Dividend Growth Predictability and the Price-Dividend Ratio*

Ilaria Piatti† and Fabio Trojani‡

June 2018

Abstract

Asymptotic tests over-reject the null of no predictability in present-value models. We develop a nonparametric testing approach in state space models, implying reliable finite sample inference under weak assumptions on price-dividend ratio and dividend shocks. We find sharp evidence of return predictability in postwar US data, but a less consistent evidence of dividend predictability, which is significant only using cash-flow proxies reflecting information from mergers and acquisitions. These findings reconcile the diverging conclusions of present-value models and common predictive regressions, in a way that is robust to the choice of the predictive variables, the sample period and alternative cash-flow proxies.

Keywords: Predictability, Present-value model, State space model, Bootstrap, Likelihood ratio test.

JEL Classification: C12, C15, C32, G12.

*We thank Karim Abadir, John Cochrane, Walter Distaso, Christian Julliard, Ralph Koijen, Tarun Ramadorai, Alberto Rossi, Oleg Rytchkov, Jules Van Binsbergen, Stijn Van Nieuwenburgh, Paolo Zaffaroni, the participants of the Arne Ryde Workshop in Financial Economics 2012, Lund, the 11th Swiss Doctoral Workshop in Finance, Gerzensee, the 2012 European Summer Symposia in Financial Markets, Gerzensee, the 7th End-of-Year Conference of Swiss Economists Abroad, Luzern, the 2015 EFA Annual Meeting, Vienna, and seminar participants at the University of Geneva and at the Imperial College, London, for valuable comments. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK, project A3) and the Swiss Finance Institute (Term Structures and Cross-Sections of Asset Risk Premia). The usual disclaimer applies.

†Saïd Business School, University of Oxford, Park End Street, Oxford OX1 1HP, UK; e-mail: ilaria.piatti@sbs.ox.ac.uk

‡University of Geneva and Swiss Finance Institute, Via Buffi 13, CH-6900 Lugano, Switzerland; e-mail: fabio.trojani@alphacruncher.com

1 Introduction

Are stock market returns and dividend growth predictable? Campbell and Shiller (1988) observation that the price-dividend ratio reflects information on both future expected returns and expected dividend growth has motivated a vast literature that studies predictability features based on predictive regressions of returns and cash flow growth on lagged dividend yields.

Univariate predictive regression results typically imply an economically significant evidence of return predictability in postwar US data, even if the statistical significance is weaker in some subperiods, and an almost constant expected dividend growth. This evidence suggests that the price-dividend ratio varies mainly because of discount rate shocks; See Campbell (1991) and Cochrane (1992), among others.[1] The Kalman filter estimation of a benchmark present-value model with latent dividend and return expectations yields both a predictable return and a predictable dividend growth, indicating that the price-dividend ratio varies because of both dividend expectation and discount rate shocks. A possible explanation for these diverging findings in benchmark present-value models is the higher discriminatory power that these models imply for detecting joint dividend and predictability structures; see, e.g., Binsbergen and Koijen (2010).

In this paper, we test the predictability hypotheses of present-value models written in state space form, based on a new approach with reliable finite-sample inference and valid asymptotic properties. Our approach relies on a general nonparametric Monte Carlo bootstrap for state space models, which avoids strong assumptions on the properties of dividend and price-dividend ratio shocks. Based on the more accurate finite sample inference provided by our approach, we study and reinterpret the diverging findings of present-value models and standard predictive regressions. We document that while asymptotic tests strongly over-reject the null of no predictability in present-value models, our testing method delivers a reliable finite-sample inference. Importantly, the empirical application of our testing methodology to benchmark present-value models, which parsimoniously aggregate dividend growth and price-dividend ratio information, produces a strong evidence of return predictability and a less consistent evidence of dividend predictability for most cash-flow proxies in the literature, consistent with the evidence produced by univariate predictive regressions.

[1] Early predictive regression studies are Rozeff (1984), Schiller (1984), Keim and Stambaugh (1986), Campbell and Shiller (1988) and Fama and French (1988). The predictive regression findings of no dividend predictability can also depend on the sample period used for estimation, as the conclusions are opposite for the prewar sample, or the cash-flow proxy used; see, e.g., Chen (2009), Boudoukh, Michaely, Richardson, and Roberts (2007) and Sabbatucci (2015).

Without a reliable testing approach, comparing the findings of present-value models with those of standard predictive regressions is challenging, as it is well-known that the predictability inference in the latter context can be difficult to interpret for a number of reasons. First, the correlation between shocks in stock returns and predictive variables, combined with the high persistence of the latter, can create finite-sample biases and a non-standard asymptotic behaviour for common tests of return or dividend predictability; see e.g., Stambaugh (1999) and Torous, Valkanov, and Yan (2004).[2] Second, the dividend yield is a noisy estimate of expected returns and expected dividend growth, as it reflects expectations of both future stock returns and future cash flows. This feature creates a standard error-in-variable (EIV) problem in univariate predictive regressions; see, e.g., Binsbergen and Koijen (2010), among others. Cochrane (2008a) stresses that the weak evidence of return predictability in earlier univariate studies is stronger if one considers the joint nature of null hypotheses on the return-dividend process. Present-value models offer a potentially powerful testing framework for such joint hypotheses, as they directly incorporate the no-arbitrage constraints on stock returns, cash flows and valuation ratios.[3] However, inference procedures for present-value models written in state space form are more recent and less studied than in more common predictive regression settings. Our paper contributes to fill this gap and to reconcile the diverging predictability findings in the literature.

Similar to predictive regressions, inference in state space models is made tractable by the existence of an asymptotic theory, which under appropriate conditions implies consistency and asymptotic normality of parameter and latent state estimates; see, e.g., Liung and Caines (1979) and Spall and Wall (1984). However, the moderate length of the time series available in many predictability studies can make the use of asymptotic inference methods potentially suspect for these models as well.[4] To improve over the conventional asymptotic inference, a

[2]Stambaugh (1999) derives an analytic expression for the bias in univariate predictive regressions. Kothari and Shanken (1997), Amihud and Hurvich (2004), Lewellen (2004), Torous, Valkanov, and Yan (2004), Campbell and Yogo (2006) and Polk, Thompson, and Vuolteenaho (2006) develop methods for hypothesis testing in univariate settings. Amihud, Hurvich, and Whang (2009) propose an analytic method for hypothesis testing in regression with multiple predictors, while Lettau and Ludvigson (2001) and Ang and Bekaert (2007), among others, use bootstrap methods in this setting.

[3]Given the time variation in the price-dividend ratio, at least one between returns and dividend growth must be predictable. Cochrane (2008a) also derives upper bounds on price-dividend ratio autocorrelations, to deliver more powerful statistics in the joint testing of return and dividend growth predictability. Recent studies estimating market expectations for returns and dividends with different present-value models include Menzly, Santos, and Veronesi (2004), Lettau and Ludvigson (2005), Ang and Bekaert (2007), Lettau and Van Niewerburgh (2008), Campbell and Thompson (2008), Pastor, Sinha, and Swaminathan (2008), Cochrane (2008a,b), Binsbergen and Koijen (2010), among others.

[4]The close relation between present-value models and their (VAR) reduced-form predictive regression representations (see, e.g., Cochrane (2008b)) also suggests that if samples must be fairly large before asymptotic theory is applicable, then this should similarly hold both in predictive regressions and present-value models. See

useful nonparametric method, which does not rely on strong assumptions about dividend and price-dividend ratio shocks, is the nonparametric bootstrap first suggested by Efron (1979). Stoffer and Wall (1991) prove that the bootstrap applied to the innovations of a time-invariant and stable state space model yields asymptotically correct results. They also demonstrate, by Monte Carlo simulation and in a number of real-data applications, that a bootstrap approach can improve over the finite-sample inference of conventional asymptotics.[5] We borrow from these insights and introduce a new class of bootstrap tests of predictability hypotheses that are computationally accessible in general present-value models with hidden expectations. We first show the asymptotic validity of our testing method and demonstrate by Monte Carlo simulation the improved finite-sample properties over conventional asymptotic tests. Key to the validity and accuracy of our approach, also compared to alternative bootstrap approaches, is that it is based on a resampling of the standardized prediction errors in the innovation form representation of the state space model.

Using the new bootstrap tests, we obtain a novel set of findings and interpretations for the predictability evidence obtained by latent variable approaches within present-value models. The detailed contributions to the literature are the following.

First, we study the finite-sample properties of tests of predictability in present-value models, without assuming a particular error distribution, such as, e.g., a normal distribution, by means of a nonparametric Monte Carlo simulation approach. We find that asymptotic tests imply large finite-sample biases, which often lead to an incorrect rejection of the null of no predictability. For instance, while according to a significance level $\alpha = 5\%$ the asymptotic critical value of the likelihood ratio test of no dividend (return) predictability is 7.81 (9.49), the Monte Carlo finite-sample critical value is 17.13 (15.99). Therefore, the fraction of incorrect rejections of the null of no time-variation in dividend (return) expectations using the asymptotic test is as large as 25.8% (60.5%). Our Monte Carlo evidence also shows that the large estimated R^2's for dividends or returns in the data can arise by chance alone, even under the null of constant expected dividend growth or expected return. These features are the consequence of the volatile point estimates for the persistence of dividend and return expectations, which stay in a close relation to the estimated R^2's. This evidence emphasizes the importance of integrating a pure

also Section A of the Supplemental Appendix.

[5]Finite sample inference could potentially be improved also using a Bayesian approach that imposes prior belief information on the relevant parameter space or the degree of predictability; see e.g., Pastor and Stambaugh (2009) and Wachter and Warusawitharana (2009). In contrast with predictive regression settings with fully observable state space, an open question in our setting with hidden expectation processes is how to translate potential prior beliefs about return and dividend growth predictability into mathematically formulated prior distributions of parameters in the observable and latent variables dynamics that enter the present value restriction.

estimation approach with a reliable testing method when quantifying the degree of dividend or return predictability.

Second, we propose a general nonparametric bootstrap test of predictability, by applying the bootstrap to the dividend and price-dividend ratio shocks in the innovation form of the state space model, simulated under the relevant null hypothesis. We prove that our bootstrap test implies a valid asymptotic inference under standard conditions and we document the finite sample improvements over the asymptotic tests. These findings show that our bootstrap testing approach better controls the probability of rejecting a null hypothesis because of chance alone, thus producing a more reliable inference in a number of applications.

Third, we apply our bootstrap tests to US stock market data, using several specifications of the predictive information set and different proxies of market cash flows. We systematically find strong evidence in favour of time-varying expected returns in postwar US data, implying test $p-$values typically well below 1%, and a clearly weaker and less robust evidence of dividend growth predictability. Indeed, the null of no dividend predictability is never statistically significant at the 5% significance level. It implies $p-$values of about 9.5% using standard proxies of market cash flows, a lowest $p-$value of 7% for cash flow proxies that incorporate information from mergers and acquisitions, but also a $p-$value as large as 16% using total payouts. These findings confirm that the postwar return (dividend) predictability evidence in benchmark present-value models is similarly strong (weak) as in standard predictive regression tests. Similarly, we find that returns are unpredictable in the prewar period, while dividend growth clearly is, very consistently with the tale of two periods documented in the literature using standard predictive regressions.

Finally, we propose an extension of our bootstrap approach for estimating the distribution of out-of-sample predictive R-squared. As for the in-sample R-squared, we find that the large estimated out-of-sample R-squared for dividends in the data can arise by chance alone, under the null of constant dividend growth expectations.

In summary, the results of our bootstrap tests reconcile the diverging predictability findings in the literature and offer a new methodology for properly assessing the in-sample and the out-of-sample predictability properties. These findings are of first-order importance, as they clearly demonstrate the difficulties in reliably testing predictability hypotheses with asymptotic tests in benchmark applications of present-value models. Such difficulties have been largely overlooked in the literature, which has virtually never applied formal bootstrap methods to study predictability in present-value models.[6]

The paper proceeds as follows. Section 2 introduces the benchmark present-value model. It then briefly discusses the data and estimation strategy, before reporting the estimation results. Section 3 presents the R-squared for returns and dividend growth implied by the present-value model. It then summarizes the diverging evidence of likelihood ratio tests of predictability based on asymptotic critical values and bootstrap critical values, respectively. Section 4 introduces the bootstrap testing approach more formally. It proves its asymptotic validity, studies the improvements in finite-sample inference compared to standard asymptotic approaches, and illustrates the advantages compared to alternative bootstrap approaches. Section 5 extends our bootstrap methodology to study the distribution of in- and out-of-sample predictive R-squared. Section 6 discusses the robustness of our main findings with respect to the inclusion of the prewar sample, different cash flow proxies and various predictive information sets. Section 7 concludes.

2 Present-Value Approach

Borrowing from Binsbergen and Koijen (2010), we introduce the benchmark cash flow and discount rate dynamics. This model offers a tractable framework to estimate the expected return and expected dividend growth processes, by parsimoniously aggregating the time-series information from dividend growth and price-dividend ratios. Even though the benchmark model restricts the information set to be spanned by the history of dividend (or returns) and price-dividend ratios, it is flexible enough to capture the essential aspects related to the estimation and testing of predictive relations.[7] Our bootstrap testing approach is applicable also in more general present-value models and the implications of broader specifications of the predictive information set are studied in detail in Appendix E.

2.1 The benchmark model

Let

$$r_{t+1} \equiv \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \tag{1}$$

[6] Moreover, our approach would allow to test essentially any parametric hypothesis in state space settings.

[7] The same setting can result from a general equilibrium framework with multiple securities and time-varying risk aversions; see, e.g., Menzly, Santos, and Veronesi (2004). Recent studies have investigated predictability in the context of the model considered in this paper, including Cochrane (2008b), Binsbergen and Koijen (2010), among others. Model extensions and different special cases have also been considered in Lettau and Ludvigson (2005), Ang and Bekaert (2007), Lettau and Van Niewerburgh (2008), Campbell and Thompson (2008), Pastor, Sinha, and Swaminathan (2008), and Yun (2012).

be the cum-dividend log market return and denote by

$$\Delta d_{t+1} \equiv \log\left(\frac{D_{t+1}}{D_t}\right), \tag{2}$$

the aggregate log dividend growth. Expected dividend growth and return, conditional on the information at time t, are denoted by $g_t \equiv E_t[\Delta d_{t+1}]$ and $\mu_t \equiv E_t[r_{t+1}]$, respectively. They follow simple autoregressive processes:

$$g_{t+1} = \gamma_0 + \gamma_1(g_t - \gamma_0) + \varepsilon^g_{t+1}, \tag{3}$$

$$\mu_{t+1} = \delta_0 + \delta_1(\mu_t - \delta_0) + \varepsilon^\mu_{t+1}. \tag{4}$$

The dividend growth rate is the expected dividend growth plus an orthogonal shock:

$$\Delta d_{t+1} = g_t + \varepsilon^d_{t+1}. \tag{5}$$

The vector of independent and identically distributed shocks $(\varepsilon^g_{t+1}, \varepsilon^\mu_{t+1}, \varepsilon^d_{t+1})'$ has covariance matrix

$$\Sigma = \begin{bmatrix} \sigma^2_g & \sigma_{g\mu} & \sigma_{gd} \\ \sigma_{g\mu} & \sigma^2_\mu & \sigma_{\mu d} \\ \sigma_{gd} & \sigma_{\mu d} & \sigma^2_d \end{bmatrix}. \tag{6}$$

The affine explicit expression for the log price-dividend ratio directly follows from a Campbell and Shiller (1988) log linearisation:

$$pd_t = A - B_1(\mu_t - \delta_0) + B_2(g_t - \gamma_0), \tag{7}$$

where A, B_1 and B_2 are simple functions of the model parameters such that, consistent with intuition, pd_t is decreasing in expected returns and increasing in expected dividend growth.[8]

2.2 Estimation results

We obtain the with- and without-dividend monthly returns on the value-weighted portfolio of all NYSE, Amex and Nasdaq stocks, in the period from January 1946 until December 2010, from the Center for Research in Security Prices (CRSP). We construct annual time series of

[8] See Binsbergen and Koijen (2010) and Appendix A. Campbell and Shiller (1988) approximation in our sample holds almost exactly, for yearly data, if annual dividends and prices are constructed as described in Section 2.2.

aggregate dividends and prices, assuming that monthly dividends are cash-reinvested at the 30-day T-bill rate. Data on 30-day T-bill rates are also obtained from CRSP.[9]

We estimate the model with a Kalman filter based on a Gaussian quasi likelihood function, from the observable time series of dividend growth Δd_{t+1} and price-dividend ratios pd_{t+1}. Due to the present-value relations, market return r_{t+1} is redundant with respect to Δd_{t+1} and pd_{t+1}.[10] The parameter estimates are reported in Table 1, with bootstrapped standard errors in parenthesis.[11] We find an unconditional expected log return (dividend growth) of $\delta_0 = 8.3\%$ ($\gamma_0 = 5.7\%$). Both expectation processes feature some degree of persistence, with autoregressive roots γ_1 and δ_1 equal to 0.304 and 0.927, respectively, and expected returns that are substantially more persistent than expected dividend growth. Finally, expected dividend growth is estimated as very volatile ($\sigma_g = 6.5\%$), while unexpected dividend growth variability is very low ($\sigma_d = 0.2\%$).

3 Dividend and Return Predictability

Let I_t denote the econometrician's information set at time t, generated by the history of dividends and price-dividend ratios. A nice feature of the Kalman filter is to provide filtered estimates of the unknown latent states μ_t and g_t, conditional on I_t. Thus, a standard measure of the degree of predictability in model (3)-(5) can be computed by the fraction of r_{t+1} and Δd_{t+1} variability explained by μ_t and g_t, respectively:

$$R^2_{Ret} = 1 - \frac{\widehat{Var}(r_{t+1} - \mu_t)}{\widehat{Var}(r_{t+1})}, \tag{8}$$

$$R^2_{Div} = 1 - \frac{\widehat{Var}(\Delta d_{t+1} - g_t)}{\widehat{Var}(\Delta d_{t+1})}, \tag{9}$$

where $\widehat{Var}$ denotes sample variances.

[9] Several studies use market-reinvested instead of cash-reinvested dividends. CRSP computes quarterly or annual return series under the stock market reinvestment assumption. Chen (2009) and Koijen and Van Nieuwerburgh (2011) note that a market reinvestment assumption can be problematic, as it transfers some of the properties of returns to cash flows, inducing a large volatility of the resulting dividend growth series and a low correlation with other measures of dividend growth.

[10] Using (r_{t+1}, pd_{t+1}) as observable variables, the estimation results are almost identical and one can always recover the missing variable using Campbell and Shiller (1988) approximation; see also Cochrane (2008b), among others. Details on the estimation procedure are collected in Appendix B. For comparability with Binsbergen and Koijen (2010) we adopt their identification assumption $\rho_{gd} = 0$. More general identification assumptions allowing for a non-zero correlation between expected and realized dividends do not affect our predictability findings, see also Piatti and Trojani (2017).

[11] Parameter standard errors are obtained using the circular block-bootstrap of Politis and Romano (1992), in order to account for the potential serial correlation in the data.

We find that $R^2_{Ret} = 8.82\%$ and $R^2_{Div} = 17.58\%$ (see also Table 2), suggesting a relatively high degree of both return and dividend growth predictability. In contrast to these findings, simple regressions of returns and dividend growth on lagged price-dividend ratios yield R^2 of about 9.9% and 0.95%, respectively. A standard interpretation for these diverging results is the noisiness of the price-dividend ratio (7) as a signal for expected returns and expected dividend growth, respectively, which creates a potential EIV problem in predictive regressions of returns and dividend growth on lagged price-dividend ratios.[12] Indeed, the large persistence of return expectations is linked to a large sensitivity of price-dividend ratios to expected return shocks ($B_1 = 10.332$) and a smaller sensitivity to dividend expectation shocks ($B_2 = 1.421$), which could obfuscate the predictive power of dividend expectations in dividend predictive regressions, leading to the low R^2.

It is important to realize that estimated R-squares are per se not sufficient to draw reliable conclusions about returns and dividend growth predictability. This issue is best addressed using an appropriate testing framework. As emphasized by Cochrane (2008a), while a point estimate produces the most likely predictability structure according to a given statistical metric, hypothesis testing is essential to control for the probability of detecting predictability structures by chance alone.

3.1 Likelihood ratio tests

Most predictability hypotheses can be formulated by means of simple parametric constrains, which can be efficiently tested with standard parametric tests. We focus on the likelihood ratio (LR) test,[13] based on the statistic

$$LR_T = 2\left(\max_{\Theta} \log \mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right) - \max_{\Theta_0} \log \mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right)\right), \qquad (10)$$

where Θ_0 is the restricted set of parameters under null hypothesis H_0 and $\log \mathcal{L}$ is the log-likelihood of the model. Evidence against H_0 is collected when LR is large:

$$\{LR_T > c_{1-\alpha}\}\ , \qquad (11)$$

relative to critical value $c_{1-\alpha}$, which is unlikely under H_0.

The finite sample distribution of the LR statistic (and the critical value $c_{1-\alpha}$) is unknown,

[12] See e.g. Binsbergen and Koijen (2010) and Lettau and Ludvigson (2005).

[13] This is without loss of generality, since it is straightforward to apply our proposed bootstrap approach to Wald and score-type tests, which are asymptotically equivalent to the likelihood ratio test.

but it can be approximated in different ways. A widely used approximation is based on the well-known asymptotic limit of this distribution as $T \to \infty$, which is a χ^2_r distribution with r degrees of freedom, where r is the number of parameter constraints defining the constrained parameter set Θ_0. Therefore, the choice $c_{1-\alpha} = \chi^2_{r,1-\alpha}$, where $\chi^2_{r,1-\alpha}$ is the $1-\alpha$ quantile of the chi-square distribution, ensures asymptotically a probability α of rejecting H_0 by chance alone:

$$\alpha = \lim_{T \to \infty} P_{H_0}(LR_T > \chi^2_{r,1-\alpha}) \ . \tag{12}$$

This paper proposes a novel approach, based on bootstrap, to approximate the finite sample distribution of the LR statistic under the null hypothesis. In this way, we obtain a novel class of bootstrap likelihood ratio tests of predictability hypotheses in state space models. We discuss in this section the diverging conclusions of asymptotic and bootstrap likelihood ratio tests in the benchmark present-value model. In Section 4, we detail our bootstrap testing procedure, prove its asymptotic validity and study the finite-sample improvements relative to asymptotic tests.

3.2 Null hypotheses and test results

Testing return and cash flow growth predictability is equivalent to examining time variation in expected returns and expected dividend growth, respectively. The null of constant return expectations is:

$$H_0 : \delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0 \ . \tag{13}$$

Similarly, the null of constant dividend growth expectations is:[14]

$$H_0 : \gamma_1 = \sigma_g = \rho_{g\mu} = 0 \ . \tag{14}$$

Table 2 shows that the asymptotic test based on statistic (10) very clearly rejects both null hypotheses at a significance level α below 0.5%. The evidence provided by the bootstrap test is strikingly different, as the bootstrap p-values of both tests are much higher than their asymptotic counterparts. This implies a null hypothesis of no return predictability that can still be rejected at a 1% significance level, but a null hypothesis of no dividend predictability that

[14] Under the null (13) (the null (14)) all price-dividend ratio variation comes from variation in expected dividend growth (returns) and the present-value model collapses to a standard linear regression of dividend growth rates (returns) on the lagged price-dividend ratio. Note that $\rho_{gd} = 0$ is imposed also in the unconstrained model for identification purposes, as for instance in Binsbergen and Koijen (2010), see Appendix B.

can be rejected only at the 10% significance level, with a p-value of about 9.5%. These findings are consistent with the significant evidence of return predictability and the weaker evidence of dividend predictability resulting from standard predictive regressions.

Appendix E shows that our main results are unaffected by extensions of the benchmark present-value model that include additional predictive variables, as the bootstrap p-values are again systematically larger than the asymptotic p-values and we can never reject the null of no dividend growth predictability at the 5% significance level. In contrast, in all specifications the null hypothesis of no return predictability is rejected with p-values lower than 0.5%. Section 6.2 shows that the main findings are even strengthened using proxies of total payouts (dividend plus repurchases) instead of cash dividends. In this case, the null of no return predictability is still rejected with a p-value of 1%, but the p-value for the null of no dividend predictability is only 16%. In contrast, the asymptotic test rejects in all cases also the null of no dividend predictability with unplausibly low $p-$values.[15]

In Section 4, we prove formally and verify by Monte Carlo simulation that bootstrap and asymptotic p-values are asymptotically equivalent. Therefore, the diverging conclusions of these two testing methods are the consequence of their different finite sample properties in testing null hypotheses of no predictability in state space models. We investigate this important aspect in more detail in the next sections.

4 Bootstrap Tests in the Present-Value Model

We obtain asymptotically valid tests that are less susceptible to finite-sample distortions and less dependent on specific distributional assumptions using a nonparametric bootstrap method.[16] In this section we first introduce our bootstrap testing method in present-value models to then prove its asymptotic validity and quantify by Monte Carlo simulation the improvements over conventional asymptotic tests.

[15] Following the suggestion of an anonymous referee, we also assessed the robustness of our findings with respect to the *pd* ratio persistence properties using an equivalent state space specification depending only on realized dividends and returns, following Rytchkov (2012). The main evidence on dividend growth predictability produced by this analysis is unchanged.

[16] As shown in Hall and Horowitz (1996) and Andrews (2002), among others, a desirable property of the bootstrap is that it may provide more accurate finite-sample approximations of the sampling distribution of standard $t-$test statistics for testing the null of no predictability in predictive regression models.

4.1 State space representation

For observed variables $Y_t := (\Delta d_t, pd_t)'$ and expanded state vector $X_t := (\hat{g}_{t-1}, \varepsilon_t^g, \varepsilon_t^\mu, \varepsilon_t^d)'$, where $\hat{g}_t := g_t - \gamma_0$, the present-value model can be written in state space form (see Appendix B):

$$X_{t+1} = FX_t + \Gamma\varepsilon_{t+1}^X \ , \tag{15}$$

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t \ , \tag{16}$$

with matrices F, Γ, M_0, M_1, M_2 that are functions of parameter vector

$$\theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_d, \rho_{g\mu}, \rho_{\mu d}, \rho_{gd})' \ .$$

Let $X_{t,t-1}$ be the best linear prediction of X_t based on observable data $\{Y_s\}_{s=1}^{t-1}$, obtained via the Kalman filter, and $\eta_t = Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t,t-1}$ the corresponding prediction error. The innovations form representation of the present-value model follows from the Kalman filter as:

$$X_{t+1,t} = FX_{t,t-1} + F\mathcal{K}_t \eta_t \ , \tag{17}$$

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_{t,t-1} + \eta_t \ , \tag{18}$$

where the Kalman gain $\mathcal{K}_t$ and the conditional covariance matrix S_t of innovation η_t are given explicitly in Appendix B. The advantage of this representation for an efficient bootstrap procedure is that it allows to easily simulate forward the dynamics of observable variables $\{Y_1, \ldots, Y_T\}$, given initial conditions Y_0, $X_{0,0}$ and random innovations $\{\eta_1, \ldots, \eta_T\}$.[17]

4.2 Nonparametric bootstrap

Let $\hat{\theta}$ and $\hat{\theta}_0$ be the unconstrained and the constrained estimators of the model parameters, obtained by maximizing the likelihood function (46) in Appendix B over the full and the H_0-constrained parameter set, $\hat{\Theta}$ and $\hat{\Theta}_0$, respectively. The observed value of the likelihood ratio statistic LR_T then follows from definition (10).[18]

We apply a nonparametric bootstrap for time series to the (standardized) innovations

[17] In practice, we first apply a nonparametric bootstrap for time series to efficiently simulate the joint distribution of innovations $\{\eta_1, \ldots, \eta_T\}$. In a second step, we simulate the joint distribution of $\{Y_1, \ldots, Y_T\}$ using the forward dynamics (17)-(18).

[18] Bootstrap inference is always conditional on the observed sample of data. With a slight abuse of notation, in the sequel we denote by LR_T the sample value of the likelihood ratio statistics.

$\{\hat{e}_t := S_t^{-1/2}(\hat{\theta})\eta_t(\hat{\theta})\}_{t=1}^T$, in order to obtain the standardized bootstrap residuals $\{\hat{e}_t^*\}_{t=1}^T$. The bootstrap residuals are used to compute a bootstrap distribution of maximum likelihood estimators $\hat{\theta}^*$:

$$\hat{\theta}^* = \arg\max_{\Theta} \log \mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right) \ , \tag{19}$$

where the sequence $\{Y_t^*\}_{t=1}^T$ is simulated with the dynamics (17)-(18) applied to the unstandardized bootstrap residuals $\{\hat{\eta}_t^* := S_t^{1/2}(\hat{\theta})e_t^*\}_{t=1}^T$. Stoffer and Wall (1991) prove that this approach gives rise to a valid bootstrap distribution for $\sqrt{T}(\hat{\theta}^* - \hat{\theta})$, which is asymptotically equivalent to the distribution of $\sqrt{T}(\hat{\theta} - \theta^\star)$, where $\theta^\star$ is the true unknown parameter value. We start from this result to construct a valid nonparametric bootstrap likelihood ratio test of null hypothesis H_0 in the present-value model.

4.3 Bootstrap likelihood ratio test

Our bootstrap likelihood ratio test for state space model (15)-(16) is based on the following six-steps algorithm.

1) Using the estimated parameter vector under null hypothesis H_0, construct the (constrained) time series of standardized innovations $\{\hat{e}_{0t}\}_{t=1}^T$, by setting:

$$\hat{e}_{0t} = S_t^{-1/2}(\hat{\theta}_0)\eta_t(\hat{\theta}_0) \ . \tag{20}$$

2) Applying a nonparametric bootstrap procedure to time series $\{\hat{e}_{0t}\}_{t=1}^T$, compute a bootstrap sample $\{\hat{e}_{0t}^*\}_{t=1}^T$ of standardized innovations.

3) Using the innovation form representation (17)-(18), construct a bootstrap sample $\{Y_t^*\}_{t=1}^T$ as follows:

$$X_{t+1,t}^* = FX_{t,t-1}^* + F\mathcal{K}_t S_t^{1/2}\hat{e}_{0t}^* \ , \tag{21}$$
$$Y_t^* = M_0 + M_1 Y_{t-1}^* + M_2 X_{t,t-1}^* + S_t^{1/2}\hat{e}_{0t}^* \ , \tag{22}$$

where matrices F, $\mathcal{K}_t$, S_t, M_0, M_1, M_2 are all evaluated in $\hat{\theta}_0$ and the initial conditions are $Y_0^* = Y_0$, $X_{0,-1}^* = X_{0,0}$.

4) Using bootstrap sample $\{Y_t^*\}_{t=1}^T$, compute constrained and unconstrained maximum likelihood point estimates $\hat{\theta}_0^*$ and $\hat{\theta}^*$, respectively, by maximizing the log likelihood function

$\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)$, while imposing and not imposing null hypothesis H_0, respectively.

5) Following definition (10), compute the value LR_T^* of the likelihood ratio statistic in the bootstrap sample, defined by:

$$LR_T^* = 2\left(\mathcal{L}\left(\hat{\theta}^*, \{Y_t^*\}_{t=1}^T\right) - \mathcal{L}\left(\hat{\theta}_0^*, \{Y_t^*\}_{t=1}^T\right)\right) \; . \tag{23}$$

6) Repeat steps 2)-5) a large number of times, B, to obtain a collection of bootstrap values of the likelihood ratio statistics, $\{LR_{T,b}^*, 1 \leq b \leq B\}$. The empirical distribution of these values provides an approximation of the distribution of the likelihood ratio statistic under the null hypothesis H_0.

Remark 1 *(i) In step 2) of the algorithm, several bootstrap procedures are applicable to the standardized innovations $\{\hat{e}_{0t}\}_{t=1}^T$. We strongly recommend the use of a time-series bootstrap, such as the circular block-bootstrap (Politis and Romano (1992)), in order to robustify the inference against a potentially left time series dependence in the finite sample distribution of standardized innovations, which might not have been fully captured by the estimated conditional moment dynamics. This approach also easily accommodates a possible non normality of standardized innovations in finite samples. ii) In some cases, it may help to exclude the random sampling of the innovations for the first 2-3 data points in step 2) of the algorithm, e.g., by setting $\hat{e}_{0t}^* = \hat{e}_{0t}$ for $t = 1, 2, 3$. This is useful to avoid start-up problems of the algorithm when the Kalman filter might have an initially transient behaviour, e.g., with large values of the Kalman gain $\mathcal{K}_t$.*

An important question is whether the proposed bootstrap test delivers correct results in large samples, i.e., whether the bootstrap likelihood ratio statistic LR_T^* follows the same asymptotic distribution as LR_T under H_0. The next theorem justifies our bootstrap likelihood ratio test of null hypothesis H_0.

Theorem 1 *Under regularity conditions detailed in Appendix C, it follows as $B, T \to \infty$: $LR_T^* \longrightarrow \chi_r^2$, in distribution.*

According to Theorem 1, the bootstrap statistic LR_T^* has an asymptotically equivalent distribution to LR_T under H_0.[19] Therefore, it gives rise to bootstrap tests with the correct significance

[19] A similar result can be proven with respect to sequences of shrinking local alternative hypotheses $H_{A,T}$, by applying the above algorithm to innovations defined by $\hat{e}_{At} = S_t^{-1/2}(\hat{\theta}_{A,T})\eta_t(\hat{\theta}_{A,T})$ in step 1), where $\hat{\theta}_{A,T}$ is the constrained maximum likelihood estimator computed under the local alternative $H_{A,T}$.

level asymptotically. The most convenient way to define a bootstrap likelihood ratio test of H_0 is by means of the so-called bootstrap $p-$value:

$$p^*(LR_T) := P^*(LR_T^* > LR_T) = \frac{1}{B}\sum_{b=1}^{B} \mathbb{I}(LR_{T,b}^* > LR_T) \ , \tag{24}$$

where P^* denotes the bootstrap probability measure. Using bootstrap $p-$values, the bootstrap test rejects H_0 whenever:

$$p^*(LR_T) < \alpha \ . \tag{25}$$

From Theorem 1, this rejection region implies the correct asymptotic size α. The interesting question is whether our bootstrap test can deliver more reliable results than the asymptotic test in finite samples. A useful property in this respect is that the inference based on bootstrap procedures applied to asymptotically pivotal statistics, such as the likelihood ratio statistic in our setting, is generally more accurate than the inference of conventional asymptotics, in the sense that the errors made are of lower order in the sample size T;[20] see Beran (1988), Davidson and MacKinnon (1999b), Hall and Horowitz (1996) and Andrews (2002), among others. As a consequence, our bootstrap likelihood ratio tests theoretically improve over the conventional asymptotic inference in finite samples. We systematically verify this important aspect in the next section using Monte Carlo simulations.

4.4 Finite-sample properties of bootstrap and asymptotic tests

We compare the finite sample properties of asymptotic and bootstrap tests in the benchmark present-value model, focusing on the finite sample probability of detecting predictive relations by chance alone. Based on the asymptotic test, this probability is given by:

$$\alpha_T := P_{H_0}(LR_T > \chi^2_{r,1-\alpha}) \ , \tag{26}$$

where P_{H_0} is the probability distribution in the present-value model under null hypothesis H_0. Based on the bootstrap test, this probability is given by:

$$\alpha_T^* := P_{H_0}(p^*(LR_T) < \alpha) \ . \tag{27}$$

[20] An asymptotically pivotal statistic is a statistic with sampling distribution asymptotically independent of nuisance parameters.

Since both these tests are valid asymptotic tests that are asymptotically equivalent, these two probabilities converge to the significance level α as $T \to \infty$. In the sequel, we investigate how reliable the target rejection probability α is as an approximation of the finite sample rejection probabilities α_T and α_T^*, in the asymptotic and the bootstrap test, respectively.

In order to simulate under probability P_{H_0}, we impose the null hypothesis H_0 in the benchmark present-value model, using the constrained ML estimator $\hat{\theta}_0$. We simulate $S = 1000$ time series of dividend growth and price dividend ratios, using our nonparametric bootstrap applied to the fitted innovations in the present-value model; see again the algorithm in Section 4.3. In this way, we simulate the empirical distribution of observed data under H_0, without imposing strong assumptions on the joint distribution of dividend and price-dividend ratio shocks.[21]

The first (second) Panel of Figure 1 displays for null hypothesis (13) (null hypothesis (14)) the quantiles of the empirical distribution of statistic (10), against the quantiles of the asymptotic χ^2_4 (χ^2_3) distribution. Apparently, the finite sample distributions of these likelihood ratio statistics under H_0 deviate substantially from their chi-squared asymptotic limit. For instance, for an asymptotic level $\alpha = 5\%$ the asymptotic critical value in the test of a constant expected dividend growth is $\chi^2_{3,0.95} = 7.81$. In contrast, the finite sample critical value is more than two times larger (17.13), indicating a strong tendency of asymptotic tests to over-reject the given null hypothesis. We can quantify the finite sample rejection probability α_T in the asymptotic tests using the following Monte Carlo estimator:

$$\hat{\alpha}_T := \frac{1}{S} \sum_{b=1}^{S} \mathbb{I}(LR_{T,s} > \chi^2_{r,1-\alpha}) \ , \tag{28}$$

where $LR_{T,s}$ is the value of the likelihood ratio statistic in each simulated bootstrap sample $s = 1, \ldots, S$ and $\mathbb{I}(A)$ denotes the indicator function of event A.[22] We find that the estimated probability of rejecting null hypothesis (14) (null hypothesis (13)) by chance alone in a test of asymptotic level $\alpha = 5\%$ is as large as $\hat{\alpha}_T = 25.8\%$ ($\hat{\alpha}_T = 60.5\%$); see the first row of Table 3. This confirms the tendency of asymptotic tests to over-reject in finite samples.

In order to quantify the finite sample rejection probability α_T^* in our bootstrap test, we apply the bootstrap testing method to the simulated data. First, for each simulated time series

[21] We avoid a parametric Monte Carlo simulation with jointly normal dividend and price-dividend ratios, because for several null hypotheses relevant to our analysis we have found the finite sample distribution of fitted model residuals under H_0 to deviate from normality. Such deviations from normality can emerge in finite samples even if the true shocks are conditionally Gaussian, e.g., in presence of a neglected heteroskedasticity or other forms of neglected time series dependence. See Section 4.5.

[22] $\mathbb{I}(A)(\omega) = 1$ ($\mathbb{I}(A)(\omega) = 0$) if and only if $\omega \in A$ ($\omega \notin A$).

indexed by $s = 1, \ldots, S$, we compute a bootstrap distribution of likelihood ratio statistics $LR^*_{T,b,s}$, $b = 1, \ldots, B$, which yields a corresponding bootstrap p-value:

$$p^*(LR_{T,s}) = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}(LR^*_{T,b,s} > LR_{T,s}) \ .$$

This bootstrap p-value is computed according to the algorithm in Section 4.3. Finally, we compute probability α^*_T by the following Monte Carlo estimator:

$$\hat{\alpha}^*_T := \frac{1}{S} \sum_{b=1}^{S} \mathbb{I}(p^*(LR_{T,s}) < \alpha) \ . \tag{29}$$

Overall, this second Monte Carlo simulation is based on a double-bootstrap simulation scheme with $2S(B+1)$ estimations of the parameters in the present-value model, which is a computationally demanding procedure. We present our Monte Carlo results for the parameter choices $S = 200$, $B = 99$ and an optimal bootstrap block size of 2.[23] Other parameter choices produce similar results. For null hypothesis (13), we obtain $\hat{\alpha}^*_T = 5\%$, which is exactly the nominal level of the test ($\alpha = 5\%$). For null hypothesis (14), we obtain an empirical size $\hat{\alpha}^*_T = 8\%$, which is clearly much closer to the nominal level ($\alpha = 5\%$) than the rejection frequency $\hat{\alpha}_T = 25.8\%$ in the asymptotic test. In summary, the bootstrap test clearly improves on the finite sample properties of the asymptotic test, by producing a much better control of the probability to detect predictive relations by chance alone.

4.5 Why a nonparametric bootstrap?

The simplest way to simulate random samples from the present-value model, under the relevant null hypothesis, is by means of a parametric Monte Carlo simulation, e.g., under a normality assumption for the innovations in dividends and returns. In contrast, our nonparametric bootstrap approach renders the inference less dependent on such distributional assumptions. This is an important property, because the finite sample distribution of estimated shocks in the present-value model can deviate from normality for some of the relevant null hypotheses.[24]

To quantify such finite sample deviations from normality, we can test the normality of

[23] We apply a data driven calibration method for the selection of the block size, similar to the one introduced in Romano and Wolf (2001) and Camponovo, Scaillet, and Trojani (2009). We choose the block size that minimizes the difference between empirical and nominal size of the bootstrap test for no return predictability.

[24] Unreported empirical evidence also shows that mean, standard deviation and autocorrelation of dividend growth, price-dividend ratio and returns implied by our nonparametric bootstrap tend to be closer to the empirical sample moments than those obtained using a parametric Monte Carlo simulation.

estimated dividend and price-dividend ratio shocks in the innovation form representation of Section 4.1. Jarque and Bera (1987) Lagrange multiplier test is among the most common tests of normality. The null hypothesis of 0 sample skewness and 3 sample kurtosis is rejected whenever the statistic

$$JB = \frac{T}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \overset{a}{\sim} \chi^2(2), \tag{30}$$

exceeds the test critical value, where S and K are sample skewness and sample kurtosis, respectively, and T is the sample size.[25]

Figure 2 plots the p-values of the JB test applied to the standardized dividend and price-dividend innovations, in the unconstrained present-value model and under several null hypotheses. We find that the p-values for dividend and price-dividend ratio shocks in the unconstrained estimation are 40.01% and 4.54%, respectively, already highlighting some degree of error non normality. Under the null of constant expected returns and constant expected dividend growth, the p-value of the test for the dividend innovation is much lower, with values of 1.43% and 6.09%, respectively. In contrast, we never reject the null of normality of the price-dividend ratio innovations in these cases. This evidence indicates that the form of the distribution of the fitted innovations in the present-value model is very sensitive to the particular null hypothesis being tested.

Note that deviations from normality can be found in finite samples even if the true shocks are conditionally Gaussian, e.g., in presence of a neglected time series dependence in the underlying dynamics. To show this concretely, we generate samples of dividend growth, returns and price-dividend ratios of the same length as the one of the observed data, i.e., 65 years, using a version of the present-value model in Piatti and Trojani (2017), which includes a simple specification for the time variation of return and dividend growth risks in a setting with Gaussian shocks. Then we estimate the benchmark present-value model with constant risks for each simulated sample and study the properties of the filtered residuals. We find that the null hypothesis of normality is rejected in more than 50% of the cases, for both the standardised dividend growth and price-dividend ratio innovations. This result suggests that the failure of the asymptotic test in finite samples could be related to a neglected time series dependence in dividend growth and return shocks. As our bootstrap test provides a more robust inference in presence of a neglected time series dependence, it allows us to avoid taking a stance on the particular form of such time

[25] While the asymptotic critical values follow from a chi-squared distribution with two degrees of freedom, it has been noted (e.g., in Deb and Sefton (1996)) that the small-sample quantiles of the test statistic are quite different from their asymptotic counterparts. Therefore, we interpolate p-values using critical values computed by Monte Carlo simulation, as provided by the Matlab function *jbtest*.

series dependence or the finite sample distribution of model shocks.

In summary, we find that a nonparametric bootstrap approach, which does not rely on distributional assumptions about dividend and price dividend ratio shocks, is more appropriate for robustly testing predictability hypotheses in the benchmark present-value model. In such a setting, the Kalman filter remains valid for error distributions different from the normal and it produces the best linear filter, even if global optimality is lost. For estimation purposes, consistency is preserved whenever the first two conditional moments implied by the filter are correctly specified, with other distributional assumptions beyond this being immaterial.

A second concern in the stock return predictability literature is the high persistence of the dividend yield. Even though when the dividend price ratio follows a nearly-integrated process the bootstrap is not valid anymore, this aspect is not a first-order issue in our analysis, due to the relatively low yearly autocorrelation of the dividend yield, which is about 0.91. In nearly integrated frameworks, a valid inference can be obtained using different resampling schemes from the one used in this paper, e.g., based on different versions of the subsampling; see Andrews and Guggenberger (2010), among others.

4.6 Understanding the difference with other informal bootstrap approaches

A key feature of our nonparametric bootstrap approach is that it is based on a resampling of the standardized residuals $\{\hat{e}_t := S_t^{-1/2}(\hat{\theta})\eta_t(\hat{\theta})\}_{t=1}^T$ in the observable dynamics of the innovation form representation (17)-(18) implied by the present-value model. As we show below, this feature is essential to obtain formal bootstrap validity. A different informal bootstrap approach, which might appear intuitive at first sight, could rely on bootstrapping the filtered residuals ε_t^X of the unobservable state space dynamics in equation (15).[26] However, this alternative approach creates an asymptotic bias that makes the resulting asymptotic bootstrap inference invalid.

We formally prove the asymptotic validity of our bootstrap tests in Section 4.3.[27] To understand the bias produced by the alternative approach, it is useful to recall a key necessary condition for bootstrap consistency, i.e., that the expected log likelihood under the bootstrap distribution is maximized at the sample parameter estimate $\hat{\theta}$:

$$\left.\frac{\partial E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right]}{\partial \theta}\right|_{\theta=\hat{\theta}} = 0, \tag{31}$$

[26] Rytchkov (2012), among others, applies a version of the latter approach to test in a robustness check the null hypothesis of constant expected returns in a present-value setting.

[27] See also Section 4.2 and Stoffer and Wall (1991) for the asymptotic validity of the resampling scheme applied to the innovation form representation.

where E^* denotes the expectation under the bootstrap distribution. The bootstrap applied directly to the estimated standardized innovations e_t enforces condition (31) since:[28]

$$E^* \left[\mathcal{L} \left(\theta, \{Y_t^*\}_{t=1}^T \right) \right] = \mathcal{L} \left(\theta, \{Y_t\}_{t=1}^T \right). \tag{32}$$

In contrast, the alternative bootstrap approach based on the filtered state dynamics does not. To show this, it is convenient to write the log-likelihood $\mathcal{L} \left(\theta, \{Y_t\}_{t=1}^T \right)$ as a function of the state innovations ε_t^X, which are those bootstrapped under the second approach. The Kalman filter in Appendix B yields $\varepsilon_{t,t}^X = \Gamma' \mathcal{K}_t \eta_t$, where $\mathcal{K}_t$ is the Kalman gain. Therefore,

$$\eta_t = \left(\Gamma' \mathcal{K}_t \right)_L^{-1} \varepsilon_{t,t}^X, \tag{33}$$

with $\left(\Gamma' \mathcal{K}_t \right)_L^{-1}$ the left inverse of the 3×2 matrix $\Gamma' \mathcal{K}_t$.[29] Substituting this expression for η_t in the log-likelihood function we can show that the difference between $\mathcal{L} \left(\theta, \{Y_t\}_{t=1}^T \right)$ and $E^* \left[\mathcal{L} \left(\theta, \{Y_t^*\}_{t=1}^T \right) \right]$ under the second bootstrap approach is explicitly given as

$$-\frac{1}{2} \sum_{t=1}^{T} vec \left(\left(\Gamma' \mathcal{K}_t \right)_L^{-1\prime} S_t^{-1} \left(\Gamma' \mathcal{K}_t \right)_L^{-1} \right)' \left[vec \left(\varepsilon_{t,t}^X \varepsilon_{t,t}^{X\prime} \right) - \frac{1}{T} \sum_{t=1}^{T} vec \left(\varepsilon_{t,t}^X \varepsilon_{t,t}^{X\prime} \right) \right] \neq 0. \tag{34}$$

It is not zero because differences from the mean in the square parenthesis are not just summed over the T observations, but first weighted by a term that varies with t and depends on the Kalman gain and the filtered covariance of the prediction errors. Therefore, under the alternative bootstrap approach, $E^* \left[\mathcal{L} \left(\theta, \{Y_t^*\}_{t=1}^T \right) \right]$ is not necessarily maximized in $\hat{\theta}$, which means that bootstrap consistency cannot be guaranteed.

We can easily quantify the asymptotic bias generated by the second bootstrap approach in our application. To this end, we first simulate $B = 500$ bootstrap samples of the observables $\{Y_t^*\}_{t=1}^T$ under both bootstrap approaches, based on the unconstrained parameter estimate $\hat{\theta}$ in Table 1. We then compute the expected log-likelihood $E^* \left[\mathcal{L} \left(\theta, \{Y_t^*\}_{t=1}^T \right) \right]$ under both bootstrap distributions, for various parameter values θ on a grid in a neighborhood of $\hat{\theta}$.

As expected, we find that when $\{Y_t^*\}_{t=1}^T$ is bootstrapped under our approach the expected log-likelihood is always maximized at the parameter value in the grid that is closest to the sample estimate $\hat{\theta}$. In contrast, the expected log-likelihood under the second approach is maximized at a parameter value that can be quite distant from $\hat{\theta}$.

[28] The analytical details of all derivations in this subsection are provided in Appendix D.

[29] The left inverse of a $m \times n$ matrix A with $m > n$ and rank n is the $n \times m$ matrix A_L^{-1} such that $A_L^{-1} A = I_n$ and is computed as $A_L^{-1} = (A'A)^{-1} A'$.

Figure 3 illustrates these biases. The largest biases arise for parameters γ_1, δ_1, and σ_μ. Indeed, the maximum of the expected log-likelihood under the second approach yields parameter values 0.211, 0.867 and 0.010, respectively, which are quite different from the corresponding components in parameter vector $\hat{\theta}$, which are 0.304, 0.927 and 0.015. Also for parameter σ_D (bottom panel) the bias is quite impressive, as the expected log-likelihood under the second bootstrap distribution is maximized at 0.0005 instead of 0.002.

These biases obviously have first-order implications also for the distribution of simulated observables $\{Y_t^*\}_{t=1}^T$ under the second bootstrap approach. Figure 4 illustrates how the distribution of observables generated by bootstrapping the filtered innovations in the latent states can fail to capture the characteristics of observed dividends and returns. In particular, a 95% confidence intervals of the simulated return volatility and of the correlation between returns and dividend growth does not include the corresponding values in the data. On the contrary, the distribution of average, standard deviation, autocorrelation and correlation of the simulated observables using the innovation-form approach is always centred around the true values. Consistent with this evidence, while our bootstrap test does not reject the null hypothesis of no dividend growth predictability, the alternative (biased) bootstrap approach would reject the hypothesis.

5 How much Predictability?

The weak evidence of dividend growth predictability produced by bootstrap likelihood-ratio tests raises the question of the interpretation of the large R-squared ($R^2_{div} = 17.58\%$) estimated in Section 3 for future dividends. Differently from standard predictive regressions, the asymptotic distribution of estimated R-squares in the present-value model is not known in closed-form. Therefore, the conventional asymptotic approach cannot be used, e.g., to quantify the probability of estimating large R-squares because of chance alone. In contrast, our bootstrap methodology can be applied with no major modification to consistently estimate such probability, under the assumption that an asymptotic distribution for the estimated R-squares exists.

Using steps 1)-3) of the algorithm in Section 4.3, we compute bootstrap estimates of parameter θ in the present-value model and obtain the bootstrap distribution of estimated R-squared statistics under a given null hypothesis H_0. Figure 5 displays the histogram of the bootstrap distribution of estimated R-squares for future returns and future dividend growth, under the null hypotheses of a constant expected dividend growth. Even though the model-implied R-squared for dividend growth under H_0 is 0%, we find that the variability of estimated R-squares is quite large. For instance, while the median estimated R^2_{Div} value is 6.02% and the most frequently estimated R-squared value is 0%, the probability of estimating a dividend R-squared of at least 17.58%, as in the data, is 11.3%. These findings highlight that finite-sample variability is important for appropriately interpreting the finite-sample distribution of estimated R-squares, as large R-squares as in the data can arise by chance alone, in a present-value model where dividend predictability is absent.

Finally, Figure 6 displays the histogram of the bootstrap distribution of estimated R-squares for future returns and future dividend growth, under the null hypothesis of a constant expected return. In this case the probability of estimating a return R-squared larger than in the data, i.e., 8.82%, is only 2.9%, supporting the return predictability evidence. However, the standard deviation of the R^2 distribution is still large under this null hypothesis and can be a concern when interpreting the amount of predictability within a present-value model using the R^2 metric.

5.1 Out-of-sample predictability

All R^2 values reported in the previous sections are estimated using in-sample data. From the perspective of real-time predictability, out-of-sample prediction is an additional important aspect. For instance, Goyal and Welch (2008) study the out-of-sample predictive power of a large set of variables for market returns and find that most of them perform worse than the historical mean.

Following Campbell and Thompson (2008) and Goyal and Welch (2008), the incremental out-of-sample predictive power for returns and dividend growth in the present-value model of Section 2 can be estimated using the metrics:

$$R^2_{Ret,OS} = 1 - \frac{\sum_{t=0}^{T} \left(r_{t+1} - \tilde{\mu}_t\right)^2}{\sum_{t=0}^{T} \left(r_{t+1} - \bar{r}_t\right)^2} , \tag{35}$$

$$R^2_{Div,OS} = 1 - \frac{\sum_{t=0}^{T} \left(\Delta d_{t+1} - \tilde{g}_t\right)^2}{\sum_{t=0}^{T} \left(\Delta d_{t+1} - \overline{\Delta d}_t\right)^2} , \tag{36}$$

where $\tilde{\mu}_t$ and $\tilde{g}_t$ are the estimated expected return and expected dividend growth in the present-value model, using observations up to time t, while $\bar{r}_t$ and $\overline{\Delta d}_t$ are the sample means of returns and dividend growth using data up to time t.

We estimate the degree of out-of-sample predictability according to measures (35) and (36), using an out-of-sample period starting in 1985. Standard predictive regressions of returns and dividend growth on the lagged price-dividend ratio yield $R^2_{Ret,OS} = -12.32\%$ and $R^2_{Div,OS} = -4.38\%$, while we obtain $R^2_{Ret,OS} = -7.31\%$ and $R^2_{Div,OS} = 5.88\%$ for the present-value model.[30] Thus, the point estimates for the benchmark present-value model might indicate an incremental degree of out-of-sample predictability for dividend growth with respect to the sample mean forecast.

Using a natural adaptation of our bootstrap method, we can estimate the distribution of out-of-sample R-squares (35) and (36) under the null of no return or dividend predictability. In particular, the distribution of the out-of-sample R-squares of returns and dividend growth under the null hypothesis H_0 is computed based on the following algorithm:

1) Using the estimated model parameters obtained using the first T years of data, under the null hypothesis H_0, denoted $\hat{\theta}_{T,0}$, construct the (constrained) time series of standardized innovations $\{\hat{e}_{0t}\}_{t=1}^{T}$, and a bootstrap sample of observations, $\{Y_t^*\}_{t=1}^{T}$ as in steps 1)-3) in Section 4.3.

2) Using bootstrap sample $\{Y_t^*\}_{t=1}^{T}$, compute unconstrained maximum likelihood point estimates $\hat{\theta}_T^*$, by maximizing the log likelihood function $\log \mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^{T}\right)$, without imposing null hypothesis H_0.

3) Based on estimated parameters $\hat{\theta}_T^*$ and filtered state using data until time T, compute the expected return and dividend growth for year $T+1$, $\tilde{\mu}_T$ and $\tilde{g}_T$, respectively.

4) Repeat steps 1)-3) for $T = T_{in}, \ldots, T_{max} - 1$, where T_{in} is the minimum length of the in-sample period and T_{max} is the length of the full sample of data.[31]

5) The out-of-sample R^2 statistics for returns and dividend growth are computed as

$$\begin{aligned} R^2_{Ret,OS} &= 1 - \frac{\sum_{T=T_{in}}^{T_{max}-1} \left(r_{T+1} - \tilde{\mu}_T\right)^2}{\sum_{T=T_{in}}^{T_{max}-1} \left(r_{T+1} - \bar{r}_T\right)^2}, \\ R^2_{Div,OS} &= 1 - \frac{\sum_{T=T_{in}}^{T_{max}-1} \left(\Delta d_{T+1} - \tilde{g}_T\right)^2}{\sum_{t=T_{in}}^{T_{max}-1} \left(\Delta d_{T+1} - \overline{\Delta d}_T\right)^2}, \end{aligned}$$

[30] Precisely, we use data between 1946 and 1985 to estimate the parameters of the model and compute expected return and expected dividend growth for 1986, which are compared to the realized return and dividend growth in the same year. We then use data between 1946 and 1986 to compute predictions for 1987 and proceed in this way until the end of the sample. Using data from 1946 to 2007 and starting the out-of-sample computations in 1972, Binsbergen and Koijen (2010) find $R^2_{Ret,OS} = 1.06\%$ and $R^2_{Div,OS} = 5.76\%$.

[31] We start our out-of-sample computations in 1985, which means that the first estimation is done using $T_{in} = 40$ years of data, and the length of the full sample in our case is $T_{max} = 65$ years.

where $\bar{r}_T$ and $\overline{\Delta d}_T$ are historical means or returns and dividend growth up until time T.

6) Repeat steps 1)-5) a large number of times, B, to obtain a collection of bootstrap values of the out-of-sample R^2 statistics. The empirical distribution of these values provides an approximation of the distribution of the $R^2_{Ret,OS}$ and $R^2_{Div,OS}$ statistics under the null hypothesis H_0.

This procedure borrows from Rodriguez and Ruiz (2009), who show how to compute non-parametric bootstrap prediction intervals in state space models, while taking into account the uncertainty linked to parameter estimation and not resorting to parametric assumptions for the shock distribution in the model.

Given the variability of estimated in-sample R^2_{Div} values highlighted previously, it is plausible that the out-of-sample R-squared distribution might inherit similar features. Figure 7 illustrates the properties of the bootstrap distributions of out-of-sample R-squares (35) and (36), generated under the null hypothesis of constant expected cash flow growth in the present-value model. Both distributions imply a large variability of estimated out-of-sample measures of predictability for returns (upper panel) and dividend growth (lower panel). Even though under H_0 expected returns are time-varying, the estimated $R^2_{Ret,OS}$ distribution puts a large mass in regions where no evidence of incremental predictability is estimated. Moreover, despite the absence of dividend predictability under the null, the distribution of estimated $R^2_{Div,OS}$ puts a significant mass of about 32% in regions of positive $R^2_{Div,OS}$ values, with a probability of estimating an out-of-sample R-squared for dividends at least as large as in the data that is about 10%.

Using the same procedure, we obtain the bootstrap distributions of out-of-sample R-squares (35) and (36), generated under the null hypothesis of constant expected return (see Figure 8). As in the previous case, both distributions imply a large variability of estimated out-of-sample R-squared. The negative out-of-sample R-squared for returns ($R^2_{Ret,OS} = -7.31\%$) is not significantly different from zero, with a probability of estimating an even smaller value that is about 34%.

Overall, these findings show that the conclusions produced by estimated common measures of out-of-sample predictability in present-value models have to be taken with caution and put in relation to the finite-sample variability of these quantities under the null of no predictability. On the one side, the limited amount of data information available can lead to a difficulty in detecting predictive relations for returns when they are there. On the other side, high out-of-sample R-squares for dividends can arise by chance alone, in a setting with constant expected dividend growth. In this respect, our nonparametric bootstrap approach provides a useful tool to better interpret also the information provided by estimated out-of-sample measures of predictability.

6 Robustness to Different Samples and Cash-flow Proxies

This Section discusses the robustness of our main findings with respect to the inclusion of the prewar sample (Section 6.1) and to the use of different cash-flow proxies (Section 6.2).

6.1 A tale of two periods

The time series of US aggregate dividend growth for the prewar and the postwar periods exhibit substantially different properties, suggesting a potential structural break in the dividend process between these two sample periods. While tests based on standard predictive regressions for the postwar sample find no evidence of dividend predictability, the evidence is reversed for the prewar sample; see Chen (2009), among others. Therefore, it is natural to test whether our bootstrap tests of predictability in present-value models can produce consistent results for both the prewar and postwar samples.

We can parsimoniously account for the structural break in the parameters of the present-value model, between the prewar and the postwar samples, by allowing the persistence and the variability of expected returns and dividend growth, parametrized by δ_1, γ_1, σ_g, σ_μ, σ_d, $\rho_{g\mu}$ and $\rho_{\mu d}$, respectively, to differ before and after 1946. The parameter estimates and p-values for the LR tests of predictability in the prewar and postwar samples are collected in Table 4.

The estimation results support the evidence of a regime shift in the parameters of the dividend process, approximately in 1946, since expected dividend growth is estimated as much more volatile in the prewar sample. Similarly, expected returns are estimated as much less persistent before 1946. Actually, the point estimate of γ_1 in the prewar sample, i.e. γ_{1p}, is even negative, consistent with a negative autocorrelation of dividend growth between 1927 and 1946. The asymptotic LR test clearly rejects the null of no dividend predictability in the prewar and the postwar samples, with a p-value of 0% and 0.06%, respectively. The asymptotic test also rejects the null of no return predictability for the postwar sample, with a p-value of 0%, but not in the prewar sample, with a p-value of 9.45%.

The results of the bootstrap test again indicate that asymptotic tests in present-value models tend to overreject the null of no dividend growth predictability, since all bootstrap p-values are larger than the p-values of the corresponding asymptotic test. However, while the null of no dividend predictability cannot be rejected in the postwar sample at the 10% significance level, it is clearly rejected in the prewar sample. In contrast, for the null of no return predictability, the asymptotic and bootstrap test draw the same conclusions for both prewar and postwar samples.

In summary, when accounting for a regime shift in the parameters of the dividend and return processes, our bootstrap test reconciles the conclusions produced by standard predictive regressions and present-value models, producing dividend predictability findings consistent with the evidence in Chen (2009), among others. The findings for the prewar sample also indicate that our bootstrap test has power to detect dividend and return predictability structures based on a quite limited data information, as it rejects the null of no dividend growth predictability in the prewar sample using the information provided by only about 20 yearly observations.

6.2 Measuring cash flows: dividends or payout?

A branch of the recent literature on predictability argues that dividends are not a good measure of total payout to investors and considers dividend growth and valuation ratios adjusted for stock repurchases and (potentially) issuances. The underlying motivation refers to the fact that firms may (partially) substitute dividends with repurchases, due e.g. to taxation or psychological reasons (dividend smoothing). This alternative way of measuring aggregate dividends and valuation ratios reflects the view of a representative investor holding the whole market (see, e.g., Bansal and Yaron (2011)), while in our paper we hold the traditional portfolio view of an investor holding one share forever.

Boudoukh, Michaely, Richardson, and Roberts (2007), among others,[32] find that total and net payout yields have a stronger predictive power for market returns than the dividend yield. More recently, Sabbatucci (2015) argues that including M&A cash flows in dividend measures increases both return and dividend growth predictability evidence from standard predictive regressions. It is therefore interesting to look at the effects of such alternative cash-flow measures for the estimation and testing of predictability hypotheses in present-value models.

Annual time series of repurchases and issuances from 1946 to 2003 are obtained from the dataset constructed by Boudoukh, Michaely, Richardson, and Roberts (2007),[33] and we thank

[32] See, e.g., Robertson and Wright (2006), Larrain and Yogo (2008) and Chen, Da, and Priestley (2012).

[33] The series are drawn from Michael Roberts' website: http://finance.wharton.upenn.edu/∼mrrobert/. Data are available since 1926, but we use data starting from 1946 for comparison with our main results in the paper. The repurchase yield is only available beginning in 1971, thus repurchases are assumed to be zero until 1970. Repurchases were of negligible size until the mid 1980s, thus this lack of data is likely to have little effect on the results. Note that annual yearly returns in Boudoukh, Michaely, Richardson, and Roberts (2007) implicitly

Serhiy Kozak for providing us with log dividend yield and dividend growth measures including M&A distribution from 1946 to 2010.[34] Figure 9 shows the dynamics of yearly cash-flow growth (upper panel) and valuation ratios (lower panel) using different measures of cash-flow: dividend (blue line), total payout (dividend plus repurchases, red line), net payout (dividend plus repurchases minus issuances, green line) and cash M&A (cash dividend plus M&A cash flows to shareholders, magenta line).[35] While dividend, total payout and cash M&A share similar patterns (correlations are around 60%), issuances seem to be more related to returns, likely because of strategic firm behaviour. Therefore, we focus our robustness checks on the total payout and cash M&A as alternatives for cash dividends. Tables 5 and 6 report the results of the estimation of the present-value model using total payout data and cash M&A, respectively, without imposing any constraint and under the null hypotheses of no return predictability and no dividend predictability. The tables also shows the values of the standard likelihood ratio statistics and the corresponding asymptotic and bootstrap p-values, in percentage. The parameter estimates are qualitatively similar to those shown in the main text and the conclusions concerning the main predictability features are also similar: bootstrap p-values are always larger than those obtained with asymptotic critical values, and they are in line with standard predictive regression results. In fact, standard predictive regressions of cash flow growth on the log price dividend ratio adjusted for M&A cash flows yield a marginally significant regression coefficient (of the right sign) and an adjusted R-squared of 3.76% (see Table 7). Using the bootstrap test, we also reject the null hypothesis of no dividend predictability at a 10% level but not at 5%. Using total payout measures, the estimated slope coefficient in standard predictive regressions is insignificant and has the wrong sign, the adjusted R-squared of the regression is only -0.29%, and our bootstrap test cannot reject the hypothesis of no dividend growth predictability with a p-value of 16%.

7 Conclusion

Univariate predictive regressions of future returns and dividend growth on predictive variables including the lagged price-dividend ratio have produced no apparent evidence of dividend pre-

assume market reinvestment of dividends. Thus, we adjust them to be consistent with our assumption of monthly dividends reinvested at the risk-free rate.

[34] They are constructed using a CRSP-based measure of M&A cash dividends, while Sabbatucci (2015) uses SDC Platinum data after 1980, which includes M&A of non-listed companies.

[35] Returns, as well as the Campbell-Shiller approximation, should also be adjusted for repurchases. However to perform these adjustments, we would need information on the time-varying number of repurchased shares. Therefore, for simplicity we abstract from the issue of time-varying capitalization.

dictability in the postwar sample, suggesting that price-dividend ratios have mostly varied because of discount rate shocks in that period. In contrast, latent variable approaches within present-value models, which parsimoniously incorporate information from the joint time-series of dividends and returns, have found a stronger evidence of a time-varying expected dividend growth.

A natural explanation for these contrasting conclusions is the error-in-variable (EIV) problem inherent to predictability studies. This paper provides sharp evidence for a different explanation, linked to the so far unexplored finite-sample properties of conventional tests of predictability in models with latent return and dividend expectations.

First, we show that conventional tests frequently reject the null of no dividend predictability because of chance alone. Moreover, we find that large estimated R-squares for dividends can arise by chance alone, even under the null of a constant expected dividend growth. These findings stress the importance of combining a pure estimation approach with a reliable testing method, when testing and quantifying the actual degree of predictability within present-value models.

Second, in order to introduce a general and more reliable testing approach, we propose a class of nonparametric bootstrap tests of predictability hypotheses in present-value models, by applying the bootstrap to the innovation form of the Kalman filter, generated under the relevant null hypothesis. We prove that the bootstrap tests imply a valid asymptotic inference and demonstrate their improved properties in finite samples.

Third, we apply our bootstrap tests to US stock market data, based on a variety of specifications of the predictive information set. In contrast to the results implied by standard asymptotic tests, we find a significant evidence in favour of time-varying expected returns, both in the prewar and the postwar samples, no evidence of time-varying dividend expectations in the postwar sample and a strong dividend predictability in the prewar sample. This evidence reconciles the diverging conclusions in the literature.

We finally propose a natural modification of our bootstrap testing method, which can be used also to test the presence of out-of-sample predictability, while controlling the probability of detecting predictive relations by chance alone. We find that the conclusions produced by estimated common measures of out-of-sample predictability in present-value models have to be taken with caution and need to be set in relation to the finite-sample variability of these quantities under the null of no predictability.

From a broader methodological perspective, our bootstrap testing approach and our results have implication for a number of potentially more general aspects. First, while our bootstrap tests can help to control more systematically the probability of rejecting a null hypothesis by chance alone, our results also indicate that the information generated by the joint time series of stock market returns and dividends might be insufficient to reliably identify time-variations in dividend expectations, i.e., tests of dividend predictability in such settings may have a low power.

A low power might arise because of the short time series available for many predictability studies or because market price-dividend ratios aggregate into a single observable signal the expectations of future dividends for different horizons, which are potentially difficult to identify separately. As shown in Binsbergen, Brandt, and Koijen (2012) and Binsbergen, Hueskes, Koijen, and Vrugt (2013), a more direct identification of dividend expectations at distinct horizons can rely on the equity yield of dividend strips, which are dividend claims for single maturities. Annual dividend growth is strongly predictable in the period from October 2002 to April 2011, with univariate predictive regression R^2s between 48% for the 5 year yield and 76% for the 1 year yield. This evidence suggests that dividend strip information can potentially improve the power of tests of dividend predictability more generally. Tests of predictability in benchmark present-value models may imply a low power also because of some form of model misspecification generated, e.g., by an unmodeled heteroskedasticity or by measurement errors in the observable variables. Piatti and Trojani (2017) and Schorfheide, Song, and Yaron (2018) study tests of predictability in more general present value models incorporating heteroskedasticity in the state dynamics and measurement errors.

Finally, our bootstrap testing method is in principle applicable more generally, in order to more reliably test the relevant null hypotheses in models estimated by a latent variable approach using their state space form. For instance, an interesting application could be the analysis of multi-horizon and multi-scale predictability. However, such settings would require a formal study of bootstrap validity under more general state dynamics than in this paper, in order to appropriately model multiple predictability scales.[36]

[36] Recent evidence in the literature shows that the aggregation properties of linear ARMA processes are not well-suited to model the multiple scales of return predictability (see, e.g., Bandi, Perron, Tamoni, and Tebaldi (2018), Ortu, Tamoni, and Tebaldi (2013) and Bianchi and Tamoni (2016)). We thank an anonymous referee for suggesting this interesting extension.

References

Amihud, Y. and C. Hurvich (2004). Predictive regression: a reduced bias estimation method. *Journal of Financial and Quantitative Analysis 39*, 813–841.

Amihud, Y., C. Hurvich, and Y. Whang (2009). Multiple-predictor regressions: Hypothesis testing. *Review of Financial Studies 22*, 413–434.

Andrews, D. (2002). Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica 70*(6), 119–162.

Andrews, D. and P. Guggenberger (2010). Application of Subsampling, Hybrid and Size-Correction Methods. *Journal of Econometrics 158*, 285–305.

Ang, A. and G. Bekaert (2007). Stock return predictability: Is it there? *Review of Financial Studies 20*(3), 651–708.

Bandi, F. M., B. Perron, A. Tamoni, and C. Tebaldi (2018). The scale of predictability. *Journal of Econometrics, forthcoming*.

Bansal, R. and A. Yaron (2011). The asset pricing-macro nexus and return-cash flow predictability. Working Paper. Fuqua School of Business, Duke University, Wharton School, University of Pennsylvania and NBER.

Beran, R. (1988). Prepivoting test statistics. *Journal of the American Statistical Association 83*, 687–697.

Bianchi, D. and A. Tamoni (2016). The dynamics of expected returns: Evidence from multi-scale time series modeling. *Working Paper*.

Binsbergen, J., M. Brandt, and R. S. Koijen (2012). On the timing and pricing of dividends. *American Economic Review 102*(4), 1596–1618.

Binsbergen, J., W. H. Hueskes, R. S. Koijen, and E. B. Vrugt (2013). Equity yields. *Journal of Financial Economics 110*(3), 503–519.

Binsbergen, J. and R. S. Koijen (2010). Predictive regressions: A present-value approach. *Journal of Finance 65*(4), 1439–1471.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance 62*(2), 877–915.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal 101*, 157–179.

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*(3), 195–228.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies 21*(4), 1509–1531.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*(1), 27–60.

Camponovo, L., O. Scaillet, and F. Trojani (2009). Robust resampling methods for time series. Swiss Finance Institute Research Paper No. 09-38.

Chen, L. (2009). On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics 92*(1), 128–151.

Chen, L., Z. Da, and R. Priestley (2012). Dividend smoothing and predictability. *Management Science 58*(10), 1834–1853.

Cochrane, J. H. (1992). Explaining the variance of price-dividend ratios. *Review of Financial Studies 5*(2), 243–280.

Cochrane, J. H. (2008a). The dog that did not bark: A defense of return predictability. *Review of Financial Studies 21*(4), 1533–1576.

Cochrane, J. H. (2008b). State-space vs. var models for stock returns. Working Paper. University of Chicago.

Davidson, R. and J. MacKinnon (1999a). *Econometric Theory and Methods*. Oxford University Press.

Davidson, R. and J. MacKinnon (1999b). The size distortion of bootstrap tests. *Econometric Theory 15*, 361–376.

Deb, P. and M. Sefton (1996). The distribution of a lagrange multiplier test of normality. *Economics Letters 51*, 123–130.

Efron, B. (1979). Bootstrap methods: Another look at the jacknife. *Annals of Statistics 7*, 1–26.

Fama, E. F. and K. R. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22*(3), 3–25.

Favero, C. A., A. E. Gozluklu, and A. Tamoni (2011). Demographic trends, the dividend-price ratio, and the predictability of long-run stock market returns. *Journal of Financial and Quantitative Analysis 46*(5), 1493–1520.

Goyal, A. and I. Welch (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Hall, P. and J. Horowitz (1996). Bootstrap critical values for tests based on generalized-method-of-moment estimators. *Econometrica 64*(4), 891–916.

Jarque, C. M. and A. K. Bera (1987). A test for normality of observations and regression residuals. *International Statistical Review 55*(2), 163172.

Keim, D. B. and R. F. Stambaugh (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics 17*(2), 357–390.

Koijen, R. S. and S. Van Nieuwerburgh (2011). Predictability of returns and cash flows. *Annual Review of Financial Economics 3*, 467–491.

Kothari, S. P. and J. Shanken (1997). Book-to-market, dividend yield, and expected market returns: A time series analysis. *Journal of Financial Economics 44*(2), 169–203.

Larrain, B. and M. Yogo (2008). Does firm value move too much to be justified by subsequent changes in cash flow? *Journal of Financial Economics 87*(1), 200–226.

Lettau, M. and S. C. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance 3*, 815–849.

Lettau, M. and S. C. Ludvigson (2005). Expected returns and expected dividend growth. *Journal of Financial Economics 76*(3), 583–626.

Lettau, M. and S. Van Niewerburgh (2008). Reconciling the return predictability evidence. *Review of Financial Studies 21*(4), 1607–1652.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics 74*(2), 209–235.

Liung, L. and P. Caines (1979). Asymptotic normality of prediction error estimatorsfor approximatesystem models. *Stochastics 3*, 29–46.

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

Ortu, F., A. Tamoni, and C. Tebaldi (2013). Long-run risk and the persistence of consumption shocks. *Review of Financial Studies 26*(11), 2876–2915.

Pastor, L., M. Sinha, and B. Swaminathan (2008). Estimating the intertemporal risk-return tradeoff using the implied cost of capital. *Journal of Finance 63*(6), 2859–2897.

Pastor, L. and R. Stambaugh (2009). Predictive systems: Living with imperfect predictors. *Journal of Finance 64*(4), 1583–1628.

Piatti, I. and F. Trojani (2017). Predictable risks and predictive regression in present-value models. *Working Paper*.

Politis, D. N. and J. P. Romano (1992). *Exploring the Limits of Bootstrap*, Chapter A circular block-resampling procedure for stationary data, pp. 263–270. Wiley, New York.

Polk, C., S. Thompson, and T. Vuolteenaho (2006). Cross-sectional forecasts of the equity premium. *Journal of Financial Economics 81*, 101–141.

Robertson, D. and S. Wright (2006). Dividends, total cash flow to shareholders, and predictive return regressions. *Review of Economics and Statistics 88*(1), 91–99.

Rodriguez, A. and E. Ruiz (2009). Bootstrap prediction intervals in state-space models. *Journal of Time Series Analysis 30*(2), 167–178.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio Management 11*(1), 68–75.

Rytchkov, O. (2012). Filtering out expected dividends and expected returns. *Quarterly Journal of Finance 2*(3), 1250012.

Sabbatucci, R. (2015). Are dividends and stock returns predictable? new evidence using m&a cash flows. Working paper. University of California, San Diego.

Schiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity 2*, 457–498.

Schorfheide, F., D. Song, and A. Yaron (2018). Identifying long-run risks: A bayesian mixed-frequency approach. *Econometrica forthcoming.*

Spall, J. and K. Wall (1984). Asymptotic distribution theory for the kalman filter state estimator. *Communications in Statistics-Theory and Methods 13*, 1981–2003.

Stambaugh, R. (1999). Predictive regressions. *Journal of Financial Economics 54*(3), 375–421.

Stoffer, D. and K. Wall (1991). Bootstrapping state-space models: Gaussian maximum likelihood estimation and the kalman filter. *Journal of the American Statistical Association 86*, 1024–1033.

Torous, W., R. Valkanov, and S. Yan (2004). On predicting stock returns with nearly integrated explanatory variables. *Journal of Business 77*, 937–966.

Wachter, J. A. and M. Warusawitharana (2009). Predictable returns and asset allocation: Should a skeptical investor time the market? *Journal of Econometrics 148*(2), 162–178.

Yun, J. (2012). Evaluating predictors within a present-value framework. Working Paper. University of Chicago Booth School of Business.

A Price-dividend ratio

In this section we present the detailed derivation of Equation (7) in the text. From Campbell and Shiller (1988) we have

$$pd_t \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}. \tag{37}$$

By iterating this equation we find:

$$\begin{aligned} pd_t &\simeq \kappa + \rho(\kappa + \rho pd_{t+2} + \Delta d_{t+2} - r_{t+2}) + \Delta d_{t+1} - r_{t+1} \\ &= \sum_{j=0}^{\infty} \rho^j \kappa + \rho^{\infty} pd_{\infty} + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}) \\ &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}), \end{aligned} \tag{38}$$

assuming that $\rho^{\infty} pd_{\infty} = \lim_{j\to\infty} \rho^j pd_{t+j} = 0$, at least in expectation. Then, we take expectation conditional to time t:

$$\begin{aligned} pd_t &\simeq \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[\Delta d_{t+j} - r_{t+j}] \\ &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[g_{t+j-1} - \mu_{t+j-1}] \\ &= \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^{j} E_t[g_{t+j} - \mu_{t+j}]. \end{aligned} \tag{39}$$

Iterating the dynamics of $\hat{\mu}_{t+1}$ and $\hat{g}_{t+1}$ and taking conditional expectation we find

$$E_t[\hat{\mu}_{t+j}] = \delta_1^j \hat{\mu}_t$$

and

$$E_t[\hat{g}_{t+j}] = \gamma_1^j \hat{g}_t.$$

Therefore,

$$\begin{aligned}
pd_t &\simeq \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^j [\gamma_0 + \gamma_1^j \hat{g}_t - \delta_0 - \delta_1^j \hat{\mu}_t] \\
&= \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho} + \frac{\hat{g}_t}{1-\rho\gamma_1} - \frac{\hat{\mu}_t}{1-\rho\delta_1} \\
&= A + B_2 \hat{g}_t - B_1 \hat{\mu}_t. \qquad (40)
\end{aligned}$$

The explicit expressions for the present-value coefficients A, B_1 and B_2 are the following:

$$\begin{aligned}
A &= \frac{\kappa + \gamma_0 - \delta_0}{1-\rho}, \\
B_1 &= \frac{1}{1-\rho\delta_1}, \\
B_2 &= \frac{1}{1-\rho\gamma_1}.
\end{aligned}$$

B Estimation Methodology

For estimation purposes, we cast the model in state space form, using demeaned state variables $\hat{\mu}_t \equiv \mu_t - \delta_0$ and $\hat{g}_t \equiv g_t - \gamma_0$. We obtain the following linear transition dynamics:

$$\hat{g}_{t+1} = \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \qquad (41)$$

$$\hat{\mu}_{t+1} = \delta_1 \hat{\mu}_t + \varepsilon^\mu_{t+1}. \qquad (42)$$

The observable variables are dividend growth Δd_{t+1} and the price-dividend ratio pd_{t+1}. Measurement equations for Δd_{t+1} and pd_{t+1} are derived from the model-implied expressions for dividend growth and price-dividend ratio. The measurement equation for dividend growth is given by (5) while log price-dividend ratio is given by (7). Note however that Equation (7) contains no error term, and as shown by Binsbergen and Koijen (2010), this feature can be exploited to reduce the number of transition equations in the model. By substituting the equation for pd_t in the measurement equation for dividend growth, we arrive at a final system with one transition equation, (41), and two measurement equations:

$$\Delta d_{t+1} = g_t + \varepsilon^d_{t+1}. \qquad (43)$$

$$pd_{t+1} = (1-\delta_1)A + B_2(\gamma_1 - \delta_1)\hat{g}_t + \delta_1 pd_t - B_1 \varepsilon^\mu_{t+1} + B_2 \varepsilon^g_{t+1}. \qquad (44)$$

We use the Kalman filter to derive the likelihood of the model and we estimate it using ML. The parameters to be estimated are the following:

$$\theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_d, \rho_{g\mu}, \rho_{\mu d}, \rho_{gd}).$$

We assume that expectation processes are stationary, therefore parameters δ_1 and γ_1 are bounded to be less than one in absolute value. The covariance matrix of the shocks, (6), has to be positive definite, thus σ_g, σ_μ and σ_d are constrained to be positive, while the correlation parameters are between -1 and 1.[37] Rytchkov (2012) shows that it is impossible to identify the whole covariance structure of shocks even when an infinitely long history of returns and dividends is given, but only one element of Σ must be fixed to identify the whole matrix. Thus, for identification purposes, we impose the constraint $\rho_{gd} = 0$, as in Binsbergen and Koijen (2010). Overall the model implies 9 free parameters to estimate. The estimation procedure is the following: We first define an expanded 4-dimensional state vector by the concatenation of the original state variable $\hat{g}$ and the process and observation noise random variables:

$$X_t = \begin{pmatrix} \hat{g}_{t-1} \\ \varepsilon_t^g \\ \varepsilon_t^\mu \\ \varepsilon_t^d \end{pmatrix},$$

which satisfies:

$$X_{t+1} = F X_t + \Gamma \varepsilon_{t+1}^X,$$

where

$$\varepsilon_{t+1}^X = \begin{pmatrix} \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^\mu \\ \varepsilon_{t+1}^d \end{pmatrix},$$

with conditional variance Σ, given in (6). Moreover,

$$F = \begin{bmatrix} \gamma_1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \Gamma = \begin{bmatrix} 0_{1\times 3} \\ I_3 \end{bmatrix},$$

[37]Moreover, the condition $\rho_{g\mu}^2 + \rho_{\mu d}^2 + \rho_{gd}^2 < 1$ has to hold for Σ to be positive definite

The measurement equation,

$$Y_t = \begin{pmatrix} \Delta d_t \\ pd_t \end{pmatrix},$$

is of the form

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t,$$

where

$$M_0 = \begin{bmatrix} \gamma_0 \\ (1-\delta_1)A \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & 0 \\ 0 & \delta_1 \end{bmatrix},$$

and

$$M_2 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ B_2(\gamma_1 - \delta_1) & B_2 & -B_1 & 0 \end{bmatrix}.$$

The steps of the filter algorithm are the following:

- Initialize with the unconditional mean and covariance of the expanded state:

$$\begin{aligned} X_{0,0} &= 0_{4\times 1}, \\ P_{0,0} &= E(X_t X_t'). \end{aligned}$$

- The time-update equations are

$$\begin{aligned} X_{t,t-1} &= F X_{t-1,t-1}, \\ P_{t,t-1} &= F P_{t-1,t-1} F' + \Gamma \Sigma \Gamma', \end{aligned}$$

- The prediction error η_t and the variance-covariance matrix of the measurement equations are then:

$$\begin{aligned} \eta_t &= Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t,t-1}, \\ S_t &= M_2 P_{t,t-1} M_2', \end{aligned} \tag{45}$$

  where Y_t is the observed value of the measurement equation at time t.

- Update filtering:

$$
\begin{aligned}
\mathcal{K}_t &= P_{t,t-1} M_2' S_t^{-1}, \\
X_{t,t} &= X_{t,t-1} + \mathcal{K}_t \eta_t, \\
P_{t,t} &= (I - \mathcal{K}_t M_2) P_{t,t-1},
\end{aligned}
$$

where $\mathcal{K}_t$ is the *Kalman gain*.

To estimate model parameters, θ, we define the log-likelihood for each time t, assuming normally distributed observation errors, as

$$
l_t(\theta) = -\frac{1}{2} \log |S_t| - \frac{1}{2} \eta_t' S_t^{-1} \eta_t,
$$

where η_t and S_t denote prediction error of the measurement series and the covariance of the measurement series, respectively, obtained from the KF. Model parameters are chosen to maximize the log-likelihood of the data series:

$$
\hat{\theta} \equiv \arg\max_{\Theta} \mathcal{L}\left(\theta, \{Y_t\}_{t=1}^{T}\right), \tag{46}
$$

with

$$
\mathcal{L}\left(\theta, \{Y_t\}_{t=1}^{T}\right) = \sum_{t=1}^{T} l_t(\theta),
$$

where T denotes the number of time periods in the sample of estimation.[38]

C Asymptotic Validity of the Bootstrap Likelihood Ratio Test

In this appendix we prove the validity of our nonparametric bootstrap likelihood ratio testing procedure, i.e., the equivalence in distribution of LR_T and LR_T^* in Equations (10) and (23), respectively, when $B, T \to \infty$, under the null hypothesis H_0. It is well known that if H_0 holds, as $T \to \infty$, LR_T follows a χ_r^2 distribution with r degrees of freedom, where r is the number of parameter constraints defining the null hypothesis H_0. Therefore, we only need to show that also LR_T^* is asymptotically χ_r^2 distributed.

Without loss of generality, let us consider for brevity the case in which the null hypothesis to be tested is formed by zero restrictions, i.e., some of the model parameters are equal to

[38] For yearly data, as in our application, T is the number of years in the sample.

zero. In such cases, the r restrictions can be written as $\theta_2 = 0_{r\times 1}$, where the parameter vector θ is partitioned as $\theta = [\theta_1' \quad \theta_2']'$, possibly after some reordering of the elements, where θ_1 is $(k-r)\times 1$ and θ_2 is $r\times 1$-dimensional.

Let $\hat{\theta}$ be the unconstrained ML estimator of θ, while the pseudo-true value of θ in the population under H_0 is denoted by $\theta^\star = [\theta_1^{\star\prime} \quad 0_{1\times r}]'$, where $\theta_1^\star$ is the *pseudo-true value* of θ_1, i.e., the maximum of the population expected log likelihood function with respect to θ_1 under the (potentially) incorrect assumption $H_0 : \theta_2 = 0_{r\times 1}$.

Stoffer and Wall (1991) show that nonparametric bootstrap applied to the (standardized) innovations $\{\hat{e}_t := S_t^{-1/2}(\hat{\theta})\eta_t(\hat{\theta})\}_{t=1}^T$ yields a distribution of bootstrap residuals $\{\hat{e}_t^*\}_{t=1}^T$, which can be used to compute a bootstrap distribution of ML estimators $\hat{\theta}^*$:

$$\hat{\theta}^* = \arg\max_{\Theta} \log \mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right) \ , \tag{47}$$

where the Monte Carlo sequence $\{Y_t^*\}_{t=1}^T$ is obtained by simulating dynamics (17)-(18) based on bootstrap residuals $\{\hat{e}_t^*\}_{t=1}^T$ (see steps 1)-3) in Section 4.3). Stoffer and Wall (1991) also provide an asymptotic justification of this procedure, showing, under general conditions, the equivalence in distribution of $\sqrt{T}(\hat{\theta}^* - \hat{\theta})$ and $\sqrt{T}(\hat{\theta} - \theta^\star)$ as $B, T \to \infty$, and assuming for simplicity $B = T$. For simplicity of notation we assume that the ML setting holds, but all results hold true with obvious modifications in a PML setting, using sandwich variance-covariance matrix estimators, see Stoffer and Wall (1991):

$$\sqrt{T}(\hat{\theta} - \theta^\star) \xrightarrow{d} N\left(0, \mathcal{I}(\theta^\star)^{-1}\right), \tag{48}$$

where $\mathcal{I}(\theta) = \operatorname{plim}_{T\to\infty} \frac{1}{T} E[-\partial^2 \log \mathcal{L}(\theta)/\partial\theta\partial\theta']$ is the asymptotic information matrix, and

$$\sqrt{T}(\hat{\theta}^* - \hat{\theta}) \xrightarrow{d} N\left(0, \mathcal{I}(\theta^\star)^{-1}\right). \tag{49}$$

The constrained ML estimator $\hat{\theta}_0$ can then be expressed as $\hat{\theta}_0 = \left[\hat{\theta}_1' \quad 0_{1\times r}\right]'$, and the asymptotic distribution of $\hat{\theta}_1$ is given by:[39]

$$\sqrt{T}(\hat{\theta}_1 - \theta_1^\star) \xrightarrow{d} N\left(0, \mathcal{I}_{11}(\theta_1^\star)^{-1}\right), \tag{50}$$

where $\mathcal{I}_{11}(.)$ is the $(k-r)\times(k-r)$ top left block of the asymptotic information matrix $\mathcal{I}(.)$ of

[39] See e.g. Davidson and MacKinnon (1999a), chapter 10.

the unrestricted model. Analogously, the constrained bootstrap Maximum Likelihood estimator $\hat{\theta}_0^*$ can be partitioned as $\hat{\theta}_0^* = \begin{bmatrix} \hat{\theta}_1^{*\prime} & 0_{1\times r} \end{bmatrix}'$ and its asymptotic distribution is given by:

$$\sqrt{T}(\hat{\theta}_1^* - \hat{\theta}_1) \xrightarrow{d} N\left(0, \mathcal{I}_{11}(\theta_1^\star)^{-1}\right). \tag{51}$$

For ease of notation, let us denote by $l(\theta, y)$ the log-likelihood of the model, i.e. $l(\theta, y) \equiv \log \mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right)$. Using a second order Taylor expansion around $\hat{\theta}^*$, the bootstrap log-likelihood $l(\hat{\theta}_0^*, y^*)$ can be written as

$$l(\hat{\theta}_0^*, y^*) = l(\hat{\theta}^*, y^*) - \frac{1}{2}(\hat{\theta}_0^* - \hat{\theta}^*)' H(\bar{\theta})(\hat{\theta}_0^* - \hat{\theta}^*). \tag{52}$$

where $H(.)$ is the Hessian matrix,[40] and $\bar{\theta} \in (\hat{\theta}_0^*, \hat{\theta}^*)$. Using (52), the bootstrap likelihood ratio statistics LR_T^* in (23) becomes

$$\begin{aligned} LR_T^* &= 2\left(l(\hat{\theta}^*, y^*) - l(\hat{\theta}_0^*, y^*)\right) \\ &= -(\hat{\theta}_0^* - \hat{\theta}^*)' H(\bar{\theta})(\hat{\theta}_0^* - \hat{\theta}^*). \end{aligned}$$

Consistency of $\hat{\theta}^*$ implies consistency of $\bar{\theta}$, and using information matrix inequality[41] we get:

$$LR_T^* \overset{a}{=} T(\hat{\theta}_0^* - \hat{\theta}^*)' \mathcal{I}(\theta^\star)(\hat{\theta}_0^* - \hat{\theta}^*). \tag{53}$$

Let we now define the score vector $g(\theta, y)$ of first derivatives of $l(\theta, y)$ with respect to the elements of θ,[42] and the asymptotic score vector $s \equiv plim T^{-1/2} g(\theta^\star, y)$. From a Taylor expansion of the likelihood equation $g(\hat{\theta}^*, y^*) = 0$ we obtain the following asymptotic equalities:

$$\begin{aligned} T^{1/2}(\hat{\theta}^* - \hat{\theta}) &\overset{a}{=} \mathcal{I}^{-1} T^{-1/2} g(\theta^\star) \\ T^{1/2}(\hat{\theta}_1^* - \hat{\theta}_1) &\overset{a}{=} \mathcal{I}_{11}^{-1} T^{-1/2} g_1(\theta^\star), \end{aligned}$$

which can be used to eliminate the estimators in (53) when we take the limit, obtaining an expression that involves only asymptotic information matrix and asymptotic score vector, as

[40] The $k \times k$ matrix of second derivatives of the log-likelihood with respect to θ.

[41] Let the asymptotic Hessian matrix be defined as $\mathcal{H}(\theta) \equiv plim \frac{1}{T} H(\theta)$. The information matrix equality, which assumes correct specification of the model, implies that $\mathcal{I}(\theta) = -\mathcal{H}(\theta)$.

[42] In the same way, $g_1(\theta, y)$ is the subvector of first derivatives of $l(\theta, y)$ with respect to the elements of θ_1

follows:

$$
\begin{aligned}
\operatorname{plim} T^{1/2}(\hat{\theta}^* - \hat{\theta}_0^*) &= \operatorname{plim} T^{1/2}(\hat{\theta}^* - \hat{\theta}) - \operatorname{plim} T^{1/2}(\hat{\theta}_0^* - \hat{\theta}) \\
&= \mathcal{I}^{-1} s - \mathcal{I}_{11}^{-1} s_1 \\
&= \mathbf{J} s, \qquad (54)
\end{aligned}
$$

where s_1 is the subvector of s that corresponds to θ_1, and

$$
\mathbf{J} \equiv \mathcal{I}^{-1} - \begin{bmatrix} \mathcal{I}_{11}^{-1} & 0_{(k-r)\times r} \\ 0_{r\times(k-r)} & 0_{r\times r} \end{bmatrix}. \qquad (55)
$$

Using (54), the probability limit of LR_T^* for $T \to \infty$ becomes:

$$
\operatorname{plim} LR_T^* = s' \mathbf{J} \mathcal{I} \mathbf{J} s. \qquad (56)
$$

Moreover, from (55), we have that

$$
\mathcal{I}\mathbf{J} = I_k - \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{bmatrix} \begin{bmatrix} \mathcal{I}_{11}^{-1} & 0_{(k-r)\times r} \\ 0_{r\times(k-r)} & 0_{r\times r} \end{bmatrix} = \begin{bmatrix} 0_{(k-r)\times(k-r)} & 0_{(k-r)\times r} \\ -\mathcal{I}_{21}\mathcal{I}_{11}^{-1} & I_r \end{bmatrix} \equiv \mathbf{Q}, \qquad (57)
$$

which implies $\mathcal{I}^{-1}\mathbf{Q} = \mathbf{J}$, $\mathbf{J}\mathbf{Q} = \mathbf{J}$ and $\mathbf{J}\mathcal{I}\mathbf{J} = \mathbf{J}$, from which we conclude that (56) can be written as

$$
\operatorname{plim} LR_T^* = s' \mathbf{J} s. \qquad (58)
$$

Now, notice that s is asymptotically $N(0, \mathcal{I})$, thus $s = \mathcal{I}^{1/2}\tilde{s}$, where $\tilde{s}$ is asymptotically standard normal. Therefore, (58) can be written as

$$
\operatorname{plim} LR_T^* = \tilde{s}' \mathcal{I}^{1/2} \mathbf{J} \mathcal{I}^{1/2} \tilde{s}, \qquad (59)
$$

which is χ^2 distributed with degrees of freedom equal to the rank of matrix $\mathcal{I}^{1/2}\mathbf{J}\mathcal{I}^{1/2}$:

$$
r(\mathcal{I}^{1/2}\mathbf{J}\mathcal{I}^{1/2}) = r(\mathcal{I}^{-1/2}\mathbf{Q}\mathcal{I}^{1/2}) = r,
$$

using the fact that $\mathcal{I}$ has full rank and that the rank of $\mathbf{Q}$ is r since its first $k - r$ rows are zero. Therefore, we can conclude that $LR_T^* \xrightarrow{d} \chi_r^2$, as we wanted to show.

D Consistency of Alternative Bootstrap Approaches

A key feature of our nonparametric bootstrap approach is that it is based on a resampling of the standardized residuals $\{\hat{e}_t := S_t^{-1/2}(\hat{\theta})\eta_t(\hat{\theta})\}_{t=1}^T$ in the observable dynamics of the innovation form representation (17)-(18) implied by the present-value model. We show that this feature is essential to obtain formal bootstrap validity. An alternative informal bootstrap approach, discussed in Section 4.6, relies on bootstrapping the filtered residuals ε_t^X of the unobservable state space dynamics. This appendix illustrates in detail the analytic derivations used to show that this alternative approach creates an asymptotic bias that makes the resulting asymptotic bootstrap inference invalid.

A necessary condition for bootstrap consistency is that the expected log-likelihood under the bootstrap distribution is maximized at the sample parameter estimate $\hat{\theta}$. The bootstrap applied directly to the estimated standardized innovations e_t enforces condition (31) by construction, since:

$$\begin{aligned} E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right] &= E^*\left[-\frac{1}{2}\sum_{t=1}^T\left(\log|S_t| + e_t^{*'}e_t^*\right)\right] \\ &= -\frac{1}{2}\sum_{t=1}^T\left(\log|S_t| + \frac{1}{T}\sum_{t=1}^T e_t' e_t\right), \end{aligned}$$

so that $\mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right) - E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right]$ is equal to

$$-\frac{1}{2}\sum_{t=1}^T\left(e_t'e_t - \frac{1}{T}\sum_{t=1}^T e_t'e_t\right) = 0$$

i.e., $\hat{\theta}$ maximizes $E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right]$.[43]

The alternative bootstrap approach instead is based on resampling the state innovations ε_t^X. From the Kalman filter in Appendix B we have:

$$\varepsilon_{t,t}^X = \Gamma' X_{t,t} = \Gamma'\left(F X_{t-1,t-1} + \mathcal{K}_t\eta_t\right) = \Gamma'\mathcal{K}_t\eta_t, \tag{60}$$

since $\Gamma' F = 0_{3\times 4}$, where $\mathcal{K}_t$ is the Kalman gain. Therefore,

$$\eta_t = \left(\Gamma'\mathcal{K}_t\right)_L^{-1}\varepsilon_{t,t}^X, \tag{61}$$

[43]See also Lemma 1 in Stoffer and Wall (1991).

with $(\Gamma'\mathcal{K}_t)_L^{-1}$ the left inverse of the 3×2 matrix $\Gamma'\mathcal{K}_t$.

The log-likelihood in the model can therefore be written as:

$$\begin{aligned}
\mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right) &= -\frac{1}{2}\sum_{t=1}^{T}\left(\log|S_t| + \eta_t' S_t^{-1}\eta_t\right) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(\log|S_t| + \varepsilon_{t,t}^{X'}\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\varepsilon_{t,t}^{X}\right) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left[\log|S_t| + Tr\left(\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\varepsilon_{t,t}^{X}\varepsilon_{t,t}^{X'}\right)\right] \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left[\log|S_t| + vec\left(\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\right)' vec\left(\varepsilon_{t,t}^{X}\varepsilon_{t,t}^{X'}\right)\right].
\end{aligned} \tag{62}$$

Accordingly, the expected log likelihood under the second bootstrap distribution is:

$$\begin{aligned}
E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right] &= -\frac{1}{2}\sum_{t=1}^{T}\left[\log|S_t| + vec\left(\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\right)' E^*\left[vec\left(\varepsilon_{t,t}^{X*}\varepsilon_{t,t}^{X*'}\right)\right]\right] \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left[\log|S_t| + vec\left(\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\right)' \frac{1}{T}\sum_{t=1}^{T} vec\left(\varepsilon_{t,t}^{X}\varepsilon_{t,t}^{X'}\right)\right],
\end{aligned} \tag{63}$$

which is different from $\mathcal{L}\left(\theta, \{Y_t\}_{t=1}^T\right)$.

The difference between (62) and (63) under the second bootstrap approach is explicitly given as

$$-\frac{1}{2}\sum_{t=1}^{T} vec\left(\left(\Gamma'\mathcal{K}_t\right)_L^{-1'} S_t^{-1}\left(\Gamma'\mathcal{K}_t\right)_L^{-1}\right)'\left[vec\left(\varepsilon_{t,t}^{X}\varepsilon_{t,t}^{X'}\right) - \frac{1}{T}\sum_{t=1}^{T} vec\left(\varepsilon_{t,t}^{X}\varepsilon_{t,t}^{X'}\right)\right] \neq 0. \tag{64}$$

It is not zero because differences from the mean in the square parenthesis are not just summed over the T observations, but first weighted by a term that varies with t and depends on the Kalman gain and the filtered covariance of the prediction errors. Therefore, under the alternative bootstrap approach, $E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right]$ is not necessarily maximized in $\hat{\theta}$.

E Broader Specifications of the Predictive Information Set

While the benchmark present-value model in Section 2 is useful for highlighting the main issues of tests of predictability hypotheses, it might not provide the most accurate description for the dynamics of dividend-return expectations and their link to price-dividend ratios. Richer specifications might improve the evidence of predictability and it is useful to study the robustness of our previous results, with respect to an enlarged specification of the predictive information set.

Several potential predictors have been considered in the literature, to improve the statistical evidence of univariate predictive regressions with the lagged price-dividend ratio.[44] Such predictive variables can naturally extend the benchmark present-value model, in order to parsimoniously aggregate the joint information generated by the time series of dividend growth, price-dividend ratios and additional predictors, following the present-value approach proposed in Yun (2012).

Using the conventional asymptotic approach, variables such as the book-to-market ratio (bm), the stock market variance ($svar$), the consumption-wealth-income ratio (cay) and the BAA-rated corporate bond yield (BAA) significantly improve the forecasts of future returns and future dividend growth in the present-value model.[45] Using our general bootstrap tests of Section 4, we study the robustness of our findings on dividend and return predictability, with respect to the choice of the predictive information set.

E.1 The present-value model with additional predictive variables: estimation results

Expected dividend growth, expected return and an additional predictive variable, z_t, follow the following first-order vector autoregression:

$$g_{t+1} = \gamma_0 + \gamma_1(g_t - \gamma_0) + \gamma_2(z_t - \xi_0) + \varepsilon^g_{t+1}, \tag{65}$$

$$\mu_{t+1} = \delta_0 + \delta_1(\mu_t - \delta_0) + \delta_2(z_t - \xi_0) + \varepsilon^\mu_{t+1}, \tag{66}$$

$$z_{t+1} = \xi_0 + \xi_1(z_t - \xi_0) + \varepsilon^z_{t+1}. \tag{67}$$

In contrast to the benchmark dynamics (3)-(4), the additional predictive variable z_t can help to better explain expected returns or expected dividend growth. As such, it appears in the

[44] Goyal and Welch (2008) and Koijen and Van Nieuwerburgh (2011) give an excellent review of this literature. Even though less studies have focused on dividend growth predictability, Lettau and Ludvigson (2005) and Favero, Gozluklu, and Tamoni (2011), among others, provide evidence that predictive variables like cay and proxies of demographics help forecasting cash flow growth.

[45] The benchmark present-value model assumes a constant return volatility. Piatti and Trojani (2017) develop a present-value approach with time-varying return and dividend growth risks to predictive regression.

price-dividend ratio implied by a standard Campbell and Shiller (1988) log linearization:

$$pd_t = A - (B_1\hat{\mu}_t + B_3\hat{z}_t) + (B_2\hat{g}_t + B_4\hat{z}_t), \tag{68}$$

where $B_3 = \frac{\delta_2}{\delta_1-\xi_1}\left(\frac{1}{1-\rho\delta_1} - \frac{1}{1-\rho\xi_1}\right)$, $B_4 = \frac{\gamma_2}{\gamma_1-\xi_1}\left(\frac{1}{1-\rho\gamma_1} - \frac{1}{1-\rho\xi_1}\right)$ and $\hat{z}_t = z_t - \xi_0$ is the demeaned additional predictive variable at time t; see, e.g., Yun (2012).

The model is again estimated in state space form with a Kalman filter.[46] For brevity, we report results only for additional predictive variables that significantly predict returns and dividend growth using standard asymptotic tests. These include the book-to-market ratio (bm), the stock market variance ($svar$), the consumption-wealth-income ratio (cay) and the corporate bond yield on BAA-rated bonds (BAA). The description of the variables is provided by Goyal and Welch (2008) and their updated time series through 2010 are available at Goyal's website.[47]

Estimated present-value model parameters and R-squares for returns and dividend growth are collected in Table 8, together with the R-squared estimated from standard predictive regressions with the additional predictive variable z_t. In each present-value model, the predictive information set enlarged by the additional predictor z_t increases the estimated R-squares for dividends and returns, relative to the findings for the benchmark model in Section 3. While estimated R-squares for returns are similar to those obtained from the standard predictive regressions in Panel C of Table 8, the estimated R-squared values for dividend growth are much higher, consistently with the findings of Section 3 for the benchmark present-value model.

E.2 Tests of constant dividend and return expectations

Cash flow predictability is again tested by testing the null hypothesis of constant expected dividend growth. In the extended present-value model, this hypothesis is equivalent to the following constraints, which are tested using a standard LR statistic that is asymptotically χ^2_5 distributed:

$$H_0 : \gamma_1 = \gamma_2 = \sigma_g = \rho_{g\mu} = \rho_{gz} = 0\ . \tag{69}$$

We test this null hypothesis for $z_t = bm$ and $z_t = svar$, which are the variables that seem to increase more model-implied dividend growth predictability, measured in terms of R-squared, compared to the benchmark model (see again Panel B of Table 8). Panel B of Table 9 shows that the asymptotic likelihood ratio test rejects null hypothesis (69) for both choices of predictive

[46] Supplemental Appendix F accurately describes the state space representation and the Kalman filter estimation procedure for the present-value model with the additional predictor z_t.

[47] See the web page http://www.hec.unil.ch/agoyal/.

variable z_t, with a p-value below 0.5%.

To apply our bootstrap testing approach, we introduce the extended vectors of observed variables $Y_t := (\Delta d_t, pd_t, z_t)'$ and state variables $X_t := (\hat{g}_{t-1}, \epsilon_t^g, \epsilon_t^\mu, \epsilon_t^d, \epsilon_t^z)'$, in order to write the present-value model (65)-(67) in state space form:

$$X_{t+1} = FX_t + Bu_{t+1} + \Gamma\epsilon_{t+1}^X \ , \tag{70}$$

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t \ , \tag{71}$$

with parameter-dependent matrices F, B, Γ, M_0, M_1, M_2 and variable $u_t := z_{t-1} - \xi_0$.

Given $X_{t,t-1}$ the best linear prediction of X_t based on data $\{Y_t\}_{s=1}^{t-1}$ and $\eta_t = Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t,t-1}$, the innovations form representation of model (65)-(67) follows from the Kalman filter:

$$X_{t+1,t} = FX_{t,t-1} + Bu_{t+1} + F\mathcal{K}_t\eta_t \ , \tag{72}$$

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_{t,t-1} + \eta_t \ . \tag{73}$$

From this dynamics, the bootstrap likelihood ratio test in the extended present-value model is performed with the algorithm presented in Section 4.3.[48]

Panel C of Table 9 shows that the bootstrap likelihood ratio test produces different conclusions from the asymptotic test. The bootstrap test p-values are always bigger than the asymptotic p-values and we can never reject null hypothesis (69) at the 5% significance level, indicating that the evidence of dividend growth predictability is similarly weak in the extended present-value models, as it was in Section 3 for the benchmark model.

The null hypothesis of no return predictability in the extended present-value model is equivalent to the following parametric constraints:

$$H_0 : \delta_1 = \delta_2 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = \rho_{\mu z} = 0 \ . \tag{74}$$

For brevity, we test again this null hypothesis using the two predictive variables that mostly increase the return predictability evidence, as measured by the model-implied R^2_{Ret}, namely

[48] To run the bootstrap algorithm in the extended present-value model, we replace Equation (17) in step 3) of the algorithm in Section 4 by the following bootstrap simulation scheme:

$$X^*_{t+1,t} = FX^*_{t,t-1} + Bu_{t+1} + F\mathcal{K}_t S_t^{1/2}\hat{e}^*_{0t} \ ,$$

using parameter matrices detailed in Supplemental Appendix F.

$z_t = bm$ and $z_t = cay$; see again Panel B of Table 8. Panel B of Table 10 shows that the asymptotic likelihood ratio test rejects null hypothesis (74) for all choices of the predictive variable z_t, with a p-value below 0.05%. The p-values for the bootstrap test are reported in Panel C of Table 10. Consistently with the asymptotic test results and the bootstrap test results in Section 3 for the benchmark model, null hypothesis (74) is again clearly rejected, with p-values of about 0.5%.

E.3 Variability of estimated R-squared values

To explain the weak evidence of dividend growth predictability and the large estimated dividend R-squares in the extended present-value models, Figure VIII of the Supplemental Appendix plots the bootstrap distribution of estimated R-squares for returns and future dividend growth, simulated under null hypothesis (69), for two different choices $z_t = bm$ (left panels) and $z_t = svar$ (right panels) of the additional predictive variable.

The bootstrap distribution of estimated R-squares under the null of constant expected dividend growth is similar to the one estimated in the benchmark model, with a large variability of estimated R-squares. The increased predictive information generated by z_t tends to rise the probability of correctly estimating an R-squared of 0% for dividend growth under the given null hypothesis for $z_t = bm$, while the distribution of R^2_{Div} displays more variability for $z_t = svar$. To illustrate, while the median estimated R^2_{Div}-value is 6.23% for $z_t = bm$ (9.41% for $z_t = svar$), the most frequently estimated R-squared value is 0%, but the probability of estimating a dividend R-squared of at least 22.32% (25.71%), as in the data, is still as large as 8.20% (10.60%). In summary, finite-sample variability again produces large estimated R-squares by chance alone, within a present-value model where dividend predictability is absent.

F Tables and Figures

Table 1: Results of the estimation of the present-value model in Section 2. The model is estimated by maximum likelihood, using yearly data from 1946 to 2010 on log dividend growth rates and log price-dividend ratio. Panel A presents estimates of the coefficients of the underlying processes. Panel B reports resulting coefficients of the present-value decomposition $pd_t = A - B_1\hat{\mu}_t + B_2\hat{g}_t$. Bootstrapped standard errors are in parentheses.

Panel A: Maximum likelihood estimates				
γ_0	δ_0	γ_1	δ_1	
0.057	0.083	0.304	0.927	
(0.009)	(0.010)	(0.337)	(0.089)	
σ_g	σ_μ	σ_D	$\rho_{g,\mu}$	$\rho_{\mu,D}$
0.065	0.015	0.002	0.231	-0.972
(0.023)	(0.024)	(0.028)	(0.419)	(0.606)
Panel B: Implied present-value parameters				
ρ	A	B_1	B_2	
0.974	3.637	10.332	1.421	
(0.004)	(0.140)	(2.418)	(6.396)	

Table 2: Constrained ML estimates of the present-value model and LR statistics for the tests of no return predictability ($H_0 : \delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$) and no dividend growth predictability ($H_0 : \gamma_1 = \sigma_g = \rho_{g\mu} = 0$). The first column reports the results of the unconstrained estimation, from Table 1. The second panel reports the R-squares for returns and dividend growth implied by the estimated present-value model. $LogL$ denotes the pseudo log-likelihood obtained, LR is the value of the Likelihood Ratio statistic computed using (10), $As - pval$ and $Boot - pval$ denote percentage p-values of the asymptotic and bootstrap LR tests, respectively.

	Unconstrained	No Ret Pred	No Div Pred
γ_0	0.057	0.072	0.055
δ_0	0.083	0.079	0.082
γ_1	0.304	0.996	0
δ_1	0.927	0	0.903
σ_g	0.065	0.002	0
σ_μ	0.015	0	0.021
σ_d	0.002	0.069	0.068
$\rho_{g\mu}$	0.231	0	0
$\rho_{\mu d}$	-0.972	0	0.357
R^2_{ret}	8.82%	0.00%	10.46%
R^2_{div}	17.58%	-0.12%	0.00%
$LogL$	230.84	215.91	224.33
LR		29.87	13.02
$As - pval$		0.00%	0.46%
$Boot - pval$		0.50%	9.50%

Table 3: Finite-sample sizes of the asymptotic and bootstrap tests, for the null hypotheses of no return predictaility ($H_0 : \delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$) and no dividend growth predictability ($H_0 : \gamma_1 = \sigma_g = \rho_{g\mu} = 0$). The finite-sample sizes of the asymptotic and bootstrap test, denoted α_T^* and α_T, respectively, are computed for a nominal size α of 5%.

	No Ret Pred	No Div Pred
Asymptotic test	60.5%	25.8%
Bootstrap test	5%	8%

Table 4: Unconstrained and constrained ML estimates of the present-value model using long sample (1927-2010) and assuming a regime shift in 1946 for the persistence, volatility and correlation parameters. The table also shows LR statistics for the tests of constant expected returns (prewar or postwar) and constant expected dividend growth (prewar and postwar). $LogL$ denotes the pseudo log-likelihood obtained and LR is the value of the Likelihood Ratio statistic. $As-pval$ and $Boot-pval$ denote percentage p-values of the asymptotic and bootstrap LR tests, respectively.

	Unconstr	No Ret Pred - pre	No Ret Pred - post	No Div Pred - pre	No Div Pred - post
γ_0	0.056	0.053	0.050	0.063	0.057
δ_0	0.087	0.094	0.089	0.112	0.100
γ_{1p}	-0.163	0.517	0.084	0	-0.135
γ_{1d}	0.382	0.442	0.999	0.298	0
δ_{1p}	0.740	0	0.739	0.546	0.483
δ_{1d}	0.951	0.960	0	0.974	0.954
σ_{gp}	0.222	0.148	0.223	0	0.233
σ_{gd}	0.064	0.064	0.007	0.058	0
$\sigma_{\mu p}$	0.051	0	0.042	0.140	0.079
$\sigma_{\mu d}$	0.012	0.012	0	0.012	0.015
σ_{dp}	0.074	0.125	0.025	0.236	0.066
σ_{dd}	0.008	0.009	0.072	0.032	0.068
$\rho_{g\mu p}$	-0.719	0	-0.208	0	-0.658
$\rho_{g\mu d}$	0.273	0.247	0	0.170	0
$\rho_{\mu dp}$	0.695	0	0.978	0.242	0.753
$\rho_{\mu dd}$	-0.962	-0.969	0	0.285	0.328
$R^2_{ret,p}$	3.81%	0.00%	1.19%	-21.76%	-1.34%
$R^2_{ret,d}$	7.35%	7.06%	0.00%	7.12%	8.37%
$R^2_{div,p}$	77.93%	71.44%	85.50%	0.00%	90.08%
$R^2_{div,d}$	23.44%	23.00%	-11.04%	15.59%	0.00%
$LogL$	284.45	280.49	262.32	261.75	275.83
LR		7.92	44.26	45.40	17.24
$As-pval$		9.45%	0.00%	0.00%	0.06%
$Boot-pval$		11%	0.00%	0.00%	14%

Table 5: Unconstrained and constrained ML estimates of the present-value model using total payout (dividend plus repurchases) as an alternative measure of dividends. The table also shows LR statistics for the tests of constant expected returns ($H_0 : \delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$) and constant expected dividend growth ($H_0 : \gamma_1 = \sigma_g = \rho_{g\mu} = 0$). $LogL$ denotes the pseudo log-likelihood obtained and LR is the value of the Likelihood Ratio statistic. $As - pval$ and $Boot - pval$ denote percentage p-values of the asymptotic and bootstrap LR tests, respectively.

	Unconstrained	No Ret Pred	No Div Pred
γ_0	0.067	0.086	0.065
δ_0	0.104	0.098	0.102
γ_1	0.316	0.994	0
δ_1	0.843	0	0.825
σ_g	0.076	0.003	0
σ_μ	0.031	0	0.033
σ_d	0.001	0.080	0.080
$\rho_{g\mu}$	0.478	0	0
$\rho_{\mu d}$	0.878	0	0.403
R^2_{ret}	16.01%	0.00%	6.56%
R^2_{div}	12.79%	0.16%	0.00%
$LogL$	200.36	187.08	196.39
LR		26.55	7.91
$As - pval$ (%)		0.00%	4.78%
$Boot - pval$ (%)		1.00%	16.00%

Table 6: Unconstrained and constrained ML estimates of the present-value model using cash M&A (cash dividends plus M&A cash flows) as an alternative measure of dividends. The table also shows LR statistics for the tests of constant expected returns ($H_0 : \delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$) and constant expected dividend growth ($H_0 : \gamma_1 = \sigma_g = \rho_{g\mu} = 0$). $LogL$ denotes the pseudo log-likelihood obtained and LR is the value of the Likelihood Ratio statistic. $As - pval$ and $Boot - pval$ denote percentage p-values of the asymptotic and bootstrap LR tests, respectively.

	Unconstrained	No Ret Pred	No Div Pred
γ_0	0.080	0.086	0.079
δ_0	0.115	0.122	0.114
γ_1	0.168	0.881	0
δ_1	0.896	0	0.904
σ_g	0.161	0.034	0
σ_μ	0.027	0	0.029
σ_d	0.068	0.173	0.177
$\rho_{g\mu}$	0.618	0	0
$\rho_{\mu d}$	0.216	0	0.682
R^2_{ret}	5.34%	0.00%	6.41%
R^2_{div}	15.21%	4.41%	0.00%
$LogL$	166.84	161.56	142.56
LR		48.55	10.57
$As - pval$ (%)		0.00%	1.43%
$Boot - pval$ (%)		10.50%	7.00%

Table 7: This table summarizes the return and dividend growth predictability evidence using different cash-flow proxies, i.e. standard cash dividends, total payout and dividends corrected for M&A related cash flows. Data are annual from 1946 to 2010 (to 2003 for total payout). The first panel shows the estimated slope coefficient of standard predictive regressions, as long as their Newey-West corrected t-statistics and p-values, and the adjusted R-squared of the regressions. The second panel shows the R-squares implied by the estimated present-value model and the p-values of asymptotic and bootstrap tests of predictability.

Panel A: Predictive Regressions Results								
	Dividends				Returns			
	β	tstat	p-value	R^2	β	tstat	p-value	R^2
Cash Dividends	-0.016	-0.644	52.22%	-0.65%	-0.122	-2.961	0.43%	8.45%
Total Payout	-0.033	-0.947	34.78%	-0.29%	-0.212	-3.840	0.03%	14.81%
Cash M&A	0.103	1.857	6.81%	3.53%	-0.110	-2.589	1.20%	4.83%

Panel B: Present-value Model R-squares and Test Results						
	R-squares		Test p-values			
			Asymptotic Test		Bootstrap Test	
	R^2_{Div}	R^2_{Ret}	*Div*	*Ret*	*Div*	*Ret*
Cash Dividends	17.58%	8.82%	0.46%	0.00%	9.50%	0.50%
Total Payout	12.79%	16.01%	4.78%	0.00%	16.00%	1.00%
Cash M&A	15.21%	5.34%	1.43%	0.00%	7.00%	10.50%

Table 8: Panel A reports estimation results of the extended present-value model, using as predictor variables the book-to-market ratio (BM), stock variance ($SVAR$), CAY and the BAA corporate bond yield (BAA), respectively. The models are estimated using annual data from 1946 to 2010. Panel B reports the model-implied R-squared values for return and dividend growth, in percentage, computed as in (8)-(9), while Panel C reports R-squared from standard OLS predictive regressions of returns and dividend growth on lagged price-dividend ratio and each predictive variable z_t:

$$\begin{aligned} r_{t+1} &= a_r + b_r pd_t + \gamma_r z_t + \epsilon^r_{t+1} \\ \Delta d_{t+1} &= a_d + b_d pd_t + \gamma_d z_t + \epsilon^d_{t+1}. \end{aligned}$$

	BM	$SVAR$	CAY	BAA
Panel A: Maximum-likelihood estimates				
γ_0	0.051	0.056	0.050	0.057
δ_0	0.071	0.129	0.077	0.090
γ_1	0.234	0.475	0.338	0.296
δ_1	0.878	0.993	0.926	0.920
σ_g	0.064	0.078	0.066	0.065
σ_μ	0.016	0.018	0.031	0.018
σ_d	0.013	0.018	0.015	0.008
$\rho_{g\mu}$	0.220	-0.454	-0.308	0.177
$\rho_{\mu d}$	-0.144	-0.758	-0.596	-0.167
ξ_0	0.487	0.020	0	0.082
ξ_1	0.913	0.418	0.733	0.939
ρ_{gz}	-0.298	-0.764	-0.528	-0.250
$\rho_{\mu z}$	0.567	0.804	0.882	0.485
ρ_{dz}	0.597	-0.594	-0.775	0.702
σ_z	0.102	0.021	0.014	0.010
δ_2	0.019	-0.285	-0.396	-0.025
γ_2	0.067	1.652	0.837	-0.014
Panel B: Model-implied R-squared				
R^2_{ret}	10.13%	9.70%	17.58%	9.65%
R^2_{div}	22.32%	25.71%	19.81%	18.29%
Panel C: Predictive regression R-squared				
R^2_{ret}	10.34%	12.24%	15.40%	13.81%
R^2_{div}	4.26%	8.02%	0.95%	3.03%

Table 9: Test of no dividend growth predictability in the context of the extended present-value model:

$$H_0 : \gamma_1 = \gamma_2 = \sigma_g = \rho_{g\mu} = \rho_{gz} = 0.$$

Panel A reports constrained estimation results, using as predictor variables the book-to-market ratio (BM) and stock variance ($SVAR$), respectively. The models are estimated using annual data from 1946 to 2010. Panel B reports the p-values of the test, using the asymptotic distribution of the LR statistic, and the effective size of the asymptotic test, for a nominal size $\alpha = 5\%$, while Panel C reports the p-values of the bootstrap test. Finite sample size computations and bootstrap tests are based on 1000 bootstrap samples.

	BM	$SVAR$
Panel A: Constrained Maximum-likelihood estimates		
γ_0	0.056	0.055
δ_0	0.078	0.094
γ_1	0	0
δ_1	0.887	0.960
σ_g	0	0
σ_μ	0.019	0.031
σ_d	0.068	0.068
$\rho_{g\mu}$	0	0
$\rho_{\mu d}$	0.382	-0.029
ξ_0	0.504	0.021
ξ_1	0.904	0.363
ρ_{gz}	0	0
$\rho_{\mu z}$	0.681	0.927
ρ_{dz}	0.136	-0.216
σ_z	0.101	0.022
δ_2	0.006	-0.701
γ_2	0	0
Panel B: Asymptotic test		
$p-value$ (%)	0.22%	0.19%
$empirical\ size$ (%)	22.70%	28.60%
Panel C: Bootstrap test		
$p-value$ (%)	7.40%	10.20%

Table 10: Test of no return predictability in the context of the extended present-value model:

$$H_0 : \delta_1 = \delta_2 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = \rho_{\mu z} = 0.$$

Panel A reports constrained estimation results, using as predictor variables the book-to-market ratio (BM) and CAY, respectively. The models are estimated using annual data from 1946 to 2010. Panel B reports the p-values of the test, using the asymptotic distribution of the LR statistic, and the effective size of the asymptotic test, for a nominal size $\alpha = 5\%$, while Panel C reports the p-values of the bootstrap test. Finite sample size computations and bootstrap tests are based on 1000 bootstrap samples.

	BM	CAY
Panel A: Constrained Maximum-likelihood estimates		
γ_0	0.070	0.072
δ_0	0.075	0.080
γ_1	0.960	0.996
δ_1	0	0
σ_g	0.008	0.003
σ_μ	0	0
σ_d	0.067	0.069
ξ_0	0.234	-0.006
ξ_1	0.997	0.857
ρ_{gz}	0.811	-0.796
$\rho_{\mu z}$	0	0
ρ_{dz}	0.215	-0.253
σ_z	0.104	0.014
δ_2	0	0
γ_2	-0.001	0.015
Panel B: Asymptotic test		
$p-value$ (%)	0.01%	0.00%
$empirical\ size$ (%)	14.10%	28.10%
Panel C: Bootstrap test		
$p-value$ (%)	0.50%	0.30%

Figure 1: The first (second) panel displays the quantiles of the empirical distribution of the LR statistics for the tests of constant expected returns (dividend growth), obtained through a nonparametric bootstrap simulation procedure, against the quantiles of the asymptotic chi-squared distribution of the statistics (dotted red line). The vertical dotted line denotes the 95% quantile of this distribution.

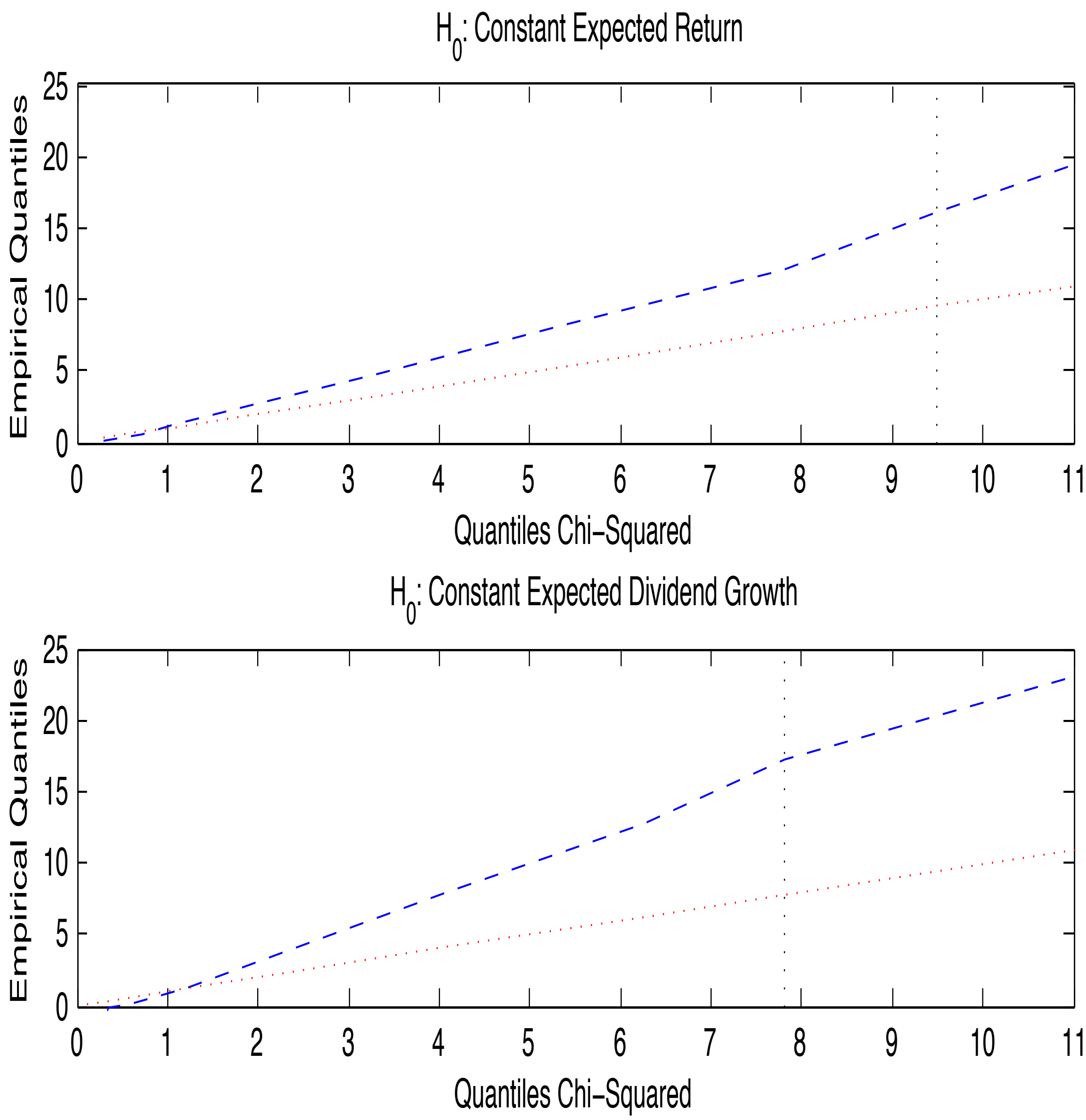

Figure 2: Test for normality of the filtered innovations for the unconstrained model and under different null hypotheses (No return predictability, no dividend growth predictability, equal autoregressive coefficients and $\delta_1 - \gamma_1 = 0.05, 0.15, \ldots, 0.95$). The two axes show the p-value of the Jarque-Bera test applied to the filtered dividend and price-dividend shocks, respectively.

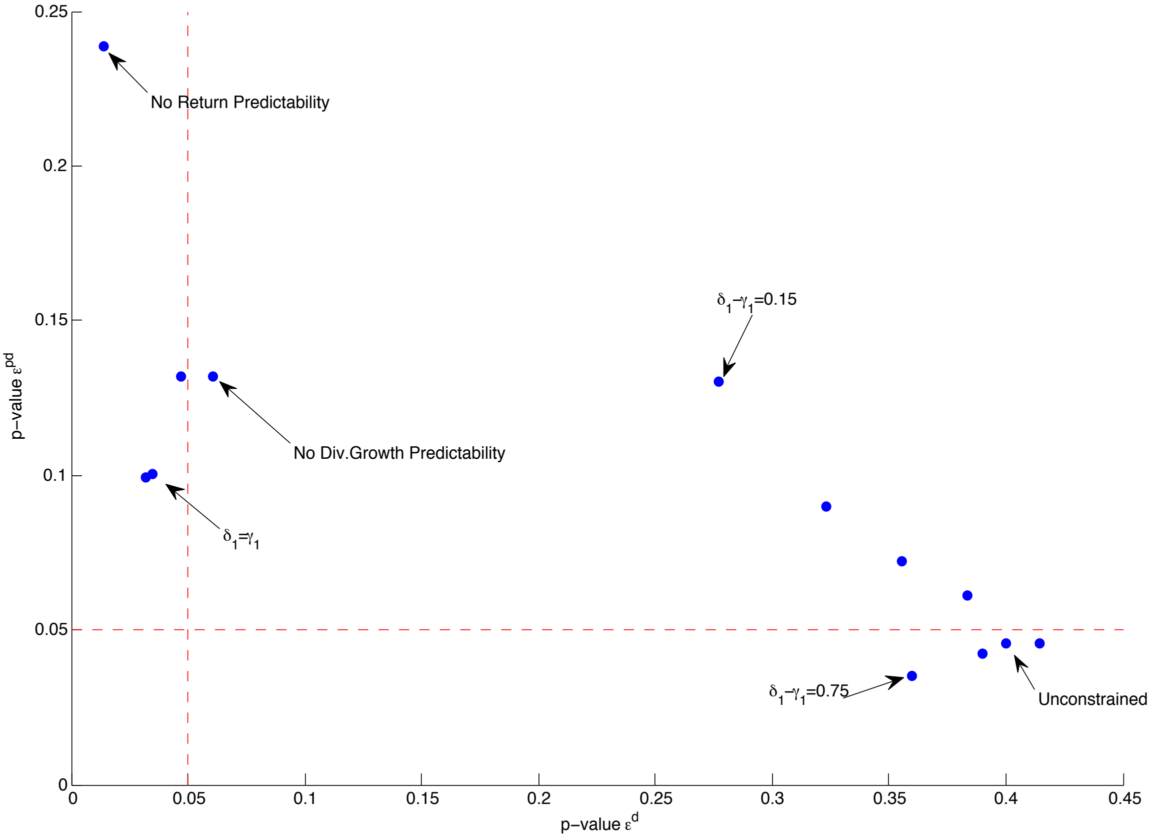

Figure 3: Average log-likelihood, i.e. $E^*\left[\mathcal{L}\left(\theta, \{Y_t^*\}_{t=1}^T\right)\right]$, over $B = 500$ bootstrap samples $\{Y_t^*\}_{t=1}^T$ generated from the innovation approach (*Innovation Form*, blue line) and the alternative approach (*State Dynamics*, red line), as a function of one specific parameter, keeping all other parameters fixed at their value in $\hat{\theta}$. The vertical line in each panel denotes the value in $\hat{\theta}$ of the parameter that is allowed to vary. The parameters considered are γ_1 (top left), δ_1 (top right), σ_g (middle left), σ_μ (middle right), and σ_D (bottom).

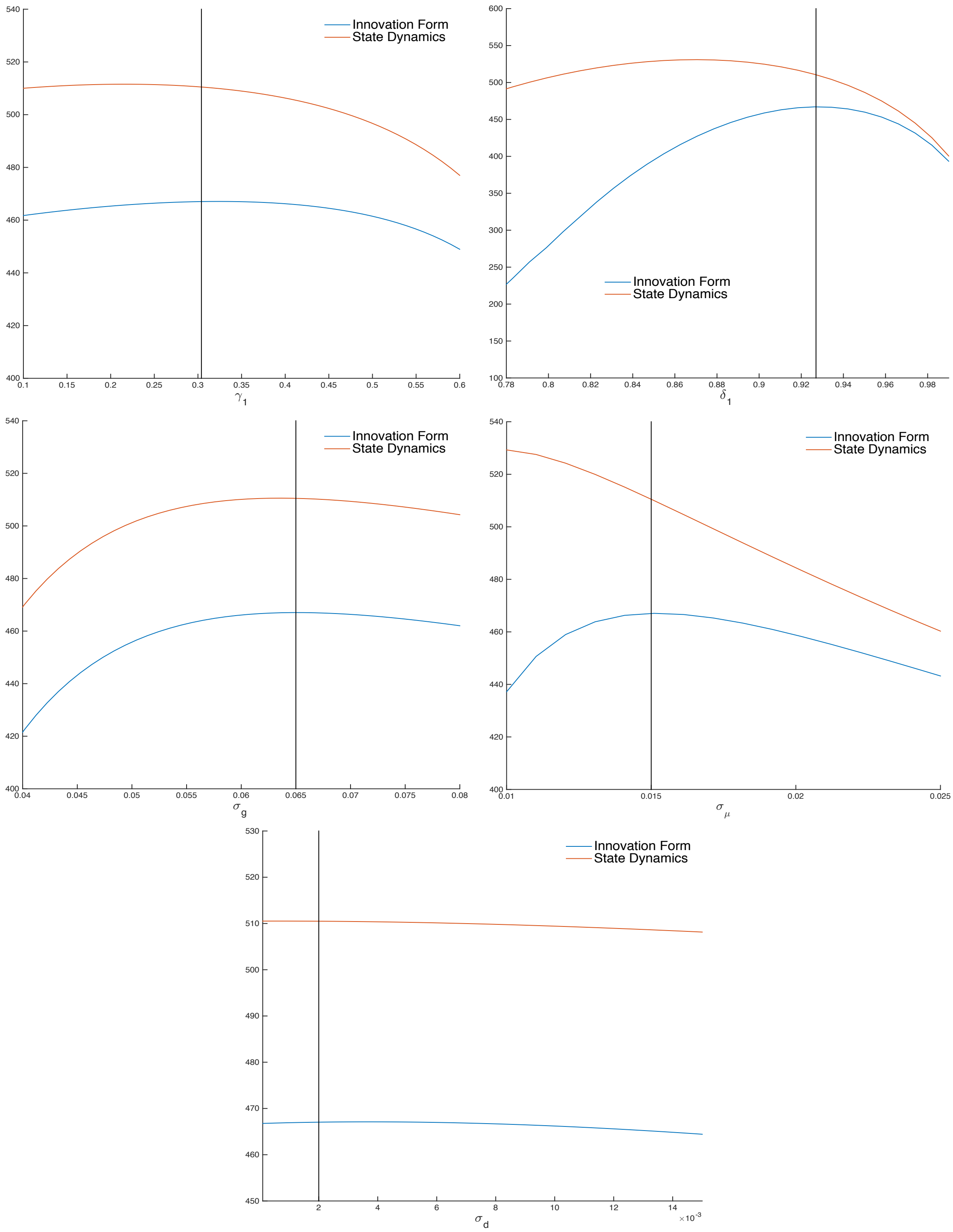

Figure 4: Distribution of selected statistics of the simulated observables over $B = 1000$ bootstrap samples $\{Y_t^*\}_{t=1}^T$ generated from the innovation approach (*Innovation Form*, blue bars) and the alternative approach (*State Dynamics*, yellow bars). The vertical red line in each panel denotes the value in the data. The statistics considered are the standard deviation of returns (top left), the autocorrelation of returns (top right) and the correlation between returns and dividend growth (bottom).

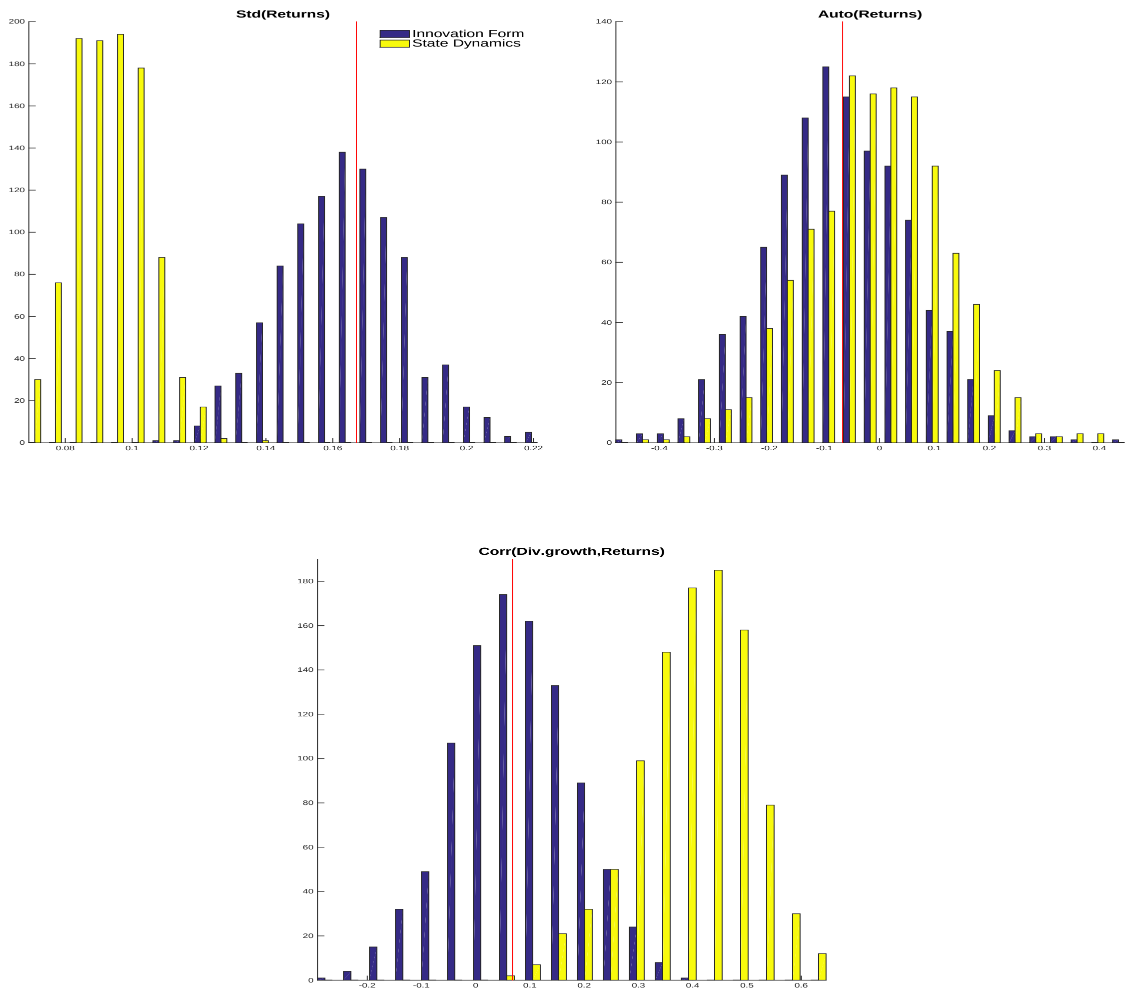

Figure 5: Bootstrapped distribution of the R-squared of returns (upper panel) and dividend growth (lower panel), starting from the estimates under the constraint of constant expected dividend growth ($\gamma_1 = \sigma_g = \rho_{g\mu} = 0$, left panels). Vertical dotted red lines and dashed black lines denote R-squared from constrained and unconstrained estimations on real data, respectively. Distributions are based on 1000 bootstrap samples.

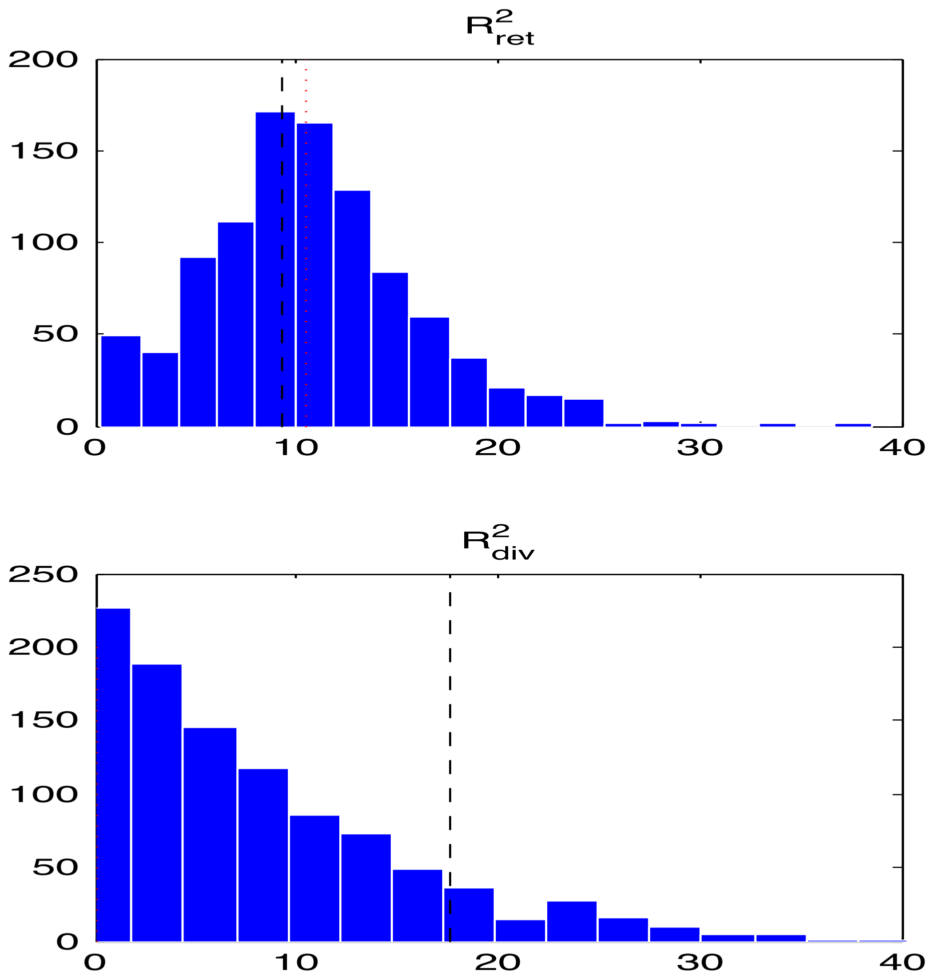

Figure 6: Bootstrapped distribution of the R-squared of returns (upper panel) and dividend growth (lower panel), starting from the estimates under the constraint of constant expected return ($\delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$). Vertical dotted red lines and dashed black lines denote R-squared from constrained and unconstrained estimations on real data, respectively. Distributions are based on 1000 bootstrap samples.

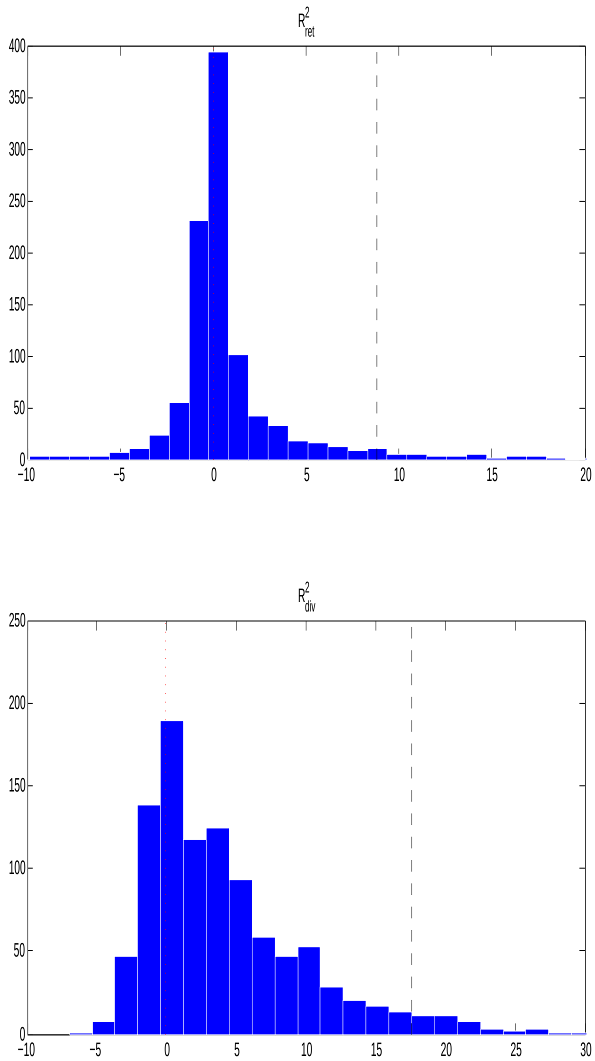

Figure 7: Bootstrapped distribution of the out-of-sample R-squared of returns (upper panel) and dividend growth (lower panel), starting from the estimates under the constraint of constant expected dividend growth ($\gamma_1 = \sigma_g = \rho_{g\mu} = 0$). Vertical dotted red lines and dashed black lines denote out-of-sample R-squared from constrained and unconstrained estimations on real data, respectively. Distributions are based on 1000 bootstrap samples.

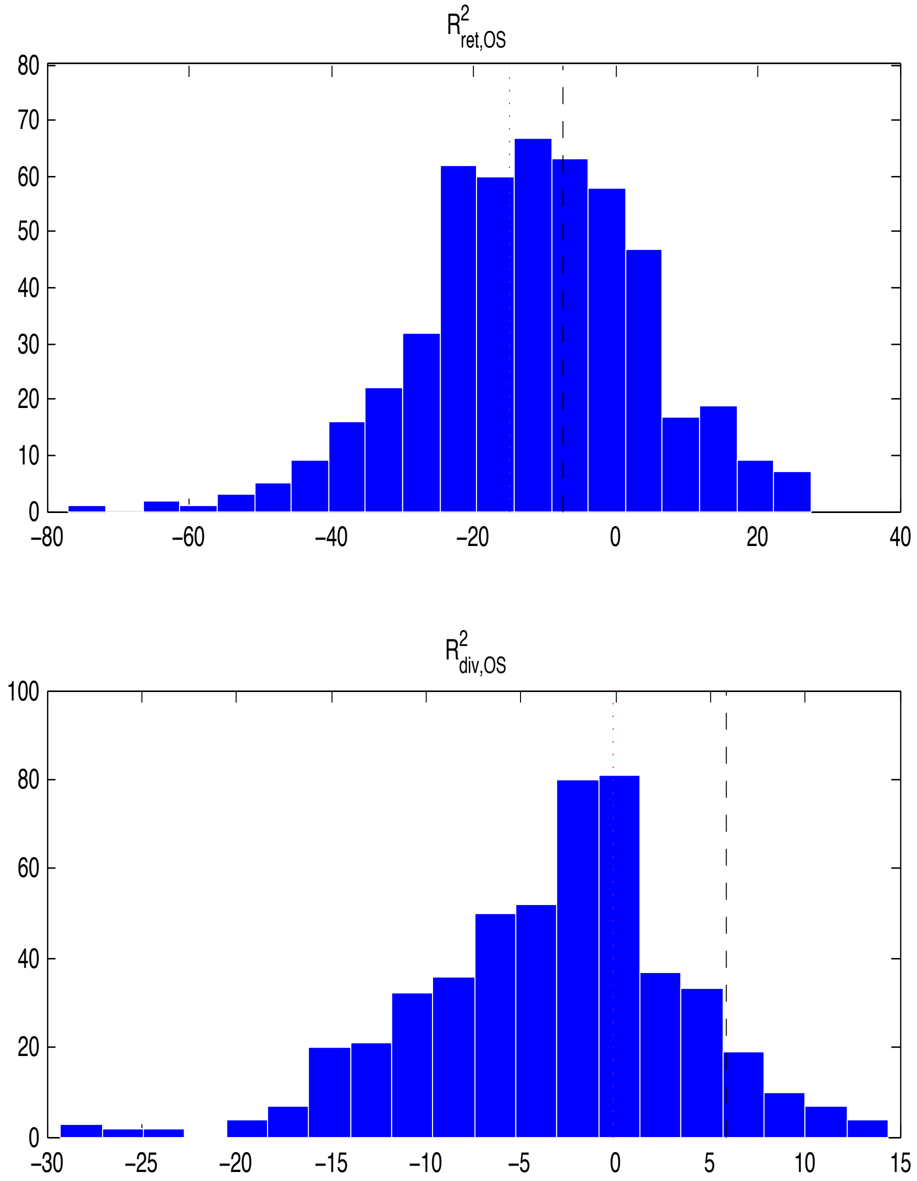

Figure 8: Bootstrapped distribution of the out-of-sample R-squared of returns (upper panel) and dividend growth (lower panel), starting from the estimates under the constraint of constant expected return ($\delta_1 = \sigma_\mu = \rho_{g\mu} = \rho_{\mu d} = 0$). Vertical red dotted lines and dashed black lines denote out-of-sample R-squared from constrained and unconstrained estimations on real data, respectively. Distributions are based on 1000 bootstrap samples.

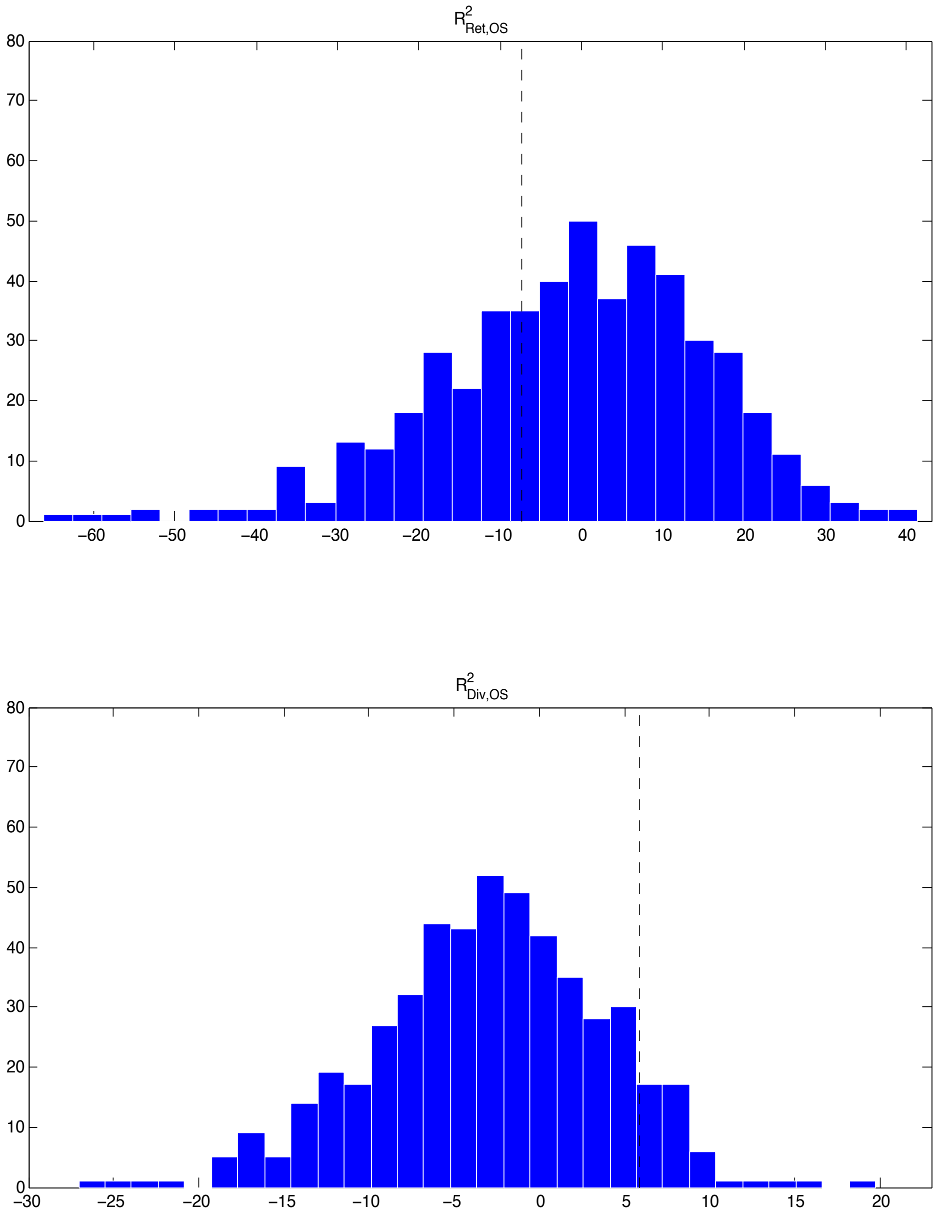

Figure 9: Yearly cash-flow growth (upper panel) and ratio of price over cash-flow (lower panel), using different measures of cash-flow: dividend (blue line), total payout (dividend plus repurchases, red line), net payout (dividend plus repurchases minus issuances, green line), and dividends including M&A cash flows (magenta line). Data are annual from 1946 to 2010 for dividends and cash M&A and to 2003 for payout and net payout.