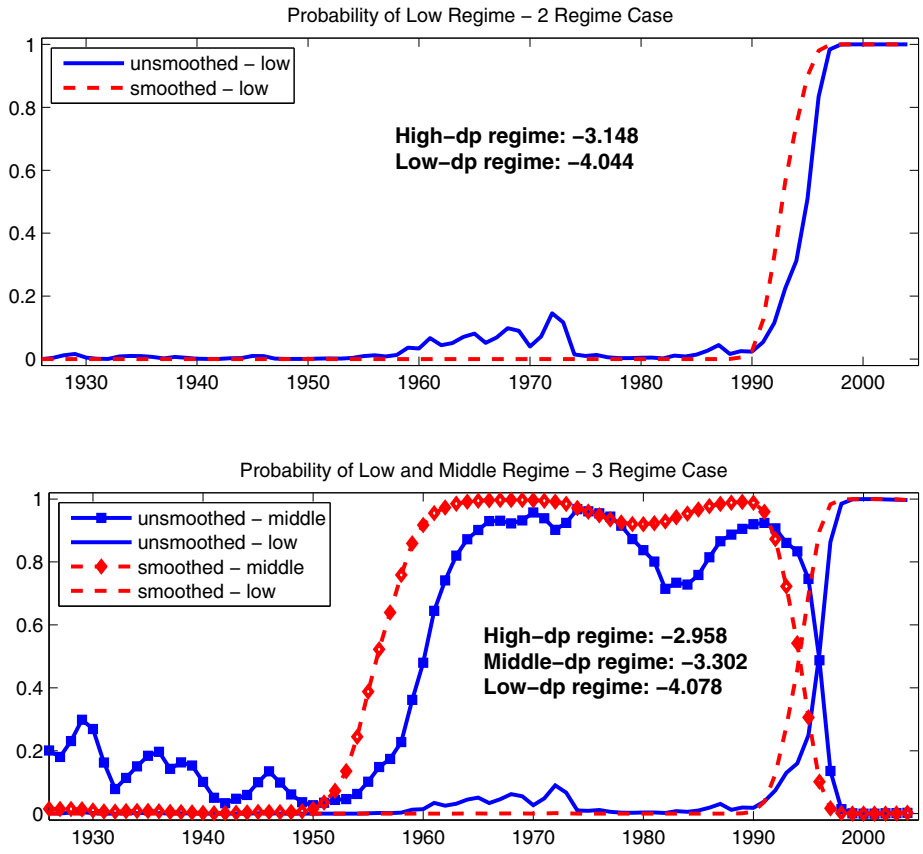

Figure 6: Regime Switching Model - Mean Dividend-Price Ratio

This figure reports the mean dividend-price ratio based on the estimation results from a Hamilton regime-switching model for the log dividend-price ratio $d - p$. The top panel is for the two-regime (one-break) case; the bottom panel is for the three-regime (two-break) case. The solid line is the weighted average of the regime means, where the weights are the unsmoothed (real-time) state probabilities. The dashed line is a weighted average of the same regime means, but the weights are the smoothed state probabilities (based on the entire sample).

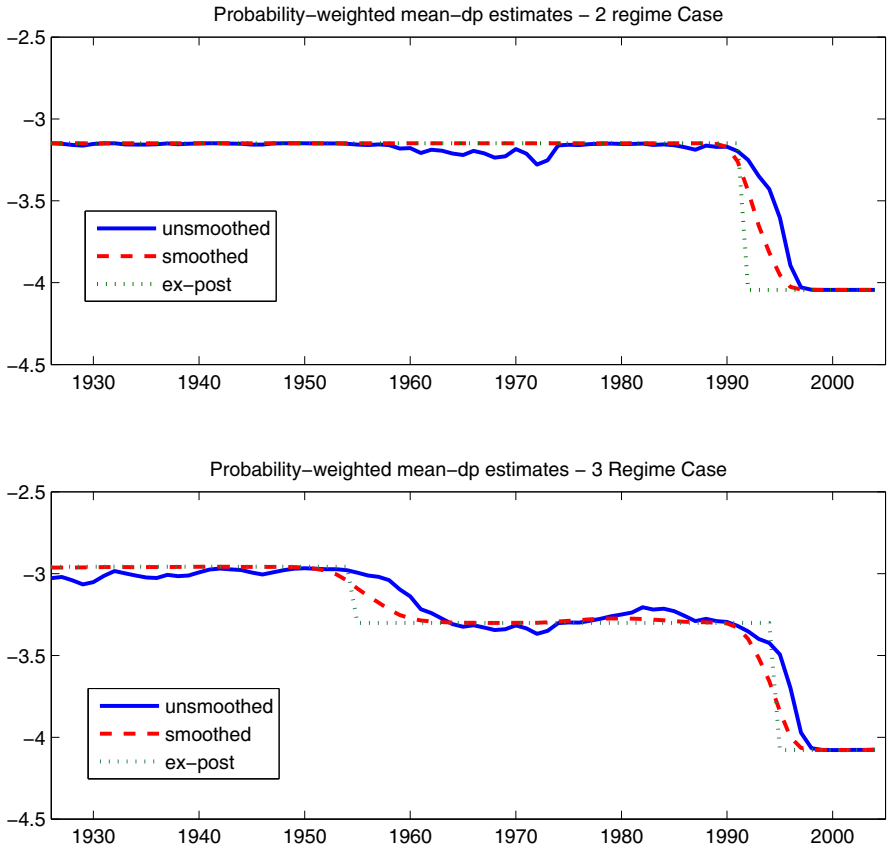

Figure 7: Estimation with Long-Horizon Moments

This figure compares the univariate OLS long-horizon regression coefficients, $\kappa_d(H) - OLS$ and $\kappa_r(H) - OLS$ to the GMM estimates that impose the present-value restriction (12), $\kappa_d(H) - GMM$ and $\kappa_r(H) - GMM$. The system contains 10 equations, 5 return and 5 dividend growth equations. The horizons (in years) are $H \in 1, 3, 5, 7, 10$. The top panel uses the unadjusted dp ratio as predictor, the middle panel uses the $\widetilde{dp}$ ratio adjusted for one break in 1991, and the bottom panel uses the $\widetilde{dp}$ ratio adjusted for two breaks in 1954 and 1994.