Predictable Risks and Predictive Regression in Present-Value Models*

Ilaria Piatti† and Fabio Trojani‡

First version: December 2010; This version: July 2011

Abstract

In a present-value model with time-varying risks, we develop a latent variable approach to estimate expected market returns and dividend growth consistently with the conditional risk features implied by present-value constraints. We find a time-varying expected dividend growth and expected return, with the explained fractions of return and dividend growth variability which are around 10% at an annual frequency. Expected return is more persistent than expected dividend growth and generates large price-dividend ratio components that mask the predictive power for future dividend growth. The model implies (i) predictive regressions consistent with a weak return predictability and a missing dividend predictability by aggregate price-dividend ratios, (ii) predictable market volatilities, (iii) volatile and often counter-cyclical Sharpe ratios and (iv) a time-varying and hump-shaped term structure of stock market risk. These findings show the importance of controlling for time-varying risks and the potential long-run effect of persistent return or dividend forecasts when studying predictive relations.

*We would like to thank Jules van Binsbergen, Christian Gouriéroux, Paolo Porchia, and the participants of the II Workshop on Games and Decisions in Reliability and Risk, Belgirate, the Workshop in Time Series and Financial Econometrics at Bocconi University, the 10th Swiss Doctoral Workshop in Finance, the 8th FINRISK Research Day, and the seminar participants of IE Business School, Madrid, and Warwick Business School, for valuable comments. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK). The usual disclaimer applies.

†University of Lugano, Via Buffi 13, CH-6900 Lugano, Switzerland; e-mail: ilaria.piatti@usi.ch

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

1 Introduction

We propose a latent variable framework with time-varying risks, to estimate expected market returns and dividend growth rates consistently with the conditional risk restrictions of present-value models. This approach aggregates information from the history of dividend growth, price-dividend ratios and market volatilities, and uncovers expected returns and dividend growth rates coherently with the conditional risk features of dividends and returns. Given exogenous latent processes for expected market returns, expected aggregate dividend growth and the variance-covariance structure of dividends and returns, we specify a Campbell and Shiller (1988) present-value model that constraints the conditional risk structure of expected return and dividend shocks, together with the implied price-dividend ratio dynamics. We finally apply a Kalman filter to estimate the model by Quasi Maximum Likelihood (QML).

We find that expected dividend growth and expected returns are both time-varying, but while expected dividend growth explains a negligible fraction of actual dividend growth (with average model-implied R^2 values below 1%), expected returns explain a large portion of future returns (with average model-implied R^2 values of about 50%). Estimated expected dividend growth is more persistent than expected returns and gives rise to a large price-dividend ratio component that masks the predictive power of valuation ratios for future returns. These findings have important economic implications. First, they produce a sharp statistical evidence for return predictability. Second, they highlight the potential presence of expected dividend growth components that are substantially more persistent than expected returns. Third, they stress the importance of accounting for persistent dividend forecasts and their long-run effects when predicting future returns with the price-dividend ratio. Using our present-value model with time-varying risks, we also uncover the potential implications of these predictive structures for the long-horizon predictability of market returns and the term structure of market risks. First, we find that the lower persistence of expected market returns is linked to a weaker model-implied predictability at longer horizons. Second, we observe that the larger uncertainty about future expected returns produces a model-implied term structure of risks that is often upward sloping and sometimes hump-shaped.

Using Monte Carlo simulations, we show that, despite the large estimated degree of return predictability, our model is broadly consistent with (i) the weak statistical evidence of return predictability in predictive regressions with aggregate price-dividend ratios, (ii) the even weaker evidence of dividend growth predictability at yearly horizons, (iii) a low real-time predictability of stock returns, (iv) predictable market risks, (v) volatile and often counter-cyclical Sharpe ratios, (vi) a stronger evidence of return predictability using long-horizon predictive regressions and (vii) a decreasing term structure of market risks, uncovered by variance ratio tests or multi-period ahead iterated VAR forecasts. Finally, we find that while the predictive power of price-dividend ratios for future returns is low and time-varying, the forecasting power of price-dividend ratios adjusted for the hidden expected dividend growth component is large and more stable over time.

Our approach builds on the recent literature advocating the use of present-value models to jointly uncover market expectations for returns and dividends, including 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), Rytchkov (2008), Cochrane (2008a), Cochrane (2008b), Ferreira and Santa-Clara (2010) and van Binsbergen and Koijen (2010), among others. We add to this literature by introducing a tractable present-value model incorporating the latent time-varying features of return and dividend risks, in which we study the implications for the identification of potentially persistent dividend growth components, the detection of predictive relations and the estimation of time-varying risk features.

Using our modeling framework, we reconcile a number of predictive regression findings in the literature. First, we show that large time-varying expected return components are compatible with the weak in-sample predictability of market returns by aggregate price-dividend ratios, as well as with both predictable market risks and high Sharpe ratio volatilities. Second, we show that our findings are consistent with Goyal and Welch (2008) observation that aggregate price-dividend ratios have no additional out-of-sample predictive power for market returns, relative to a straightforward sample mean forecast. Third, our results indicate that a present-value model with time-varying risks is able to identify persistent dividend components in price-dividend ratios, which can be related to the long-run implications of expected dividend growth, studied in Bansal and Yaron

(2004), Lettau and Ludvigson (2005) and Menzly, Santos, and Veronesi (2004), among others. In contrast, the model with constant risks tends to identify a less persistent expected dividend growth process, which explains a large fraction of future dividend growth (with average model-implied R^2 values of about 99%). Fourth, the persistent dividend component in price-dividend ratios is responsible for the weak and time-varying predictability evidence of standard predictive regressions. We show that price-dividend ratios adjusted by this component produce a strong and more robust evidence in favour of return predictability, by eliminating a large fraction of the time-instabilities noted by Lettau and Van Niewerburgh (2008), among others, within standard predictive regression models.[1] Fifth, a framework featuring time-varying risks can potentially help to reconcile some of the implications for the term structure of market risks and the long-horizon predictability features. Our findings show that even if from an investor's perspective the average term structure of market risks can be increasing, as motivated, e.g., by Pastor and Stambaugh (2010),[2] the term structure of risks uncovered by multi-period ahead VAR forecasts can be decreasing, as shown, e.g., in Campbell and Viceira (2005). Similarly, even if the term structure of long-horizon predictability can be decreasing from an investor's perspective, the one uncovered by multi-period ahead VAR forecasts can be increasing, as emphasized, e.g., by Cochrane (2008a). Finally, we provide independent evidence on the importance of time-varying risk features to uncover predictive return relations within present-value models. Using a particle filter approach, Johannes, Korteweg, and Polson (2011) estimate a set of Bayesian predictive regressions of market returns on aggregate payout yields. They show that models with return predictability and time-varying risks can produce a large additional economic value, from the perspective of a Constant Relative Risk Aversion investor maximizing the predictive utility of her terminal wealth. In contrast, models with constant risks imply no substantial economic gain in incorporating predictability features. Consistently with these findings, our model estimates a large degree of return predictability, which is hardly uncovered by the setting

[1] This last finding supports the intuition, put forward in Lacerda and Santa-Clara (2010), among others, that price-dividend-ratios adjusted by a smooth real-time proxy of dividend expectations can have a large and more robust predictive power for future returns.

[2] In a Bayesian predictive regression setting with time-varying expected returns and volatility, Johannes, Korteweg, and Polson (2011) also find a sometimes increasing term structure of market risk.

with constant risks.

The paper proceeds as follows. Section 2 introduces our present-value model with time-varying return and dividend risks. In Section 3, we discuss our data set and the estimation strategy, while Section 4 presents estimation results. In Section **??**, we analyse the model implications and show that they are consistent with a number of predictive regression findings in the literature. Section 5 discusses additional implications of the model and Section 6 concludes.

2 Present-Value Model

As shown in Cochrane (2008a), among others, dividend growth and returns are better studied jointly in order to understand their predictability features. Following Campbell and Shiller (1988), this section introduces a present-value model with time-varying risks for the joint dynamics of aggregate dividends and market returns. We denote by

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

the cum-dividend log market return, and by

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

the aggregate log dividend growth. Expected return and dividend growth, conditional on investors' information set at time t, are denoted by $\mu_t \equiv E_t[r_{t+1}]$ and $g_t \equiv E_t[\Delta d_{t+1}]$, respectively, while the conditional variance-covariance or returns and dividend growth is denoted by Σ_t.

μ_t, g_t and Σ_t follow exogenous latent processes that model the time-varying second-order structure of returns and dividends:

$$\begin{pmatrix} \Delta d_{t+1} \\ r_{t+1} \end{pmatrix} = \begin{pmatrix} g_t \\ \mu_t \end{pmatrix} + \Sigma_t^{1/2} \begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix}, \tag{3}$$

where $(\varepsilon^D_{t+1}, \varepsilon^r_{t+1})'$ is a bivariate iid process. Expected returns and expected dividends follow simple linear autoregressive processes, allowing for the potential presence of a risk-in-mean effect linked to Σ_t:[3]

$$
\begin{aligned}
g_{t+1} &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \varepsilon^g_{t+1}, & (4)\\
\mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + Tr(\Lambda(\Sigma_t - \mu^{\Sigma})) + \varepsilon^{\mu}_{t+1}, & (5)
\end{aligned}
$$

with real valued parameters $\gamma_0, \gamma_1, \delta_0, \delta_1$ and symmetric 2×2 parameter matrices Λ and μ^{Σ}. $Tr(.)$ denotes the trace of a matrix, i.e., the sum of its diagonal components. Parameter μ^{Σ} is the unconditional mean of stationary variance-covariance process Σ_t, while parameter Λ captures the potential presence of a risk-in-mean effect linked to the time-varying risks of returns and dividends. Shocks $(\varepsilon^g_{t+1}, \varepsilon^{\mu}_{t+1})'$ have zero conditional means, but they feature a potentially time-varying risk structure, which has to be consistent with the present-value constraints imposed on the dynamics of dividends, returns and price-dividend ratios, discussed in detail below. The case $\Lambda = 0$ corresponds to a model with no risk-in-mean effect. In this case, the conditional mean of (g_{t+1}, μ_{t+1}) has a simple linear autoregressive structure. However, process (g_{t+1}, μ_{t+1}) does not follow a standard linear autoregressive process with constant risk, as for instance the one studied in van Binsbergen and Koijen (2010), because also in this case shocks $(\varepsilon^g_{t+1}, \varepsilon^{\mu}_{t+1})'$ feature a degree of heteroskedasticity, induced by present-value constraints when Σ_t is time-varying.

We specify the dynamics of Σ_t by a simple autoregressive process that implies tractable price-dividend ratio formulas also in presence of a risk-in-mean effect. Precisely, we assume that Σ_t follows a Wishart process of order one (see Gourieroux, Jasiak, and Sufana (2009) and Gourieroux (2006)):

$$
\Sigma_{t+1} = M\Sigma_t M' + kV + \nu_{t+1}, \tag{6}
$$

with integer degrees of freedom $k > 1$, a 2×2 matrix M of autoregressive parameters and a 2×2 symmetric and positive-definite volatility of volatility matrix V. Note that for

[3]A large literature studies the relation between conditional mean and conditional volatility of stock returns. See, e.g., Pastor, Sinha, and Swaminathan (2008), Campbell (1987), Breen, Glosten, and Jagannathan (1989), French, Schwert, and Stambaugh (1987), Schwert (1989), Whitelaw (1994), Ludvigson and Ng (2007), Ghysels, Santa-Clara, and Valkanov (2005), Bollersev, Engle, and Wooldridge (1988), Glosten, Jagannathan, and Runkle (1993), Brandt and Kang (2004), Gallant, Hansen, and Tauchen (1990) and Harrison and Zhang (1999). An excellent review of this literature is provided by Lettau and Ludvigson (2010).

$k > n - 1$ process Σ_t takes positive semi-definite values, making dynamics (6) a naturally suited model for multivariate time-varying risks. The conditional distribution of Σ_{t+1} is Wishart and completely characterized by the (affine) Laplace transform:

$$\Psi_t(\Gamma) = E_t\left[\exp Tr(\Gamma\Sigma_{t+1})\right] = \frac{\exp Tr\left[M'\Gamma(I_2 - 2V\Gamma)^{-1}M\Sigma_t\right]}{\left[\det(I_2 - 2V\Gamma)\right]^{k/2}}\ , \tag{7}$$

which implicitly defines the conditional distribution of zero mean 2×2 error term ν_{t+1} in model (6). Under process (6), the unconditional mean μ^{Σ} is the unique solution of the (implicit) steady state equation:

$$\mu^{\Sigma} = kV + M\mu^{\Sigma}M'. \tag{8}$$

Finally, it can be shown that the dynamic dependence structure between risk factors in this model is quite flexible, with, e.g., both conditional and unconditional correlations that are unrestricted in sign.

2.1 Price-dividend ratio

Let $pd_t \equiv \log \frac{P_t}{D_t}$ denote the log price-dividend ratio. To derive the expression for the price-dividend ratio implied by our model, we follow Campbell and Shiller (1988) log linearization approach:[4]

$$r_{t+1} \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t, \tag{9}$$

where $\overline{pd} = E[pd_t]$, $\kappa = \log(1+\exp(\overline{pd})) - \rho\overline{pd}$ and $\rho = \frac{\exp(\overline{pd})}{1+\exp(\overline{pd})}$. By iterating this equation using dynamics (4)-(6), we obtain a log price-dividend ratio that is an affine function of μ_t, g_t and Σ_t. For convenience of interpretations and in order to obtain pd_t expressions that are easily manageable in our Kalman filter estimation, we directly express pd_t as an affine function of a demeaned expected return and dividend growth ($\hat{\mu}_t = \mu_t - \delta_0$ and $\hat{g}_t = g_t - \gamma_0$) and a demeaned half vectorized covariance matrix ($\hat{\Sigma}_t = vech(\Sigma_t - \mu^{\Sigma})$).

[4] Expression (9) is obtained from a first order Taylor expansion of (1) around the unconditional mean of pd. The approximation error is related to the variance of the price-dividend ratio (see, e.g., Engsted, Pedersen, and Tanggaard (2010)), which is time-varying in our model. However, even in our setting, the approximation turns out to be very precise.

Proposition 1 (Price-dividend ratio) *Under model (3)-(6), the log price-dividend ratio takes the affine form:*

$$pd_t = A - B_1\hat{\mu}_t + B_2\hat{g}_t + B_3\hat{\Sigma}_t, \tag{10}$$

with

$$A = \frac{\kappa + \gamma_0 - \delta_0}{1-\rho}, \tag{11}$$

$$B_1 = \frac{1}{1-\rho\delta_1}, \tag{12}$$

$$B_2 = \frac{1}{1-\rho\gamma_1}, \tag{13}$$

and 1×3 *vector* B_3*, which depends only on parameters* $\rho, \delta_1, \Lambda, M$ *through an expression given explicitly in Appendix A.2.*

Price-dividend ratio pd_t is an affine function of expected returns, expected dividend growth and dividend-return variance-covariance risk. The dependence of pd_t on covariance matrix Σ_t reflects the potential presence of a risk-in-mean effect when $\Lambda \neq 0$. According to intuition, pd_t is decreasing in expected returns and increasing in expected dividend growth. The dependence on Σ_t is more ambiguous and depends on parameters that jointly affect the expected return, expected dividend and variance-covariance risk dynamics.

2.2 Time-varying risks in the present-value model

For Quasi Maximum Likelihood estimation with a Kalman Filter, we assume independence between shocks to returns and dividends $(\varepsilon^D_{t+1}, \varepsilon^r_{t+1})'$ and shocks to time-varying risk ν_{t+1}, in equations (3) and (6), respectively, where we assume $(\varepsilon^D_{t+1}, \varepsilon^r_{t+1})'$ to follow a bivariate standard normal distribution.

Time-varying risks in dynamics (3) and (6) have implications for the conditional risk features of expected returns and expected dividend growth in equations (4) and (5) of our present-value model. Let

$$\tilde{\varepsilon}^D_{t+1} = e_1'\Sigma_t^{1/2}\begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix} \tag{14}$$

and

$$\tilde{\varepsilon}^{r}_{t+1} = e_2' \Sigma_t^{1/2} \begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix} \tag{15}$$

be the total shocks to dividends and returns in dynamics (3), where e_i denotes the $i-$th unit vector in $\mathbb{R}^2$. Campbell and Shiller (1988) approximation (9) implies, together with the explicit PD expression (10):

$$\tilde{\varepsilon}^{r}_{t+1} = \tilde{\varepsilon}^{D}_{t+1} + \rho \varepsilon^{pd}_{t+1}, \tag{16}$$

and

$$\varepsilon^{pd}_{t+1} = B_2 \varepsilon^{g}_{t+1} - B_1 \varepsilon^{\mu}_{t+1} + B_3 \varepsilon^{\Sigma}_{t+1}, \tag{17}$$

where $\varepsilon^{\Sigma}_{t+1} = vech(\nu_{t+1})$. The redundancy of return shocks in equation (16) implies that the state dynamics of our present-value model (3)-(6), can be fully described by the joint dynamics of state vector $(\Delta d_{t+1}, pd_{t+1}, \hat{\Sigma}_t, \hat{g}_t, \hat{\mu}_t)$. Moreover, equation (17) implies that the distribution of the shocks in expected returns and expected dividends is constrained: One of the shocks ε^{g}_{t+1} or ε^{μ}_{t+1} can be defined as a linear combination of the others and an identification assumption has to be imposed.

For identification purposes, we assume that ε^{g}_{t+1} is independent of $(\varepsilon^{D}_{t+1}, \varepsilon^{r}_{t+1})'$ and distributed as $N(0, \sigma_g^2)$. Under this assumption, the conditional variance of return expectation shock

$$\varepsilon^{\mu}_{t+1} = \frac{1}{\rho B_1} (\tilde{\varepsilon}^{D}_{t+1} - \tilde{\varepsilon}^{r}_{t+1} + \rho B_2 \varepsilon^{g}_{t+1} + \rho B_3 \varepsilon^{\Sigma}_{t+1}) \tag{18}$$

can be computed explicitly:

$$Var_t(\varepsilon^{\mu}_{t+1}) = \frac{1}{(\rho B_1)^2} (\Sigma_{11,t} + \Sigma_{22,t} - 2\Sigma_{12,t}) + \left(\frac{B_2}{B_1} \right)^2 \sigma_g^2 + \frac{1}{B_1^2} B_3 Var_t(\varepsilon^{\Sigma}_{t+1}) B_3' \ , \tag{19}$$

where time-varying 3×3 covariance matrix $Var_t(\varepsilon^{\Sigma}_{t+1})$ is an affine function of Σ_t, given in closed-form in Appendix A.1.

In summary, the variance-covariance matrix for the vector of shocks $(\tilde{\varepsilon}^{D}_{t}, \tilde{\varepsilon}^{r}_{t}, \varepsilon^{g}_{t}, \varepsilon^{\Sigma}_{t})'$ in our present-value model is given by:

$$Q_t = \begin{bmatrix} \Sigma_t & 0_{2\times 1} & 0_{2\times 3} \\ 0_{1\times 2} & \sigma_g^2 & 0_{1\times 3} \\ 0_{3\times 2} & 0_{2\times 1} & Var_t(\varepsilon^{\Sigma}_{t+1}) \end{bmatrix} . \tag{20}$$

3 Data and Estimation Strategy

This section describes our data set and introduces our estimation strategy based on a Quasi Maximum Likelihood estimation with a Kalman filter.

3.1 Data

We obtain the with-dividend and without dividend monthly returns on the value-weighted portfolio of all NYSE, Amex and Nasdaq stocks from January 1946 until December 2009 from the Center for Research in Security Prices (CRSP). We use this data to construct annual series of aggregate dividends and prices. We assume that monthly dividends are reinvested in 30-day T-bills and obtain annual series for cash-reinvested log dividend growth. Data on 30-day T-bill rates are also obtained from CRSP.

In order to produce useful information to identify latent time-varying risk components in our present-value model, we consider proxies for the yearly realized volatility of market returns, which can be measured with a moderate estimation error, because market returns are available on a daily frequency. We download daily returns of the value-weighted portfolio of all NYSE, Amex and Nasdaq stocks from 1946 until the end of 2009 from CRSP, and compute a proxy for the yearly realized return variance as the sum of squared daily market returns over the corresponding year:

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2,$$

where $r_{i,t}$ is the market return on day i of year t and N_t is the number of return observations in year t. We do not correct for autocorrelation effects in daily returns (see French, Schwert, and Stambaugh (1987)) nor we subtract the sample mean from each daily return (see Schwert (1989)), since we found the impact of these adjustments to be negligible.

3.2 State space representation

The relevant state variables in model (3)-(6) are the expected return and dividend growth μ_t, g_t and variance-covariance matrix Σ_t. We propose a Kalman filter to estimate the model parameters together with the values of these latent states. To this end, we cast the model in state space form, using demeaned state variables $\hat{\mu}_t$, $\hat{g}_t$ and $\hat{\Sigma}_t$ defined in Section 2.1. In this way, we obtain the following linear transition dynamics with heteroskedastic error terms for present-value model (3)-(6):

$$\begin{aligned}
\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \\
\hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + N'\hat{\Sigma}_t + \varepsilon^\mu_{t+1}, \\
\hat{\Sigma}_{t+1} &= S\hat{\Sigma}_t + \varepsilon^\Sigma_{t+1},
\end{aligned}$$

where 1×3 vector N is a function only of parameter Λ and 3×3 matrix S is a function only of parameter M, both specified explicitly in Appendix A.1.

Observable variables in our model are dividend growth Δd_{t+1}, the price-dividend ratio pd_{t+1} and the market realized volatility RV_{t+1}. Note that while the market return r_{t+1} produces redundant information, relative to linear combinations of Δd_{t+1} and pd_{t+1}, the market realized volatility produces useful information to identify time-varying risk structures, summarized by state $\hat{\Sigma}_t$. This is a sharp difference of our setting, relative to present-value models with constant risks, in which dividend growth and price-dividend ratio provide sufficient information to identify the latent state dynamics.

Measurement equations for Δd_{t+1}, pd_{t+1}, RV_{t+1} are derived from the model-implied expressions for dividend growth, price-dividend ratio and the conditional variance of returns. The measurement equation for dividend growth follows from the first row of dynamics (3):

$$\Delta d_{t+1} = \gamma_0 + \hat{g}_t + \tilde{\varepsilon}^D_{t+1} \ . \tag{21}$$

To obtain a measurement equation for the market realized variance, we model the conditional variance of market returns, $\Sigma_{22,t}$, as an unbiased predictor of RV_{t+1}:

$$RV_{t+1} = \Sigma_{22,t} + \varepsilon^{RV}_{t+1} = \mu^\Sigma_{22} + \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\hat{\Sigma}_t + \varepsilon^{RV}_{t+1}, \tag{22}$$

where the measurement error is such that $\varepsilon^{RV}_{t+1} \sim iidN(0, \sigma^2_{RV})$.

The measurement equation for the log price-dividend ratio in equation (10) contains no error term. As shown by van 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 two transition equations (one of which is vector valued),

$$\hat{\mu}_{t+1} = \delta_1 \hat{\mu}_t + N' \hat{\Sigma}_t + \varepsilon^{\mu}_{t+1}, \tag{23}$$

$$\hat{\Sigma}_{t+1} = S \hat{\Sigma}_t + \varepsilon^{\Sigma}_{t+1}, \tag{24}$$

and three measurement equations:

$$\Delta d_{t+1} = \gamma_0 + \frac{1}{B_2} \left(pd_t - A + B_1 \hat{\mu}_t - B_3 \hat{\Sigma}_t \right) + \tilde{\varepsilon}^{D}_{t+1}, \tag{25}$$

$$RV_{t+1} = \mu^{\Sigma}_{22} + \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \hat{\Sigma}_t + \varepsilon^{RV}_{t+1}, \tag{26}$$

$$\begin{aligned} pd_{t+1} = {} & (1-\gamma_1)A + B_1(\gamma_1 - \delta_1)\hat{\mu}_t + [B_3(S - \gamma_1 I_3) - B_1 N'] \hat{\Sigma}_t + \gamma_1 pd_t \\ & + B_2 \varepsilon^{g}_{t+1} - B_1 \varepsilon^{\mu}_{t+1} + B_3 \varepsilon^{\Sigma}_{t+1}. \end{aligned} \tag{27}$$

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

$$\Theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, M, k, V, N, \sigma_g, \sigma_{RV}).$$

For identification purposes, we impose some parameter constraints. M is assumed lower triangular, with positive diagonal elements less than one. V is assumed diagonal with positive components and $k \geq 2$ is integer. Parameters δ_1 and γ_1 are bounded to be less than one in absolute value, while σ_g and σ_{RV} are constrained to be positive. Overall, the most general version of our present-value model contains 15 parameters. A restricted model with no risk-in-mean effect ($\Lambda = 0$) implies 12 parameters to estimate. Details on the estimation procedure are presented in Appendix B.

4 Results

We estimate our model and consider first the case where no risk-in-mean effect is present ($\Lambda = 0$). This is useful, because in this case the dependence of price-dividend ratio pd_t on $\hat{\mu}_t$ and $\hat{g}_t$ in Proposition 1 is identical to the dependence obtained in the model with constant dividend and return risks. Thus, this setting allows us to obtain simple interpretations for the additional effect of time-varying risks on dividend and return predictability features.

We focus on the structural quantification of the predictability implications of present-value models with time-varying risks, i.e., the characterization of the dynamic features of processes μ_t, g_t and Σ_t for expected returns, expected dividend growth and time-varying risks. First, we quantify the estimated degree of model-implied predictability for returns, dividend growth and return volatility. Second, we analyse the implications of the estimated price-dividend ratio decomposition for the predictability features of returns and dividends by aggregate valuation ratios. Third, we evaluate the consistency of the model implications with a number of well-known predictive regression findings in the literature.

4.1 Estimation results

Table 1, Panel A, presents our QML estimation results for present-value model (23)-(27). The value of the quasi log-likelihood is 853.28.[5] We can formally reject the null hypothesis that expected dividends are constant (i.e. $\gamma_1 = 0$ or 1 and $\sigma_g = 0$) at conventional significance levels. The unconditional expected log return is $\delta_0 = 9\%$, while the unconditional expected growth rate of dividends is $\gamma_0 = 5.6\%$. Expected return features an high autoregressive root, $\delta_1 = 0.907$, which is an indication of a highly persistent process, having an half-life of about 7.5 years. Expected dividend growth are persistent, but less persistent than expected return, with an autoregressive root $\gamma_1 = 0.44$ and an half-life of 1.2 years.[6] For comparison, the estimated persistency of expected returns (expected dividend growth) in a model with constant risks is slightly larger (lower), with an estimated root $\delta_1 = 0.923$ ($\gamma_1 = 0.368$) and half-life 9 years (1.1 years).[7] Estimation results

[5]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 use eight years blocks. Results are unchanged using the stationary bootstrap in Politis and Romano (1994).

[6]The first order autoregressive coefficient is equivalent to $1 - \lambda\Delta t$, where λ is the mean reversion speed and Δt is one year in our setting. The half-life is defined as $\frac{\ln 2}{\lambda}$.

[7]To derive the implications for the model with constant risks, we estimate the model in van Binsbergen and Koijen (2010) for the case of cash-reinvested dividends, using data for the sample period 1946-2009. Our parameter estimates are very similar to their ones, which are based on the sample period 1946-2007. Detailed estimation results are given in Table II of the Supplemental Appendix, which is available from the authors on request.

also indicate persistent dividend and return risks. The autoregressive matrix M in the risk dynamics (6) features both a quite persistent and a less persistent component, with estimated eigenvalues $M_{11} = 0.523$ and $M_{22} = 0.999$, respectively, and a slightly negative out-of-diagonal element $M_{21} = -0.071$. The low estimated degrees of freedom parameter $k = 5$ indicates a fat tailed distribution for the components of Σ_t.

4.2 Basic predictability features

In order to quantify the degree of predictability implied by present-value model (3)-(6), we can measure the fraction of variability in r_t, Δd_t and RV_t explained by μ_{t-1}, g_{t-1} and $\Sigma_{22,t-1}$, respectively.[8] We present in Figure 1 the estimated expected return, expected dividend growth and return variance implied by our present-value model. In each panel, we also plot the fitted values of an OLS regression of r_t, Δd_t and RV_t on the lagged log price-dividend ratio, as well as the actual value of these variables.[9]

The second panel in Figure 1 highlights apparent differences between the expected dividend growth estimated by our present-value model and those of a standard predictive regression: The model-implied expected dividend growth varies more over time and follows more closely the actual dividend growth. A different figure arises for returns in the first panel of Figure 1, where the expected returns estimated by the present-value model and the one implied by the predictive regression are quite smooth and close to each other. These findings are consistent with the different persistence features of expected return and dividend growth estimated by the present-value model with time-varying risks. Finally, the third panel in Figure 1 shows that the filtered conditional variance of returns estimated by the model is a quite good predictor of future realized variances, consistently with the large evidence of predictability in returns second moments produced by the literature.

[8] Let I_t denote the econometrician's information set at time t, generated by the history of dividends, price-dividend ratios and realized volatilities up to time t. Given estimated parameter $\widehat{\Theta}$, the Kalman filter provides expressions to compute filtered estimates of the unknown latent states μ_{t-1}, g_{t-1} and Σ_{t-1}, conditional on I_{t-1}.

[9] The predictive regression for returns takes the form $r_{t+1} = a_r + b_r pd_t + \varepsilon^r_{t+1}$. The predictive regression for dividend growth is $\Delta d_{t+1} = a_d + b_d pd_t + \varepsilon^d_{t+1}$. The one for realized variance is $RV_{t+1} = a_{RV} + b_{RV} pd_t + \varepsilon^{RV}_{t+1}$.

We can quantify the degree of predictability in returns, dividend growth and returns variance within our present-value model and a standard predictive regression, by the following sample R^2 goodness-of-fit measures:

$$R^2_{Ret} = 1 - \frac{\widehat{Var}(r_{t+1} - \mu_t)}{\widehat{Var}(r_{t+1})}, \tag{28}$$

$$R^2_{Div} = 1 - \frac{\widehat{Var}(\Delta d_{t+1} - g_t)}{\widehat{Var}(\Delta d_{t+1})}, \tag{29}$$

$$R^2_{RV} = 1 - \frac{\widehat{Var}(RV_{t+1} - \Sigma_{22,t})}{\widehat{Var}(RV_{t+1})}, \tag{30}$$

where $\widehat{Var}$ denotes sample variances and μ_t, g_t, $\Sigma_{22,t}$ are, with a slight abuse of notation, the estimated expected return, expected dividend growth and conditional return variance in the present-value model and the standard predictive regression model, respectively.

The results in Table 2 show that the estimated R^2 for returns in the present-value model is about 10.38%. The estimated R^2 for dividends is about 11.59%, while the one for the realized variance of returns is about 13.59%. Therefore, expected returns, expected dividend growth and conditional variances of returns seem to explain a relatively large fraction of actual returns, dividend growth and realized variances in our model. The predictability results of standard predictive regressions are consistent with the evidence in the literature. While the R^2 for returns is about 10.5%, the one for dividends is about 1.1%. Finally, the R^2 of predictive regressions for realized variance is about 12.2%. In summary, while the model-implied return predictability is close to the one implied by a standard predictive regression of returns on lagged price-dividend ratio, the dividend growth predictability uncovered by standard predictive regressions is much lower than the model-implied dividend growth predictability. The estimated structure of the price-dividend ratio decomposition in our model offers an intuition for this finding: Since price-dividend ratios are only noisy signals of expected dividend growth, which are contaminated by an expected return component, these predictive regressions are affected by the well-known EIV problem; see also van Binsbergen and Koijen (2009). According to the estimated parameters in Panel B of Table 1, the expected return (expected dividend growth) loads negatively (positively) on price-dividend ratios, with an estimated coefficient $-B_1 = -8.109$ ($B_2 = 1.743$). Therefore, the smooth expected return component has a large loading on the model-implied price-dividend ratio. This large loading is associated with a large fraction of the price-dividend ratio that is driven by expected return shocks. Therefore, the large and persistent expected return component in price-dividend ratios likely obfuscates the predictive power of expected dividend growth for actual dividend growth. Since the expected return component is difficult to estimate from actual returns, due to a very low signal-to-noise ratio, isolating it from aggregate price-dividend ratios in a model-free way is a potentially difficult task. Our model offers a natural way to isolate it, in order to quantify the degree of predictability that is potentially generated in predictive regressions, using aggregate price-dividend ratios adjusted by a smooth proxy of expected return.

4.3 Interpretation of Predictability Results

In order to explain the empirical evidence on time-varying risk-return tradeoffs, model predictions, such as the relatively high return and dividend growth predictability, are more realistically addressed in relation to their consistency with a number of well-known predictive regression findings in the literature. In this section, we test the main model implications for (i) the predictability features of standard predictive regressions with aggregate price-dividend ratios, (ii) long-term predictability properties and (iii) real-time predictability patterns.

We follow a Monte Carlo simulation approach. Starting from the parameter estimates in Section 4.1, we test by Monte Carlo simulation whether model implications are broadly consistent with data-derived implications. We simulate 10000 paths of length 64 years for all state variables and observable variables in our model, following the steps given below:

- Take parameter estimates in Section 4.1.
- Generate 10000 random time series of all shocks in the model, using their conditional covariance matrix (20) and constraint (18).
- Using simulated shocks, obtain recursively the latent states g_t, μ_t and Σ_t from equation (4), (5) and (6), respectively.
- For each simulated sample, compute the actual return and dividend growth from dynamics (3), the actual price-dividend ratio from formula (10) and the actual

realized variance of returns from identity (22).

4.3.1 Joint dividend-return predictability features

The predictive regression results in the data indicate the presence of return predictability (with an R^2 of about 10.49%) and a weak dividend predictability (with an R^2 of about 1.06%) by aggregate price-dividend ratios. As emphasized in Cochrane (2008a), this joint evidence implies sharp restrictions that are useful to validate or test the ability of a model in generating appropriate predictability properties. We follow this insight and compute by Monte Carlo simulation the model-implied joint distribution of estimated R^2's for dividend, return and realized volatility predictive regressions with lagged log price-dividend ratios. Table 3 (columns SV) reports confidence intervals for the estimated degree of predictability in OLS predictive regressions, if the world would be well represented by our model. Our time-varying risks model implies OLS predictive regression results in line with the empirical evidence. For instance, the median OLS R^2s for return and dividend predictive regressions are about 13.84% and 0.91%, respectively, and are very similar to the 10.49% and 1.06% OLS R^2s estimated on real data. Overall, real data OLS R^2s for return, dividend and realized variance predictive regressions are all well inside the 80% confidence interval of estimated OLS R^2s simulated from our present-value model. These results also indicate that the degree of predictability uncovered by standard predictive regressions, relative to true model-implied one, can be strongly downward biased for dividend growth.

It is useful to compare the predictability implications of the model with time-varying return and dividend risks with those of present-value models with constant risks. Columns CV in Table 3 summarize the results of the same simulation exercise for the present-value model with constant risks studied in van Binsbergen and Koijen (2010). Since in this setting market volatilities are constant, the table only contains results for dividend and return predictability features. The median R^2 implied by OLS predictive regressions for returns (dividends) is about 7.34% (3.61%), which is approximately 30% lower (320% higher) than the R^2 estimated in the data. In the model with constant risks, the marginal probability of observing a simulated R^2 for dividend growth predictive regression larger than the one in the data is about 70.8%, while the same probability in the model with time-varying risks is about 47%.

Additional useful predictability insights can be derived from the joint predictability features of dividends and returns. Figure 2 presents scatter plots for the simulated joint distribution of R^2s in standard OLS regressions of returns and dividend growth on the lagged log price-dividend ratio. Right (left) panels present results for the model with time-varying (constant) risks. In each panel, the vertical and horizontal straight red lines report R^2s estimated on real data. The right panel of Figure 2 shows that the model with constant risks tends to generate frequently, i.e., in 57% percent of the cases, R^2 for returns smaller than in the data and R^2 for dividends larger than in the data. That is, the model structure tends to produce frequently an indication of a stronger dividend predictability and a weaker return predictability than in the data. These features are less pronounced in the model with time-varying risks (left panel), where the probability of such R^2-combinations is nearer to 25% (18%). In summary, these findings show that the joint distribution of R^2 coefficients implied by the present-value model with time-varying risks is less biased towards finding particular dividend and return predictability structures that are less consistent with the empirically observed ones.

The basic intuition for the potentially different degrees of predictability implied by standard predictive regressions, relative to the latent expected return and dividend growth processes in our model, is provided in Cochrane (2008b), who derives the relation between state-space models and their observable VAR counterparts in settings with constant risks. He shows that in these models the conditional mean of the predicted variable can contain, in addition to the lagged dividend yield, a long moving average of dividend growths, price-dividend ratios or returns, which would be easy for standard predictive regressions to miss. In a context with time-varying risks, such moving average components can include return and dividend shocks featuring an heteroskedasticity of unknown form that is potentially difficult to model. These features can further increase the difficulties of obtaining efficient predictive regression parameter estimates using linear regression methods.[10]

[10]Using the Kalman filter in Appendix B, we borrow from van Binsbergen and Koijen (2010) to derive approximate expressions for the observable model-implied VAR representation with respect to the econometrician's information set. Such VAR contains several lag polynomials of returns, dividend growth rates and return realized variances. Analytic expressions for the VAR coefficients, as well as the

4.3.2 Long-horizon predictive regressions

Cochrane (2008a) shows how to derive regression coefficients of long-horizon returns and dividend growth on price-dividend ratio, implied by yearly predictive regressions. By applying recursively the following regressions,

$$\begin{aligned} r_{t+1} &= a_r + b_r pd_t + \varepsilon^r_{t+1} \\ \Delta d_{t+1} &= a_d + b_d pd_t + \varepsilon^d_{t+1} \\ pd_{t+1} &= a_{pd} + \phi pd_t + \varepsilon^{pd}_{t+1} \end{aligned}$$

the regression coefficient of long-run returns, $\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$, on pd_t is

$$b^{lr}_r = \frac{b_r}{1-\rho\phi}.$$

Similarly, the regression coefficient of long-run dividend growth is

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

We compute, as in Section 4.3.1, these regression coefficients from the data and by simulation.[11] We find that the model with time-varying risks produces with a frequency of almost 35% estimated long-run coefficients within one standard deviation of the observed sample values (jointly), compared to a frequency of 19% when we simulate from the constant risks model. Figure 3 reproduces, in the left (right) panel, the joint distribution of predictive regression coefficients of dividend growth on log price-dividend ratio, b_d, and long-horizon predictive coefficients of returns, $b^{lr}_r = \frac{b_r}{1-\rho\phi}$, for a model with time-varying (constant) risks. We find that the model with time varying risks tends to be more in line with standard long-horizon predictive regression results than a model with constant risks, which tends to generate a slightly excessive dividend predictability. For instance, at a confidence level of 95%, the long-horizon predictive coefficient for dividend growth, b^{lr}_d, is significantly different from zero 13.7% of the times in the model with constant risks, and only 7% of the times in the model with time-varying risks.

derivations and definitions for the arising errors terms are presented in Section C of the Supplemental Appendix.

[11] Using yearly data from 1946 to 2009 we find $\phi = 0.9162$, $b_r = -0.127$ and $b_d = -0.0166$, so that $b^{lr}_r = -1.1599$ (with a standard deviation of 0.4339) and $b^{lr}_d = -0.1518$ (with a standard deviation of 0.1877). Standard deviations are obtained by the delta method from the standard deviations of b_r, b_d and ϕ. Note that b^{lr}_r and b^{lr}_d satisfy the approximate identity $b^{lr}_d - b^{lr}_r = 1$; see Cochrane (2008a).

4.3.3 Out-of-sample predictability

From the perspective of real-time forecasting, out-of-sample prediction is more relevant than in-sample prediction. Goyal and Welch (2008) study the out-of-sample explanatory power of a large set of predictive variables for market returns, finding that most of them perform worse than the historical mean in forecasting future returns. As explained in Cochrane (2008a), among others, a weak out-of-sample forecasting power does not imply a rejection of the null of predictability itself, but it rather raises important doubts about the practical usefulness of such return forecasts in forming real-time portfolios, given the persistence of forecasting variables and the short span of available data. In our setting, disappointing out-of-sample performance could be explained, e.g., by the difficulty to estimate the latent state variables, exacerbated by the presence of time-varying risks and a highly persistent expected return.

Given the degree of return and dividend growth predictability implied by the estimation results in Table 1, a useful reality check for our present-value model is the absence of excessive incremental out-of-sample forecasting power, relative to a simple mean forecast, when using simple predictive regressions based on aggregate price-dividend ratios. Following Goyal and Welch (2008), we quantify incremental out-of-sample predictive power by the metric:

$$R^2_{i,OS} = 1 - \frac{MSE_{i,A}}{MSE_{i,M}}, \tag{31}$$

where $MSE_{i,A}$ ($MSE_{i,M}$) is the out-of-sample mean squared forecast error of the predictive regression model (historical mean) for returns ($i = r$) and dividend growth ($i = d$), respectively. We simulate 10000 paths of observables and state variables from our estimated present-value models and compute the joint Monte Carlo distribution of $(R^2_{d,OS}, R^2_{r,OS})$ realizations. The scatter plot in Figure 4 summarizes our findings and shows that, under the estimated present-value model, it is unlikely that predictive regressions for returns or dividends can produce a significantly larger out-of-sample predictive power than the historical mean. The estimated probability of the event $\{R^2_{d,OS} \leq 0, R^2_{r,OS} \leq 0\}$ is about 70%, while the estimated probability of the event $\{R^2_{d,OS} > 0, R^2_{r,OS} > 0\}$ is less than 1.7%.

5 Additional Implications

In this section, we study additional implications of the present-value model with time-varying risks, by focusing on (i) time-varying risk features, (ii) the term structure of long-horizon predictability and (iii) the term structure of market risks.

5.1 Basic time-varying risk features

The present-value model with time-varying risks implies a number of useful implications for conditional second moments of returns and dividends, which can be investigated in more detail using the estimated model parameters in Table 1. In this section, we focus on the dynamics of conditional Sharpe ratios and the time-varying co-movement features implied by the model.

5.1.1 Conditional Sharpe ratio dynamics

Using the estimated states $\hat{\mu}_{t-1}$ and $\hat{\Sigma}_{t-1}$ for latent expected returns and variance-covariance risks in our Kalman filter, we find a relatively good degree of variability in both expected market returns and market risk. The average negative correlation between expected returns and return volatilities is about -0.36 , even if in some subperiods these variables tend to move in the same direction (see Figure 5). Therefore, we cannot draw a unique conclusion on the direction of the link between conditional mean and volatility of returns, which seems instead to vary over time.

Conditional Sharpe ratios are defined as the ratio of conditional excess expected returns and conditional volatility, which requires assumptions on the risk-less interest rate r_t^f:

$$SR_t = \frac{E_t(r_{t+1}) - r_t^f}{\sqrt{Var_t(r_{t+1})}} = \frac{\mu_t - r_t^f}{\sqrt{\Sigma_{22,t}}}.$$

To compute our proxy for SR_t, we fix r_t^f as the annualized 30-day T-Bill rate at time t. Figure 6 shows that conditional Sharpe ratios estimated by our model are often counter-cyclical, consistently with the empirical evidence, and quite volatile, which is a useful implication in order to account for part of the "Sharpe ratio volatility puzzle" highlighted in Lettau and Ludvigson (2010), among others. At the same time, we find that the conditional Sharpe ratio implied by a model with constant risks is both less countercyclical and not sufficiently volatile.

5.1.2 Time-varying return, dividend and price-dividend ratio correlations

The model-implied conditional correlation between returns and expected returns is:

$$corr_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^\mu_{t+1}) = \frac{Cov_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^\mu_{t+1})}{\sqrt{Var_t(\tilde{\varepsilon}^r_{t+1}) Var_t(\varepsilon^\mu_{t+1})}}, \tag{32}$$

where $Cov_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^\mu_{t+1}) = \frac{1}{\rho B_1}(\Sigma_{12,t} - \Sigma_{22,t})$, using (18), $Var_t(\tilde{\varepsilon}^r_{t+1}) = \Sigma_{22,t}$ and $Var_t(\varepsilon^\mu_{t+1})$ is given in equation (19). The correlations of returns and dividend growth with the price-dividend ratio are:

$$corr_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^{pd}_{t+1}) = \frac{\Sigma_{22,t} - \Sigma_{12,t}}{\sqrt{\Sigma_{22,t}(\Sigma_{22,t} + \Sigma_{11,t} - 2\Sigma_{12,t})}}, \tag{33}$$

$$corr_t(\tilde{\varepsilon}^D_{t+1}, \varepsilon^{pd}_{t+1}) = \frac{\Sigma_{12,t} - \Sigma_{11,t}}{\sqrt{\Sigma_{11,t}(\Sigma_{22,t} + \Sigma_{11,t} - 2\Sigma_{12,t})}}. \tag{34}$$

Figure 7 reproduces the time series of correlations (32), (33) and (34) in our model, using estimated parameters in Table 1 and the corresponding filtered states in our Kalman filter. We find that the estimated correlation (32) is negative (with a mean of about -0.80), as expected, but it varies substantially over time, especially after the late sixties. Similarly, the average correlation between price-dividend ratio and returns (dividend growth) is positive (negative) with a mean of about 0.97 (-0.25), but the degree of correlation variability increases after the late sixties. While average conditional dividend correlations are roughly consistent with the (unconditional) sample correlation of about -0.25, the average correlation with returns is substantially different from the sample correlation of 0.07. This feature follows from the distinct structure of conditional and unconditional price-dividend ratio variances. Monte Carlo simulations confirm this difference of conditional and unconditional correlations in the model with time-varying risks, with a sample correlation of about 0.075 (-0.29) between log price-dividend ratio and returns (log dividend growth) in line with the empirical evidence; see also Section 4.3.1.

5.2 Term structure of long-horizon predictability

By applying recursively equations (3)-(5), we obtain the following explicit expressions for the model-implied n-year return and dividend growth, in the case where $\Lambda = 0$:

$$\sum_{j=1}^{n} \rho^{j-1} r_{t+j} = \frac{1-\rho^n}{1-\rho}\delta_0 + \frac{1-(\rho\delta_1)^n}{1-\rho\delta_1}\hat{\mu}_t + \sum_{j=1}^{n-1} \rho^j \frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}\varepsilon^{\mu}_{t+j} + \sum_{j=1}^{n} \rho^{j-1}\tilde{\varepsilon}^{r}_{t+j}, \tag{35}$$

$$\sum_{j=1}^{n} \rho^{j-1} \Delta d_{t+j} = \frac{1-\rho^n}{1-\rho}\gamma_0 + \frac{1-(\rho\gamma_1)^n}{1-\rho\gamma_1}\hat{g}_t + \sum_{j=1}^{n-1} \rho^j \frac{1-(\rho\gamma_1)^{n-j}}{1-\rho\gamma_1}\varepsilon^{g}_{t+j} + \sum_{j=1}^{n} \rho^{j-1}\tilde{\varepsilon}^{D}_{t+j}. \tag{36}$$

The model-implied expected n-year return and dividend growth follow as:

$$E_t\left[\sum_{j=1}^{n} \rho^{j-1} r_{t+j}\right] = \frac{1-\rho^n}{1-\rho}\delta_0 + \frac{1-(\rho\delta_1)^n}{1-\rho\delta_1}\hat{\mu}_t, \tag{37}$$

$$E_t\left[\sum_{j=1}^{n} \rho^{j-1} \Delta d_{t+j}\right] = \frac{1-\rho^n}{1-\rho}\gamma_0 + \frac{1-(\rho\gamma_1)^n}{1-\rho\gamma_1}\hat{g}_t. \tag{38}$$

Left panel in Figure 8 plots the estimated term structure of return predictability implied by formula (37). The term structure is quite time-varying at short horizons, but stabilizes with the horizon, around a long-term expected market return of approximately 6%.

The term structures of expected returns and dividend growth in equations (37) and (38) have direct implications for the fraction of long-horizon market returns and dividend variation that can be anticipated within the model. We can quantify by Monte Carlo simulation the percentage variation of n-year returns and dividend growth explained by the model, measured using corresponding R^2s for horizons of $n = 5, 10, 20$ and 30 years.[12] For each horizon n, blue lines in Figure 9 represent the median (solid line) and the 10%- and 90%-quantile (dashed lines) of the model-implied R^2 distribution of returns, both for the model with time-varying risks (left panel) and the model with constant risks (right panel).

We find that the term structure of model-implied predictive power for returns is increasing in both cases, even if for the time-varying risks model it slightly decreases for

[12] We simulate 10000 paths of observables and state variables, as in Section 4.3.1, but we fix the path length to 200 years, instead of 64 years, in order to obtain a sufficient number of simulated long-term returns and dividend growth also for very long horizons.

horizons longer than 20 years. Using the same simulated data, we can also quantify the term structure of long-horizon predictability uncovered by computing R^2 implied by multi-period ahead iterated VAR forecasts.[13] The red lines in Figure 9 display the corresponding Monte Carlo quantiles. Consistent with what is often found in real data for post-war aggregate returns, we find that the uncovered term structures of predictability are increasing for both models.[14] The two bottom panels in Figure 9 report the 10%-, 50%- and 90%-quantile in the Monte Carlo distribution of estimated coefficients in standard predictive regression of n-year returns on lagged log price-dividend ratio, showing that the null hypothesis of no long-horizon return predictability is typically rejected in both models.

Figure 10 reports results for long-horizon dividend growth predictability. The model with time-varying risks features lower and less significant dividend predictability at all horizons, and both models imply a decreasing term structure. For both models, multi-period ahead iterated VAR forecasts uncover a slightly increasing term structure of R^2s with wide confidence intervals, but estimated direct predictive regression coefficients (bottom panels), obtained by estimating simple OLS regressions of n-year dividend growth on the log price-dividend ratio, cannot reject the null of no predictability, as it is usually found in real data.

5.3 Term structure of market risks

Siegel (2008) reports that unconditional (sample) variances realized over long investment horizons are lower than short-horizon variances on a per-year basis. Based on an estimated VAR model for returns and predictors, Campbell and Viceira (2005) conclude that also the term structure of conditional variances is decreasing with the investment horizon. Taking a slightly different view, Pastor and Stambaugh (2010) show that from the perspective of an investor subject to parameter uncertainty and imperfect predictors stocks can be more risky over longer horizons.

The model-implied conditional variance of a n-year return in the setting with time-

[13] Analytic expressions for n-year expected return and dividend growth implied by VAR forecasts are provided in Appendix A.4.

[14] See also, e.g., Chen (2009).

varying risks is derived from equation (35) as follows:

$$
\begin{aligned}
Var_t\left[\sum_{j=1}^{n}\rho^{j-1}r_{t+j}\right] &= \sum_{j=1}^{n-1}\rho^{2j}\left(\frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}\right)^2 Var_t(\varepsilon^{\mu}_{t+j}) + \sum_{j=1}^{n}\rho^{2(j-1)}Var_t(\tilde{\varepsilon}^{r}_{t+j}) + \\
&+ 2\sum_{j=1}^{n-1}\rho^{2j-1}\frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}Cov_t(\varepsilon^{\mu}_{t+j},\tilde{\varepsilon}^{r}_{t+j}), \qquad (39)
\end{aligned}
$$

where $Var_t(\varepsilon^{\mu}_{t+j})$, $Var_t(\tilde{\varepsilon}^{r}_{t+j})$ and $Cov_t(\varepsilon^{\mu}_{t+j},\tilde{\varepsilon}^{r}_{t+j})$ are affine functions of the variance-covariance state Σ_t, given explicitly in Appendix A.5. The model-implied term structure of per-period market risk is thus time-varying. Figure 8 (right panel) plots its estimated dynamics. We find an hump-shaped average term-structure of market risk and the hump is particularly evident in periods of high short-term volatility.

To understand these findings, it is useful to split conditional variance (39) in its three components: A first term reflecting uncertainty about future expected returns, a second term capturing the risk of future return shocks and a third part reflecting the mean reversion of returns, due to the negative correlation between realized and expected return shocks. Figure 11 plots the term structure of market risk and its three components estimated for years 1946, 1986 and 2008, which are associated with an increasing level of one year market volatility. Consistent with the intuition that return mean reversion tends to produce a decreasing term structure of risk, we find that in all cases the mean reversion component has a strongly negative term structure effect, which is however often offset by the impact of the other two components for long horizons. The largest contribution to the average positively-sloped term structure of market risk arises from the uncertainty about future expected returns. As highlighted by Pastor and Stambaugh (2010), the effect of this component is often underestimated or neglected and its relative contribution is positively linked to the degree of predictability in returns, which is large for long horizons. The term structure effect of return shock risk is typically positive and decreasing with the horizon.

The bottom right panel of Figure 11 reports for comparison the decomposition of the (constant) term structure of market risks in the model with constant volatility.[15] In this

[15]The expression for the conditional variance in the constant risks case is analogous to expression (39), but with $Var_t(\varepsilon^{\mu}_{t+j})$, $Var_t(\tilde{\varepsilon}^{r}_{t+j})$ and $Cov_t(\varepsilon^{\mu}_{t+j},\tilde{\varepsilon}^{r}_{t+j})$ that are constant functions only of the model parameters.

case, we find that the effect of future expected return risk is not large enough to offset the impact of the other two term structure components, leading to a downward sloping term structure of risk. Finally, it is interesting to note that even if the time-varying risks model suggests a larger average market risk at very long horizons, it might be difficult to identify this feature without appropriate assumptions about the latent risk dynamics. To illustrate this feature, Figure 12 presents sample variance ratios for horizons from 2 to 30 years, computed from our 64-year sample of observed annual log returns, together with the 10%, 50% and 90%-quantile of the variance ratio's Monte Carlo distribution (obtained from 10000 samples of returns), in the time-varying risks (upper panel) and the constant risks (lower panel) models. For each model, median variance ratios decrease with the horizon and observed sample values are inside the 80% confidence interval of the Monte Carlo simulation.[16]

6 Conclusion

We introduce a tractable latent variable approach with time-varying risks to predictive regressions, in which expected market returns and aggregate dividend growth rates are consistent with the conditional risk features of returns and dividends in a Campbell and Shiller (1988) present-value model. Given exogenous latent processes modeling expected returns, expected aggregate dividend growth and the conditional variance-covariance features of dividends and returns, we use filtering methods to uncover their joint dynamics, as well as the implications for the long-horizon predictability of market returns and the term structure of market risks.

We find that expected dividend growth and expected returns are both time-varying. However, while expected dividends predict a small fraction of actual dividend growth (with average R^2 values of about 0.5%), expected returns explain a large portion of future returns (with average R^2 values of about 47%). The expected dividend growth estimated

[16]The width of these confidence intervals increases rapidly with the horizon, due to the decreasing number of long-horizon returns. The variance ratio at horizon n is defined as the sample variance of n-year returns, divided by n times the sample variance of 1-year returns. Calculations are based on overlapping returns and unbiased variance estimates, as for instance in equation (2.4.37) of Campbell, Lo, and MacKinlay (1997).

by our model is substantially more persistent than expected returns and gives rise to a large price-dividend ratio component that masks the predictive power of valuation ratios for future returns. At the same time, the low persistence of expected market returns produces a moderate model-implied predictability at longer horizons, while the large uncertainty of future expected returns induces an often upward sloping, sometimes hump-shaped, term structure of market risks.

Through these mechanics, our model implies a variety of predictive features that are consistent with a number of findings in the literature. These include (i) the weak return predictability by aggregate price-dividend ratios, (ii) an even weaker predictability of dividends by aggregate price-dividend ratios, (iii) the predictability of market volatilities, (iv) the large volatility and counter-cyclicality of aggregate Sharpe ratios, (v) a stronger evidence of return predictability using standard long-horizon predictive regressions and (vii) an often decreasing term structure of market risks uncovered by variance ratio tests or multi-period ahead iterated VAR forecasts.

Our analysis shows that the degree of dividend and return predictability uncovered by present-value models in a latent variable framework can be quite sensitive to the assumption of time-varying return and dividend risks: Our estimations for the setting with constant risks imply an expected dividend growth less persistent that expected returns, a weak return predictability (with average R^2 values of about 6%) and a large dividend predictability (with average R^2 values of about 99%). The sensitivity of tests for predictability to the assumption of constant or time-varying risks can be illustrated by estimating the model with constant risks on time-series of data simulated from the model with time-varying risks. We find that while the persistence of expected dividend growth and expected returns in the time-varying risk model implies autoregressive roots of 0.996 and 0.541, the average estimated root is 0.602 and 0.914, respectively, with 90% Monte Carlo confidence intervals that do no contain the true model parameters. These point estimates imply a substantially biased estimated degree of persistence, which in turn influences (i) the uncovering of return and dividend predictability relations, (ii) the decomposition of the price-dividend ratio unconditional variation, (iii) the form of the term structure of long-horizon predictability and (iv) the shape of the term structure of market risks.

Our findings also demonstrate the importance of considering time-varying risks and the potential long-run effects of persistent dividend forecasts for reconciling the predictive regression results in the literature. In our model, price-dividend ratios produce a weak and time-varying evidence for return predictability, but price-dividend ratios adjusted by the persistent dividend component have a larger and more time-consistent predictive power in standard predictive regressions, suggesting that smooth real-time proxies of dividend expectations, as proposed in, e.g., Lacerda and Santa-Clara (2010), could prove useful to construct adjusted valuation ratios with better predictive power for future returns. Our model also produces additional evidence on the importance of time-varying risks to uncover predictive return relations. Consistently with the empirical evidence in Johannes, Korteweg, and Polson (2011), we show that while our setting with time-varying risks implies a large degree of return predictability, such a predictability is hardly uncovered by the model with constant risks.

Finally, an important question concerns the extent to which estimated return predictability features could be exploited in real time, e.g., to build successful dynamic portfolio strategies. Given the relatively low estimated degree of persistence of expected market returns in our model, a key aspect is likely the identification of relevant predictive variables, spanning the information set available to investors for building their return expectation μ_t. A first interesting approach in this direction can make use of cross-sectional information on individual stocks, in order to better span investors' conditional information set. Kelly and Pruitt (2011) propose and estimate a predictive factor model with constant risks, in which cross-sectional information from individual stock price-dividend ratios is aggregated to forecast market returns and dividends. In a similar spirit, Brennan and Taylor (2010) extract aggregate discount rate news from equity portfolio returns, based on individual stock characteristics like size and book-to-market. A second useful way of enlarging the predictive conditional information set, e.g., to compute proxies of dividend forecasts, can make use of either synthetic prices of dividend strips, which can be synthesized from options and futures data, or direct quotes for swaps, futures or options on dividends, which have been recently introduced in several exchanges; see, for instance, van Binsbergen, Brandt, and Koijen (2010) and van Binsbergen, Hueskes, Koijen, and Vrugt (2011). This enriched predictive information set can prove useful also for a more accurate identification of the dynamics of the term structure of market risks in present-value models with stochastic risks.

References

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

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Bollersev, T., R. Engle, and J. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy 96*(1), 116–131.

Brandt, M. W. and Q. Kang (2004). On the relation between the conditional mean and volatility of stock returns: A latent var approach. *Journal of Financial Economics 72*(1), 217–257.

Breen, W., L. R. Glosten, and R. Jagannathan (1989). Economic significance of predictable variations in stock index returns. *Journal of Finance 44*(5), 1177–1189.

Brennan, M. J. and A. P. Taylor (2010). Predicting the market using information from equity portfolio returns. Working Paper. UCLA and Manchester University.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics 18*(2), 373–399.

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

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 L. M. Viceira (2005). The term structure of the risk-return tradeoff. *Financial Analysts Journal 61*(1), 34–44.

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.

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.

Engsted, T., T. Q. Pedersen, and C. Tanggaard (2010). The log-linear return approximation, bubbles, and predictability. CREATES Research Papers 2010-37, School of Economics and Management, University of Aarhus.

Ferreira, M. A. and P. Santa-Clara (2010). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, forthcoming.

French, A., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*(1), 3–29.

Gallant, A. R., L. P. Hansen, and G. Tauchen (1990). Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution. *Journal of Econometrics 45*(2), 141–179.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a risk-return tradeoff after all! *Journal of Financial Economics 76*(3), 509–548.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*(5), 1779–1801.

Gourieroux, C. (2006). Continuous time wishart process for stochastic risk. *Econometric Reviews 25*, 177–217.

Gourieroux, C., J. Jasiak, and R. Sufana (2009). The wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics 150*, 167–181.

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.

Harrison, P. and H. H. Zhang (1999). An investigation of the risk and return relation at long horizons. *Review of Economics and Statistics 81*(3), 399–408.

Johannes, M., A. Korteweg, and N. Polson (2011). Sequential learning, predictive regressions, and optimal portfolio returns. Columbia University, Stanford Universit and University of Chicago.

Kelly, B. and S. Pruitt (2011). Market expectations in the cross section of present values. Working Paper. University of Chicago, Booth School of Business and Federal Reserve.

Lacerda, F. and P. Santa-Clara (2010). Forecasting dividend growth to better predict returns. Working Paper. University of Chicago, NBER and Universidade Nova de Lisboa.

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. C. Ludvigson (2010). *Handbook of Financial Econometrics*, Volume 1, Chapter Measuring and Modeling Variation in the Risk-Return Trade-off, pp. 617–690. North-Holland, Amsterdam.

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

Ludvigson, S. C. and S. Ng (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics 83*(1), 171–222.

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

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 (2010). Are stocks really less volatile in the long run? Working Paper. University of Chicago Booth School of Business, NBER and CEPR, and the Wharton School, University of Pennsylvania.

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.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*, 13031313.

Rytchkov, O. (2008). Filtering out expected dividends and expected returns. Working Paper. Temple University.

Schwartz, E. S. and A. B. Trolle (2010). An empirical analysis of the swaption cube. Working Paper. Ecole Polytechnique de Lausanne and Swiss Finance Institute, UCLA and NBER.

Schwert, G. (1989). Why does stock market volatility change over time? *Journal of Finance 44*(5), 1115–1153.

Siegel, J. J. (2008). *Stocks for the Long Run* (4th ed.). McGraw Hill.

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

van Binsbergen, J., M. Brandt, and R. S. Koijen (2010). On the timing and pricing of dividends. Working Paper. Stanford GSB, Fuqua School of Business, and Chicago Booth.

van Binsbergen, J., W. H. Hueskes, R. S. Koijen, and E. B. Vrugt (2011). A term structure of growth. Working Paper. Northwestern Kellogg, APG Asset Management, Chicago Booth and Hedge Fund - Singapore.

van Binsbergen, J. and R. S. Koijen (2009). Likelihood-based estimation of exactly-solved present-value models. Working Paper. University of Chicago, Booth School of Business.

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

Whitelaw, R. F. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance 49*(2), 515–541.

A Present-value model

A.1 Main notation

The state variables of the model are:

$$\begin{aligned}\hat{\mu}_t &= \mu_t - \delta_0, \\ \hat{g}_t &= g_t - \gamma_0, \\ \hat{\Sigma}_t &= vech(\Sigma_t - \mu^{\Sigma}),\end{aligned}$$

where μ^{Σ} is the solution of (8), which is such that

$$vech(\mu^{\Sigma}) = [I_3 - L_2(M \otimes M)D_2]^{-1} k L_2 vec(V),$$

where I_2 is the identity matrix of dimension two, D_2 and L_2 are 2-dimensional duplication and elimination matrices, respectively, i.e for a symmetric 2×2 matrix A:

$$D_2 vech(A) = vec(A), \quad L_2 vec(A) = vech(A),$$

where *vec* denotes vectorization and *vech* half-vectorization.
The dynamics of the state variables are obtained from (4)-(6) as follows:

$$\begin{aligned}\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \\ \hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + N' \hat{\Sigma}_t + \varepsilon^{\mu}_{t+1}, \\ \hat{\Sigma}_{t+1} &= S \hat{\Sigma}_t + \varepsilon^{\Sigma}_{t+1},\end{aligned}$$

where $N = D_2' vec(\Lambda)$ and $S = L_2(M \otimes M)D_2$.
In terms of these demeaned states, the dynamics of realized returns and dividend growth in equation (3) is the following:

$$\begin{aligned}\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \tilde{\varepsilon}^D_{t+1} \\ r_{t+1} &= \delta_0 + \hat{\mu}_t + \tilde{\varepsilon}^r_{t+1}\end{aligned}$$

where

$$\tilde{\varepsilon}^D_{t+1} = e_1' \Sigma_t^{1/2} \begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix},$$

and

$$\tilde{\varepsilon}^r_{t+1} = e_2' \Sigma_t^{1/2} \begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix}.$$

Since ε^{μ}_{t+1} is a linear combination of the other shocks (see equation (18)), to complete the specification of the model we only need to specify the conditional covariance matrix of

$$\begin{pmatrix} \tilde{\varepsilon}^D_{t+1} \\ \tilde{\varepsilon}^r_{t+1} \\ \varepsilon^g_{t+1} \\ \varepsilon^\Sigma_{t+1} \end{pmatrix},$$

which is given by:

$$Q_t = \begin{bmatrix} \Sigma_t & 0_{2\times 1} & 0_{2\times 3} \\ 0_{1\times 2} & \sigma_g^2 & 0_{1\times 3} \\ 0_{3\times 2} & 0_{2\times 1} & Var_t(\varepsilon^\Sigma_{t+1}) \end{bmatrix}, \tag{40}$$

where $Var_t(\varepsilon^\Sigma_{t+1})$ in equations (19) and (20) is given by:

$$Var_t(\varepsilon^\Sigma_{t+1}) = L_2(I_4 + K_{2,2})[M\Sigma_t M' \otimes V + k(V \otimes V) + V \otimes M\Sigma_t M']L_2',$$

with $K_{2,2}$ being the commutation matrix of order two, i.e. the 4×4 matrix such that, for any 2×2 matrix A, $K_{2,2}vec(A) = vec(A')$.

A.2 Price-dividend ratio

In this section we present the detailed derivation of equation (10) in the text. From Campbell-Shiller approximation (9) we have

$$pd_t \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}. \tag{41}$$

By iterating this equation we find:

$$\begin{aligned} pd_t &\simeq \kappa + \rho(\kappa + \rho pd_{t+2} + \Delta d_{t+2} - r_{t+2}) + \Delta d_{t+1} - r_{t+1} \\ &= \sum_{j=0}^{\infty} \rho^j \kappa + \rho^\infty pd_\infty + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}) \\ &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}), \end{aligned} \tag{42}$$

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:

$$\begin{aligned} pd_t &\simeq \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[\Delta d_{t+j} - r_{t+j}] \\ &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[g_{t+j-1} - \mu_{t+j-1}] \\ &= \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^{j} E_t[g_{t+j} - \mu_{t+j}]. \end{aligned} \tag{43}$$

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

$$\begin{aligned} E_t[\hat{\mu}_{t+j}] &= \delta_1^j \hat{\mu}_t + N' S^{j-1} \sum_{k=0}^{j-1} (\delta_1 S^{-1})^k \hat{\Sigma}_t \\ &= \delta_1^j \hat{\mu}_t + N' S^{-1} (S^j - \delta_1^j I_3)(I_3 - \delta_1 S^{-1})^{-1} \hat{\Sigma}_t \end{aligned}$$

and

$$E_t[\hat{g}_{t+j}] = \gamma_1^j \hat{g}_t. \tag{44}$$

Therefore,

$$\begin{aligned} pd_t &\simeq \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^j [\gamma_0 + \gamma_1^j \hat{g}_t - \delta_0 - \delta_1^j \hat{\mu}_t - N' S^{-1} (S^j - \delta_1^j I_3)(I_3 - \delta_1 S^{-1})^{-1} \hat{\Sigma}_t]. \\ &= \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho} + \frac{\hat{g}_t}{1-\rho\gamma_1} - \frac{\hat{\mu}_t}{1-\rho\delta_1} + \\ &\quad + N' \left[(\rho S^2 - (1+\rho\delta_1) S + \delta_1 I_3)^{-1} + \frac{1}{1-\rho\delta_1} (S - \delta_1 I_3)^{-1} \right] \hat{\Sigma}_t \\ &= A + B_2 \hat{g}_t - B_1 \hat{\mu}_t + B_3 \hat{\Sigma}_t, \end{aligned} \tag{45}$$

if $\max |eig(\rho S)| < 1$, which is always true since $\rho < 1$, $S = L_2(M \otimes M) D_2$ and for M to be stationary its eigenvalues (its diagonal elements, since M is assumed to be triangular) are constrained to be lower than 1. The explicit expressions for the present-value coefficients A, B_1, B_2 and B_3 are the following:

$$\begin{aligned} A &= \frac{\kappa + \gamma_0 - \delta_0}{1-\rho}, \\ B_1 &= \frac{1}{1-\rho\delta_1}, \\ B_2 &= \frac{1}{1-\rho\gamma_1}, \\ B_3 &= N' \left[(\rho S^2 - (1+\rho\delta_1) S + \delta_1 I_3)^{-1} + B_1 (S - \delta_1 I_3)^{-1} \right]. \end{aligned}$$

A.3 Asymptotic bias in standard predictive regressions

We have shown in Section **??** that standard predictive regressions of either returns or dividend growth rates on the lagged log price-dividend ratio suffer from an error-in-variables (EIV) problem, which does not disappear as the sample size increases (see Figure **??**). Indeed, the true model for aggregate stock returns is:

$$r_{t+1} = \delta_0 + \hat{\mu}_t + \tilde{\varepsilon}^r_{t+1}, \tag{46}$$

but we wrongly assume the following model to hold:

$$r_{t+1} = a_r + b_r pd_t + \varepsilon^r_{t+1}, \tag{47}$$

where $pd_t = A - B_1\hat{\mu}_t + B_2\hat{g}_t$,[17] and we try to estimate the true parameter $b_r = -1/B_1$ from (47). The p-limit of the OLS slope coefficient is the following:[18]

$$\hat{b}_r \longrightarrow \frac{Cov(pd_t, r_{t+1})}{Var(pd_t)}, \tag{48}$$

where

$$\begin{aligned} Cov(pd_t, r_{t+1}) &= Cov(A - B_1\hat{\mu}_t + B_2\hat{g}_t, \delta_0 + \hat{\mu}_t + \tilde{\varepsilon}^r_{t+1}) \\ &= -B_1 Var(\hat{\mu}_t) + B_2 Cov(\hat{g}_t, \hat{\mu}_t) \\ Var(pd_t) &= B_1^2 Var(\hat{\mu}_t) + B_2^2 Var(\hat{g}_t) - 2B_1B_2 Cov(\hat{g}_t, \hat{\mu}_t) \end{aligned}$$

so that

$$\hat{b}_r \longrightarrow \frac{1}{-B_1 + \frac{B_2^2 Var(\hat{g}_t) - B_1B_2Cov(\hat{g}_t,\hat{\mu}_t)}{B_2Cov(\hat{g}_t,\hat{\mu}_t) - B_1Var(\hat{\mu}_t)}}, \tag{49}$$

and the unconditional variances and covariance of demeaned expected return and dividend growth are the following:

$$\begin{aligned} Var(\hat{\mu}_t) &= \frac{(1 \quad -2 \quad 1)vech(\mu^{\Sigma})}{\rho^2 B_1^2(1-\delta_1^2)} + \frac{B_2^2\sigma_g^2}{B_1^2(1-\delta_1^2)}, \\ Var(\hat{g}_t) &= \frac{\sigma_g^2}{1-\gamma_1^2}, \\ Cov(\hat{g}_t, \hat{\mu}_t) &= \frac{B_2\sigma_g^2}{B_1(1-\gamma_1\delta_1)}. \end{aligned}$$

[17] Remind that, as in Sections 4 and 5, we consider the case in which $B_3 = 0_{1\times 3}$.

[18] Note that here we denote with $\hat{b}_r$ the OLS estimate of the slope coefficient b_r in (47).

Thus, the OLS slope coefficient in the regression of returns on lagged price-dividend ratio is biased and converges to a value that, at the estimated parameters, is lower then the true one in absolute value, resulting in less evidence for return predictability, but at the estimated parameters the bias is small due to the relative persistence of expected dividend growth and returns.

The model for aggregate log dividend growth is:

$$\Delta d_{t+1} = \gamma_0 + \hat{g}_t + \tilde{\varepsilon}^D_{t+1}, \tag{50}$$

while the wrong model is:

$$\Delta d_{t+1} = a_D + b_D pd_t + \varepsilon^D_{t+1}, \tag{51}$$

and we try to estimate the true parameter $b_D = 1/B_2$ from (51). The p-limit of the OLS slope is the following:

$$\hat{b}_D \longrightarrow \frac{Cov(pd_t, \Delta d_{t+1})}{Var(pd_t)}, \tag{52}$$

where

$$\begin{aligned} Cov(pd_t, \Delta d_{t+1}) &= Cov(A - B_1\hat{\mu}_t + B_2\hat{g}_t, \gamma_0 + \hat{g}_t + \tilde{\varepsilon}^D_{t+1}) \\ &= B_2 Var(\hat{g}_t) - B_1 Cov(\hat{g}_t, \hat{\mu}_t) \end{aligned}$$

so that

$$\hat{b}_D \longrightarrow \frac{1}{B_2 + \frac{B_1^2 Var(\hat{\mu}_t) - B_1 B_2 Cov(\hat{g}_t, \hat{\mu}_t)}{B_2 Var(\hat{g}_t) - B_1 Cov(\hat{g}_t, \hat{\mu}_t)}}, \tag{53}$$

Therefore, the OLS slope coefficient in the regression of dividend growth on lagged price-dividend ratio is also biased. This bias is negative and, at the estimated parameters, much more significant than the one for standard return regressions.

A.4 Long-horizon predictability implied by VAR

Let us consider the following VAR model:

$$\begin{aligned} r_{t+1} &= a_r + b_r pd_t + \varepsilon^r_{t+1}, \\ \Delta d_{t+1} &= a_d + b_d pd_t + \varepsilon^d_{t+1}, \\ pd_{t+1} &= a_{pd} + \phi pd_t + \varepsilon^{pd}_{t+1}. \end{aligned}$$

By applying recursively these regressions, we obtain the following expected n-year return:

$$\begin{aligned} E_t\left[\sum_{j=1}^{n}\rho^{j-1}r_{t+j}\right] &= \frac{1-\rho^n}{1-\rho}a_r+\rho b_r a_{pd}\frac{1-(\rho\phi)^{n-1}}{(1-\rho\phi)(1-\rho)}+ \\ &\quad -\rho^n b_r a_{pd}\frac{1-\phi^{n-1}}{(1-\rho)(1-\phi)}+b_r\frac{1-(\rho\phi)^n}{1-\rho\phi}pd_t, \end{aligned} \tag{54}$$

while the expected n-year dividend growth is given by:

$$\begin{aligned} E_t\left[\sum_{j=1}^{n}\rho^{j-1}\Delta d_{t+j}\right] &= \frac{1-\rho^n}{1-\rho}a_d+\rho b_d a_{pd}\frac{1-(\rho\phi)^{n-1}}{(1-\rho\phi)(1-\rho)}+ \\ &\quad -\rho^n b_d a_{pd}\frac{1-\phi^{n-1}}{(1-\rho)(1-\phi)}+b_d\frac{1-(\rho\phi)^n}{1-\rho\phi}pd_t. \end{aligned} \tag{55}$$

A.5 Term structure of conditional variances

The conditional variance of model-implied n-year returns, in equation (35), is the following:

$$\begin{aligned} Var_t\left[\sum_{j=1}^{n}\rho^{j-1}r_{t+j}\right] &= Var_t\left[\sum_{j=1}^{n-1}\rho^{j}\frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}\varepsilon^{\mu}_{t+j}+\sum_{j=1}^{n}\rho^{j-1}\tilde{\varepsilon}^{r}_{t+j}\right] \\ &= \sum_{j=1}^{n-1}\rho^{2j}\left(\frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}\right)^2 Var_t(\varepsilon^{\mu}_{t+j})+\sum_{j=1}^{n}\rho^{2(j-1)}Var_t(\tilde{\varepsilon}^{r}_{t+j})+ \\ &+ 2\sum_{j=1}^{n-1}\rho^{2j-1}\frac{1-(\rho\delta_1)^{n-j}}{1-\rho\delta_1}Cov_t(\varepsilon^{\mu}_{t+j},\tilde{\varepsilon}^{r}_{t+j}), \end{aligned}$$

where

$$\begin{aligned} Var_t(\varepsilon^{\mu}_{t+j}) &= \frac{1}{\rho B_1^2}(1 \quad -2 \quad 1)vechE_t(\Sigma_{t+j-1})+\left(\frac{B_2}{B_1}\right)^2\sigma_g^2+\frac{1}{B_1^2}B_3Var_t(\varepsilon^{\Sigma}_{t+j})B_3', \\ Var_t(\tilde{\varepsilon}^{r}_{t+j}) &= (0 \quad 0 \quad 1)vechE_t(\Sigma_{t+j-1}), \\ Cov_t(\varepsilon^{\mu}_{t+j},\tilde{\varepsilon}^{r}_{t+j}) &= \frac{1}{\rho B_1}(0 \quad 1 \quad -1)vechE_t(\Sigma_{t+j-1}), \end{aligned} \tag{56}$$

and

$$\begin{aligned} E_t(\Sigma_{t+j-1}) &= M^j\Sigma_t(M^j)'+kV(j), \\ V(j) &= V+MVM'+\ldots+M^{j-1}V(M^{j-1})', \\ Var_t(\varepsilon^{\Sigma}_{t+j}) &= L_2(I_4+K_{2,2})[ME_t(\Sigma_{t+j-1})M'\otimes V+k(V\otimes V)+V\otimes ME_t(\Sigma_{t+j-1})M']L_2'. \end{aligned}$$

Note that non-contemporaneous correlation between return and expected return shocks are equal to zero and that the conditional variance of long-run returns is an affine functions of the variance-covariance state Σ_t.

B Kalman Filter

In this section we describe the estimation procedure of the model in Section 2.

We first define an expanded 11-dimensional state vector by the concatenation of the original state variables and the process and observation noise random variables:

$$X_t = \begin{pmatrix} \hat{\mu}_{t-1} \\ \hat{\Sigma}_{t-1} \\ \tilde{\varepsilon}_t^D \\ \tilde{\varepsilon}_t^r \\ \varepsilon_t^g \\ \varepsilon_t^\Sigma \\ \varepsilon_t^{RV} \end{pmatrix},$$

which satisfies:

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

where

$$\varepsilon_{t+1}^X = \begin{pmatrix} \tilde{\varepsilon}_{t+1}^D \\ \tilde{\varepsilon}_{t+1}^r \\ \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^\Sigma \\ \varepsilon_{t+1}^{RV} \end{pmatrix},$$

with conditional variance

$$\tilde{Q}_t \equiv Var_t(\varepsilon_{t+1}^X) = \begin{bmatrix} Q_t & 0_{6\times1} \\ 0_{1\times6} & \sigma_{RV}^2 \end{bmatrix},$$

and Q_t, is given in (20). Moreover,

$$F = \begin{bmatrix} \delta_1 & N' & \frac{1}{\rho B_1} & -\frac{1}{\rho B_1} & \frac{B_2}{B_1} & \frac{B_3}{B_1} & 0 \\ 0_{3\times1} & S & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & I_3 & 0_{3\times1} \\ \multicolumn{7}{c}{0_{7\times11}} \end{bmatrix}, \quad \text{and} \quad \Gamma = \begin{bmatrix} 0_{4\times7} \\ I_7 \end{bmatrix},$$

The measurement equation,

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

is of the form

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

where

$$M_0 = \begin{bmatrix} \gamma_0 - \frac{A}{B_2} \\ (1-\gamma_1)A \\ vec(e_2 e_2')' vec(\mu^{\Sigma}) \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & \frac{1}{B_2} & 0 \\ 0 & \gamma_1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and

$$M_2 = \begin{bmatrix} \frac{B_1}{B_2} & -\frac{B_3}{B_2} & 1 & 0 & 0 & 0_{1\times 3} & 0 \\ -B_1(\delta_1 - \gamma_1) & B_3(S - \gamma_1 I_3) - B_1 N' & -\frac{1}{\rho} & \frac{1}{\rho} & 0 & 0_{1\times 3} & 0 \\ 0 & vec(e_2 e_2')' D_2 & 0 & 0 & 0 & 0_{1\times 3} & 1 \end{bmatrix}.$$

The steps of the filter algorithm are the following:

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

$$\begin{aligned} X_{0,0} &= 0_{11\times 1}, \\ P_{0,0} &= E(X_t X_t'). \end{aligned}$$

- The time-update equations are

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

  where $\tilde{Q}_t$ is computed using the updated state $X_{t,t-1}$.

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

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

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

- Update filtering:

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

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

To estimate model parameters,Θ, we define the log-likelihood for each time t, assuming normally distributed observation errors, as[19]

$$l_t(\Theta) = -\frac{1}{2}\log|S_t| - \frac{1}{2}\eta_t' S_t^{-1}\eta_t, \tag{57}$$

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:

$$\Theta \equiv \arg\max_{\Theta} \mathcal{L}\left(\Theta, \{Y_t\}_{t=1}^{T}\right), \tag{58}$$

with

$$\mathcal{L}\left(\Theta, \{Y_t\}_{t=1}^{T}\right) = \sum_{t=1}^{T} l_t(\Theta), \tag{59}$$

where T denotes the number of time periods in the sample of estimation.[20]

[19] Approximating the true transition density with a Gaussian, makes this a QML procedure. While QML estimation has been shown to be consistent in many settings, it is in fact not consistent in a Kalman filter setting since the conditional covariance matrix $\tilde{Q}_t$ in the recursions depends on the Kalman filter estimates of the volatility state variables rather than the true, but unobservable, values. However, simulation results in several papers have shown this issue to be negligible in practice. See also Schwartz and Trolle (2010).

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

C Tables and Figures

Table 1: Estimation results. We present results of the estimation of the present-value model in equations (3)-(6), imposing $\Lambda = 0_{2\times 2}$. The model is estimated by quasi maximum-likelihood using yearly data from 1946 to 2009 on log dividend growth rates, log price-dividend ratio and realized variance of returns. Panel A presents estimates of the coefficients of the underlying processes. Panel B reports resulting coefficients of the present-value model in equation (10). Bootstrapped standard errors are in parentheses.

Panel A: Quasi maximum-likelihood estimates					
γ_0	δ_0	γ_1	δ_1	M_{11}	M_{21}
0.056	0.090	0.440	0.907	0.523	-0.071
(0.007)	(0.008)	(0.020)	(0.220)	(0.304)	(0.694)
M_{22}	k	V_{11}	V_{22}	σ_g	σ_{RV}
0.999	5	0.0028	0.0001	0.056	0.0215
(0.242)	(2.729)	(0.0041)	(0.0011)	(0.0195)	(0.0087)

Panel B: Implied present-value parameters			
ρ	A	B_1	B_2
0.967	3.364	8.109	1.743
(0.0038)	(0.1273)	(2.6188)	(1.2767)

Table 2: Sample R-squared values of returns, dividend growth and realized variance of returns, computed using equations (28)-(30). In the first row, R^2 are computed from our present-value model, estimated using yearly data from 1946 to 2009, while the second row gives results for a standard OLS predictive regression of observed returns, dividend growth and realized variance on price-dividend ratio.

R-squared values			
	R^2_{Ret}	R^2_{Div}	R^2_{RV}
Present-value model	10.38	15.04	13.59
OLS	10.49	1.06	12.16

Table 3: 10%-, 50%- and 90%-quantile of the empirical distribution of R-squared values of returns, dividend growth and realized variance of returns, computed using (28)-(30), from OLS regressions of simulated returns, dividend growth and realized variance of returns on lagged log price-dividend ratio. Distributions are based on 10000 simulations of length 64 years of the state variables and observables in our time-varying risks model (SV), using the estimated parameters in Table 1, and in the constant risks model (CV) of van Binsbergen and Koijen (2010). Parameters used in the simulations are estimated using yearly price-dividend ratio and dividend growth from 1946 to 2009.

	R^2_{Ret}		R^2_{Div}		R^2_{RV}	
	SV	CV	SV	CV	SV	CV
10%	6.78	2.57	0.03	0.14	0.12	-
50%	13.84	7.34	0.91	3.61	3.42	-
90%	22.71	14.54	5.14	15.88	20.88	-

Figure 1: Expected vs Realized yearly returns, dividend growth and variance of returns. These graphs show the model-implied (filtered) series (red lines) of expected returns μ_t (first panel), expected dividend growth g_t (second panel) and expected return variance $\Sigma_{22,t}$ (third panel), as well as the realized (blue lines) return , r_{t+1}, log dividend growth, Δd_{t+1}, and variance of returns, RV_{t+1}, respectively. The three panels also show the fitted values (green lines) of an OLS regression of realized quantities (r_{t+1}, Δd_{t+1} and RV_{t+1}, respectively) on the lagged log price-dividend ratio.

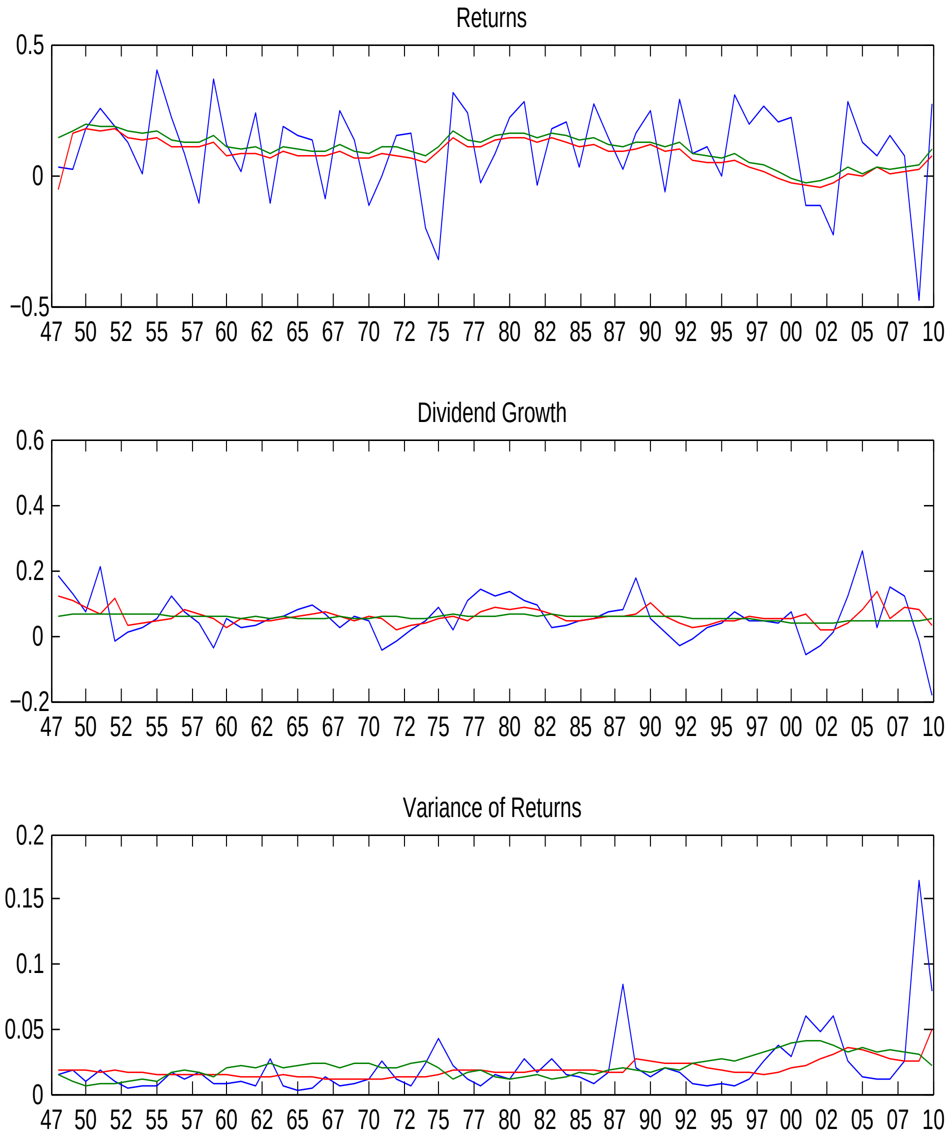


Figure 2: Joint distribution of R^2 of OLS predictive regressions of returns and dividend growth on log price-dividend ratio, obtained simulating 10000 paths of returns, dividend growth and price-dividend ratio from the model with time-varying risks (left) and from the model with constant risks (right). Straight red lines correspond to the observed values in the data. Green dashed lines correspond to median values in the simulations. Only one thousand simulations are plotted for clarity.

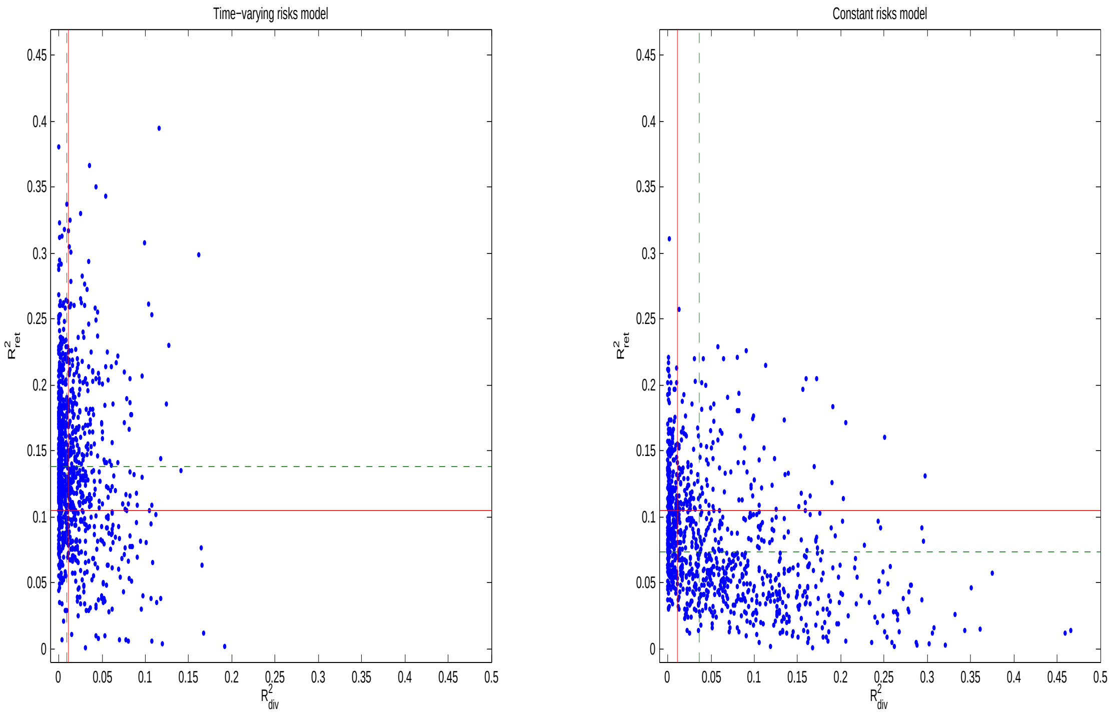


Figure 3: Joint distribution of forecasting coefficient of OLS predictive regressions of dividend growth on log price-dividend ratio, b_d, and long-run predictive coefficient for returns, $b_r^{lr} = \frac{b_r}{1-\rho\phi}$, obtained simulating 10000 paths of returns, dividend growth and price-dividend ratio from our model (left) and from a model with constant risks (right). Straight red lines correspond to the observed values in the data. Green dashed lines correspond to median values in the simulations. Only one thousand simulations are plotted for clarity.

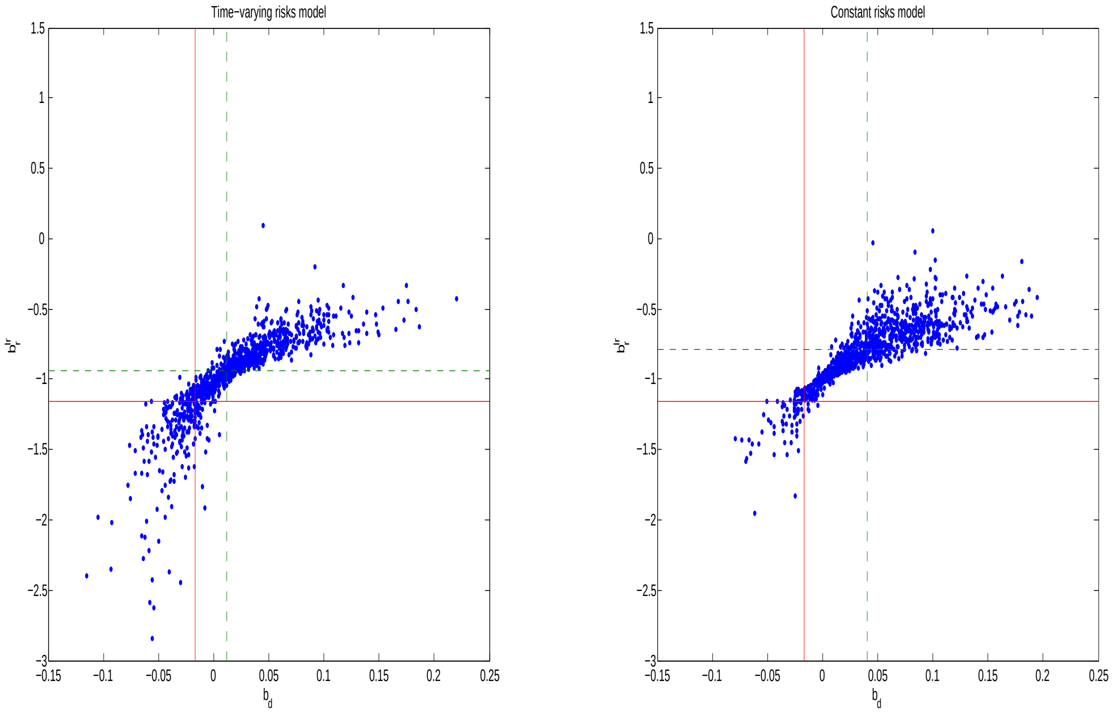


Figure 4: Joint distribution of out-of-sample R^2 of OLS predictive regressions of returns and dividend growth on log price-dividend ratio, obtained from equation (31), by simulating 10000 paths of returns, dividend growth and price-dividend ratio from our model. Straight red lines correspond to the observed values in the data. Only one thousand simulations are plotted for clarity.

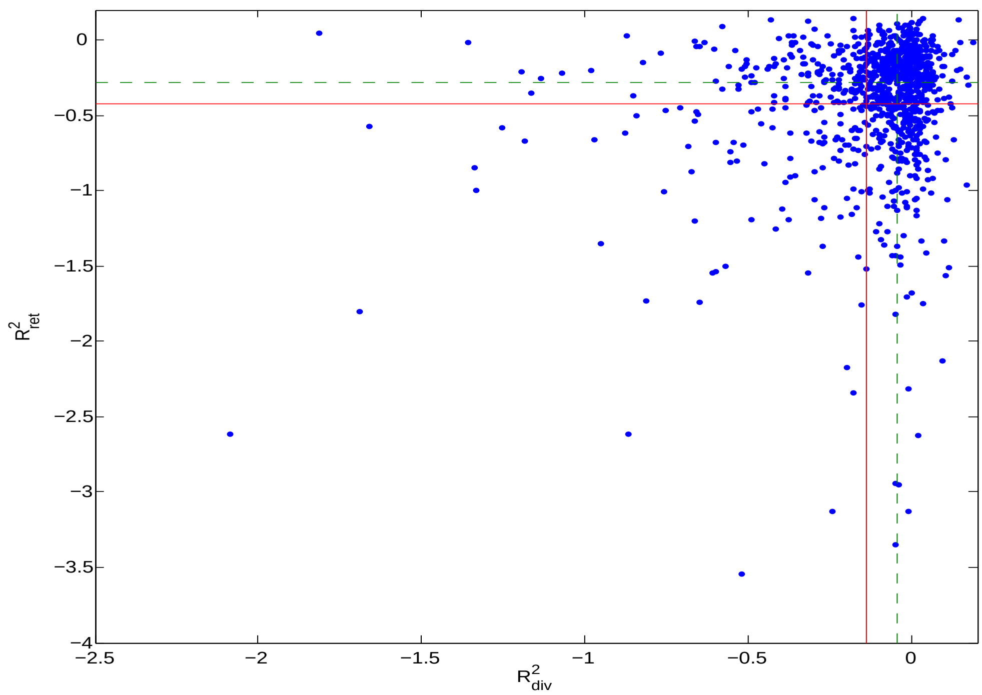


Figure 5: Risk-Return tradeoff. Filtered values of conditional expected returns, μ_t (blue line, left axis) against filtered conditional volatility of returns, $\sqrt{\Sigma_{22,t}}$ (shaded green line, right axis).

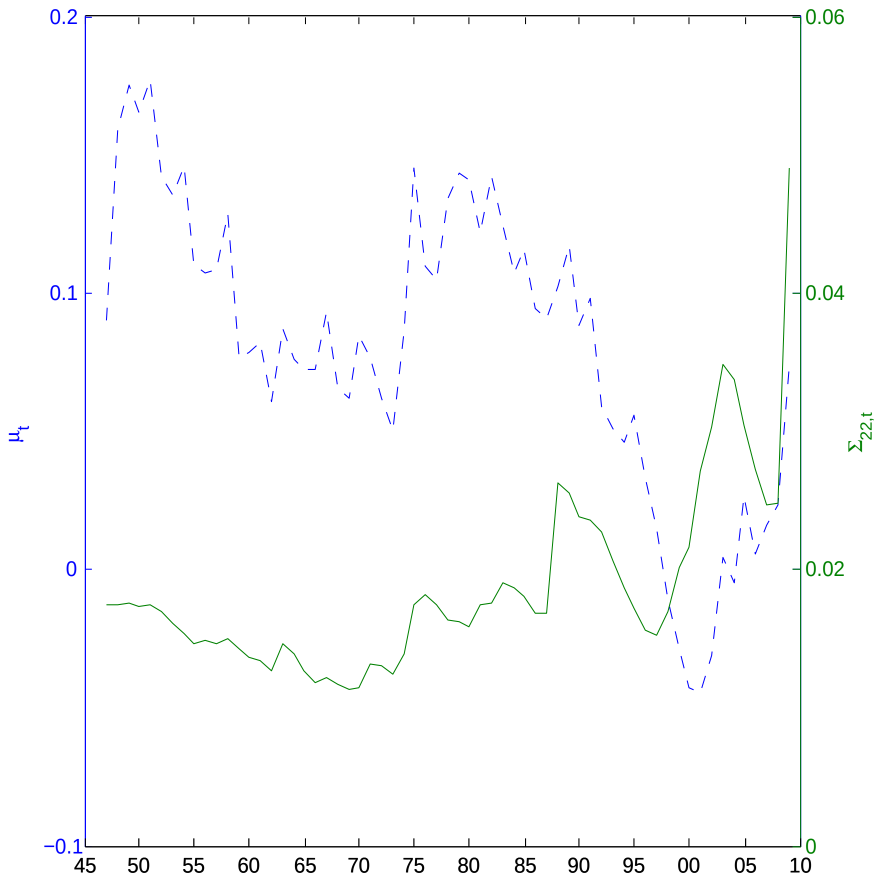


Figure 6: The blue line shows the conditional Sharpe ratio implied by our model, obtained from filtered values of conditional expected returns and conditional volatility of returns, using as risk-free rate the annualized 30-day T-Bill rate at each time t. The red line is obtained in the same way, but for a version of the model with constant risks. Shaded areas corresponds to NBER recessions.

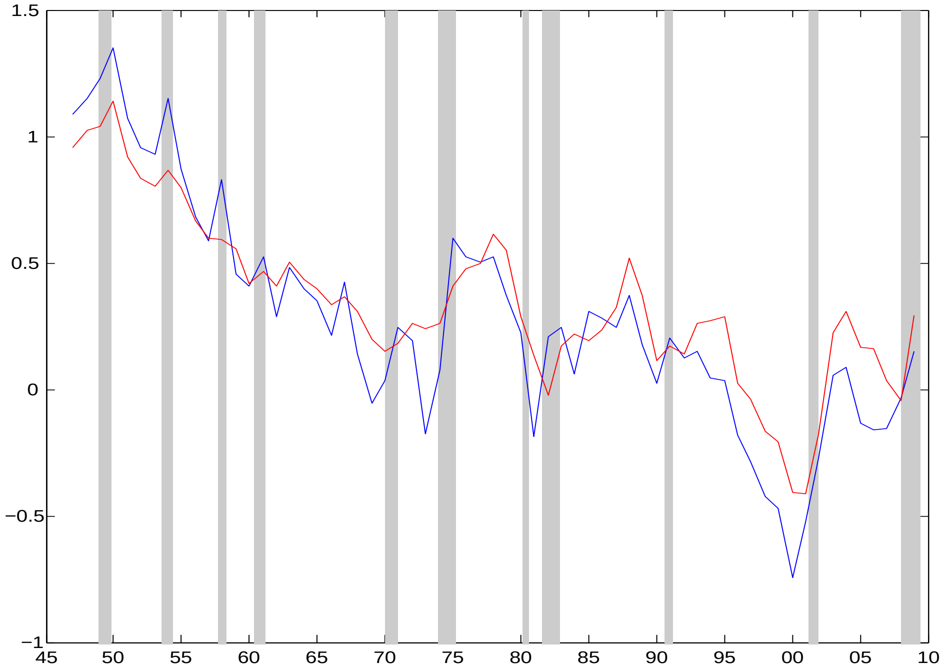
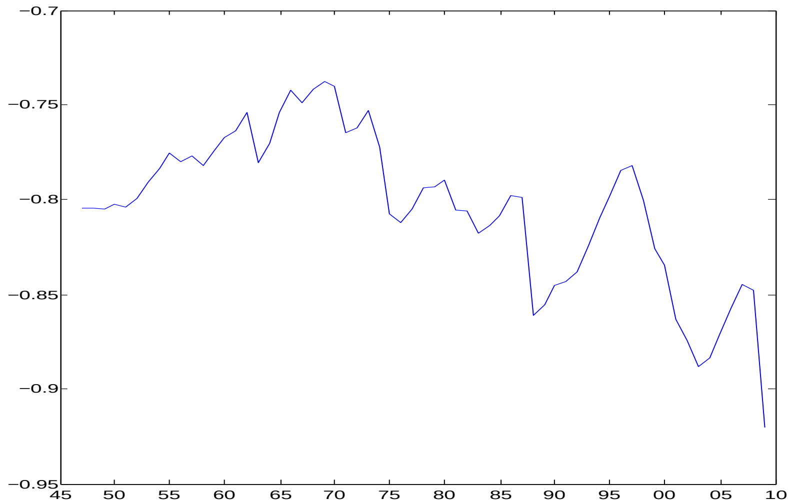


Figure 7: The upper panel shows conditional correlation between expected and unexpected returns, $corr_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^{\mu}_{t+1})$, obtained from estimated present-value parameters and filtered values of the covariance state Σ_t. Shaded areas correspond to NBER recessions. The lower panel displays conditional correlation between shocks in log price-dividend ratio and returns (blue, left axis), $corr_t(\tilde{\varepsilon}^r_{t+1}, \varepsilon^{pd}_{t+1})$, and conditional correlation between shocks in log price-dividend ratio and dividend growth (green, right axis), $corr_t(\tilde{\varepsilon}^D_{t+1}, \varepsilon^{pd}_{t+1})$.



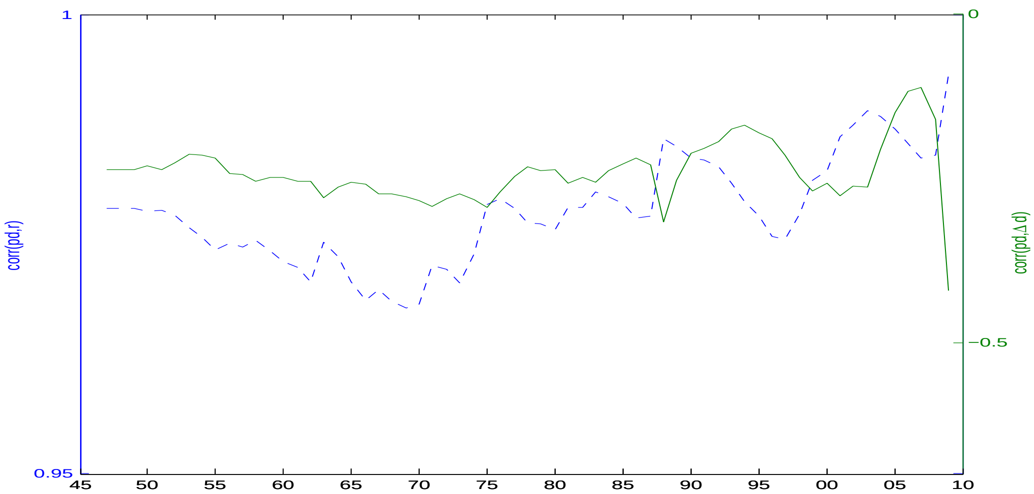


Figure 8: Dynamics of the term structure of the conditional per-period expected long-horizon return (left panel) and standard deviation of long-horizon returns (right panel), from equations (37) and (39), respectively, computed using estimated parameters and filtered state. We consider horizons of 2 to 30 years.

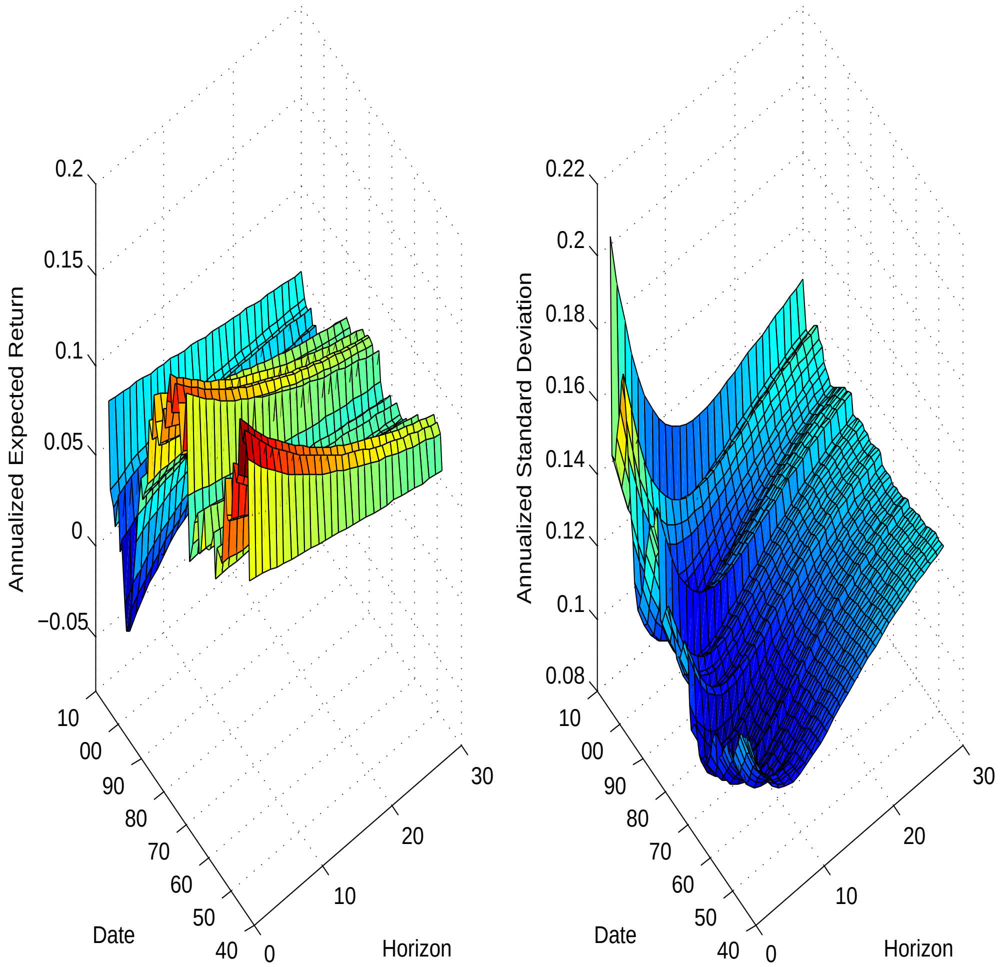


Figure 9: Upper panels show the term structure of predictive R^2 of model-implied long-horizon expected returns (blue lines), and R^2 implied by n-period ahead VAR forecasts (red lines), obtained by simulating 10000 paths of length 200 years of returns, price-dividend ratio and expected returns from our time-varying risks model (left) and from the constant risks model (right). Lower panels display quantiles of the distribution of the simulated coefficients in OLS regressions of n-year returns on log price-dividend ratio. Solid lines show the median of the distribution of R-squared, while dashed lines denote the 10%- and 90%-quantiles, respectively.

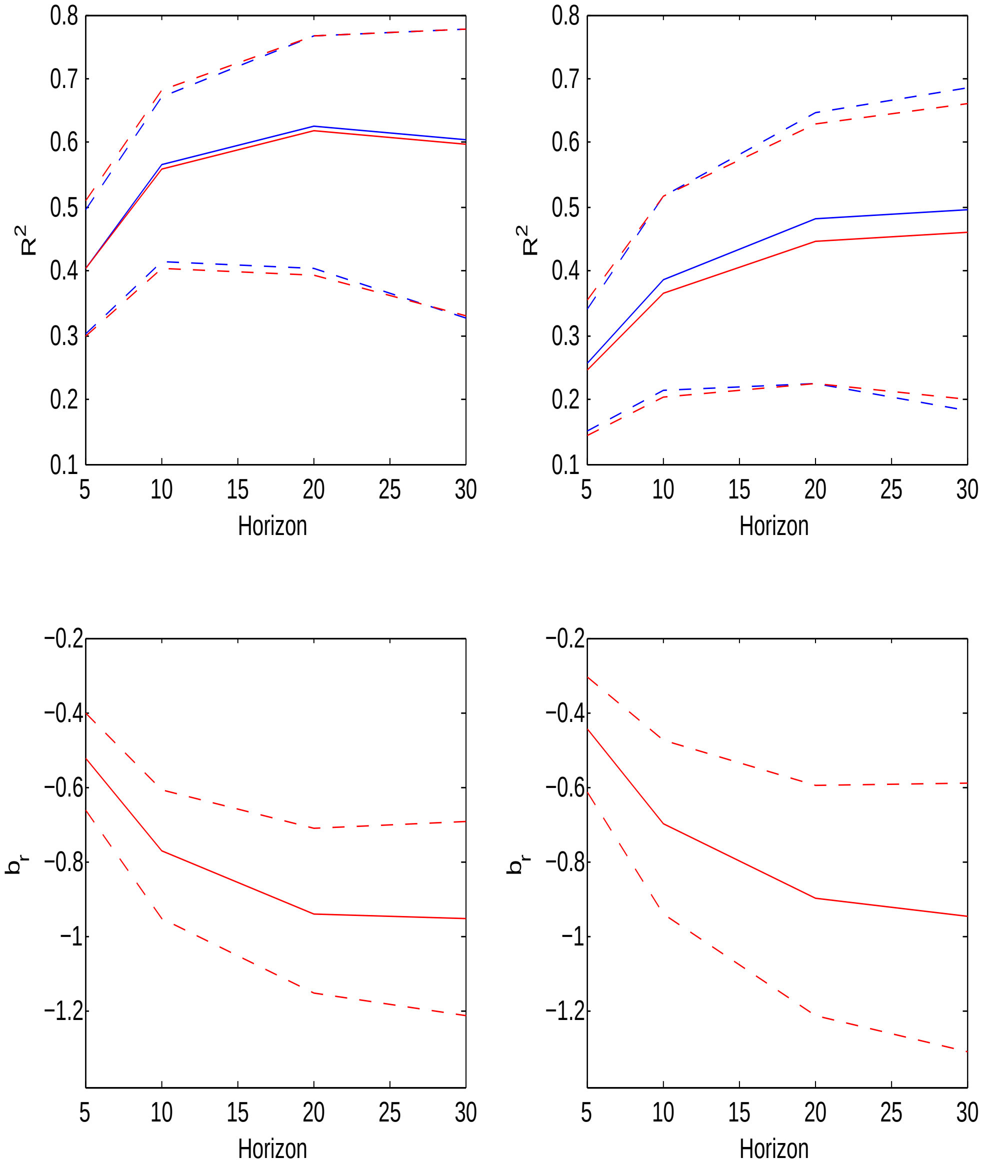


Figure 10: Upper panels show the term structure of predictive R^2 of model-implied long-horizon expected dividend growth (blue lines), and R^2 implied by n-period ahead VAR forecasts (red lines), obtained by simulating 10000 paths of length 200 years of dividend growth, price-dividend ratio and expected dividend growth from our time-varying risks model (left) and from the constant risks model (right). Lower panels display quantiles of the distribution of the simulated coefficients in OLS regressions of n-year log dividend growth on log price-dividend ratio. Solid lines show the median of the distribution of R-squared, while dashed lines denote the 10%- and 90%-quantiles, respectively.

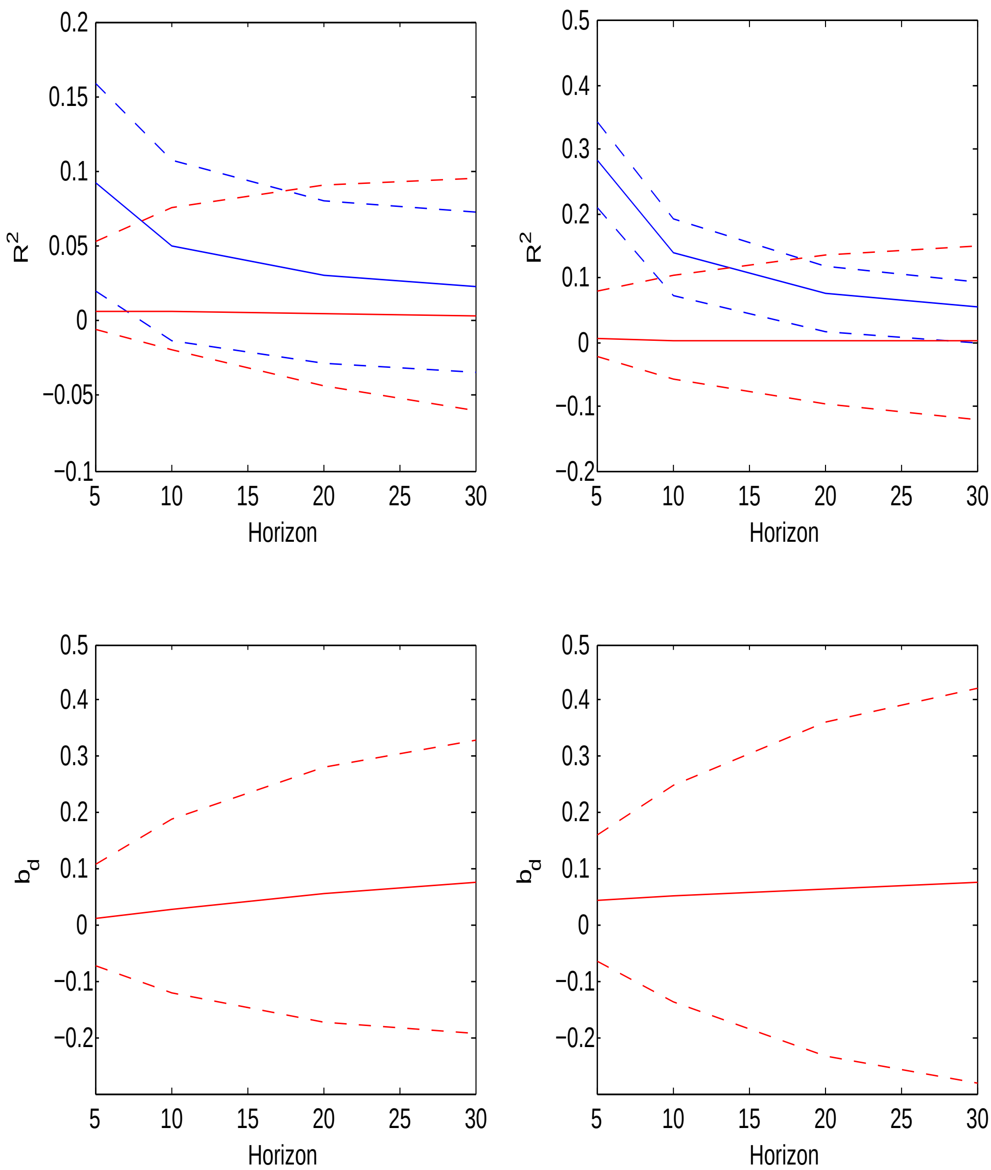


Figure 11: Decomposition of the term structure of the conditional per-period variance of long-horizon returns, computed using estimated parameters and filtered states. The blue line denotes the component of the variance that is due to uncertainty about future expected returns, the red line denotes the component due to future return shocks, while the green line denotes the mean reversion component. The black dashed line denotes the total conditional variance, for horizons of 2 to 30 years. The first three panels show the decomposition implied by our model at different points in time, while the last (bottom right) panel considers the term structure estimated for the constant risks model.

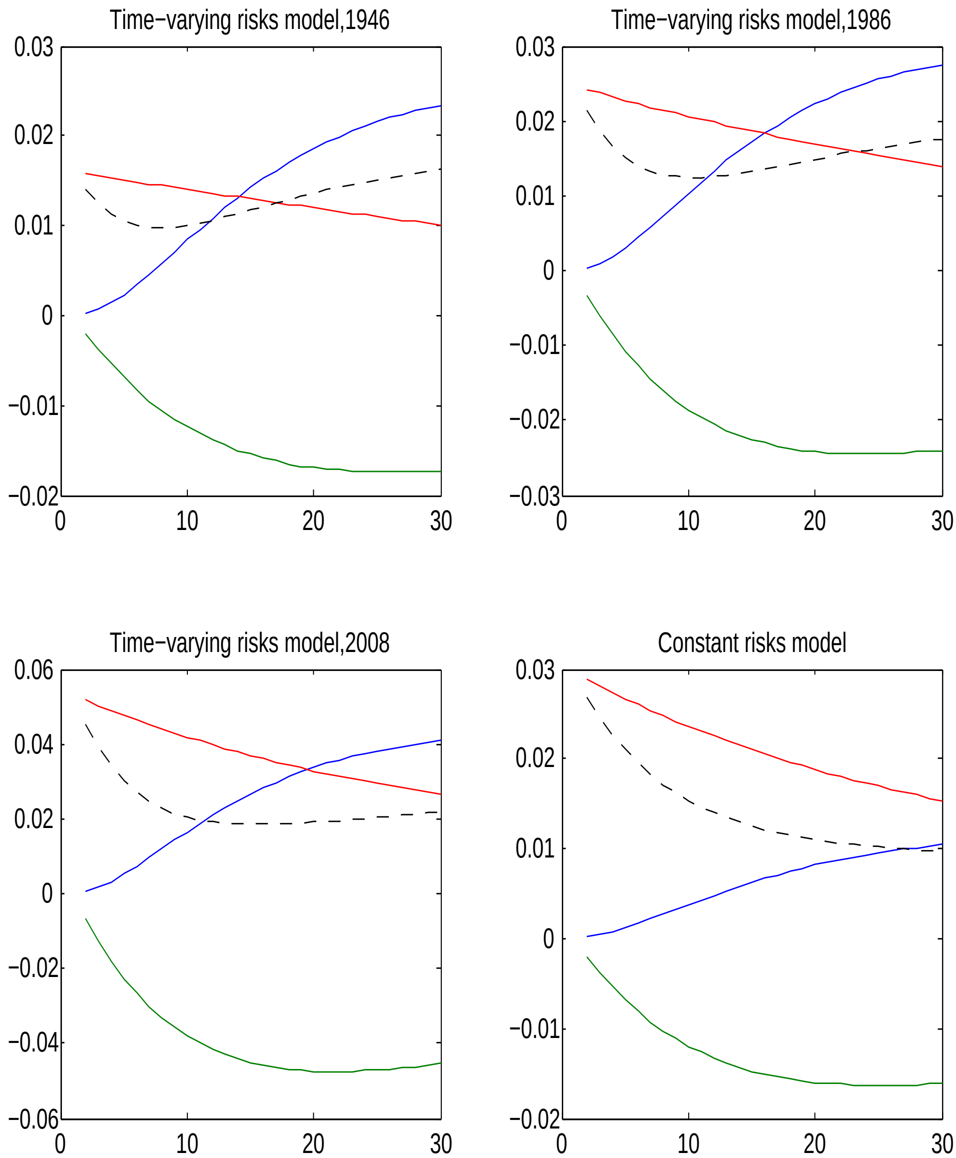


Figure 12: Sample variance ratios for horizons of 2 to 30 years, computed from the 64-year sample of annual log stock market returns (red line) and from 10000 samples of returns simulated from the model (blue lines). Solid blue line denotes the median variance ratios of the 10000 simulations, while dashed lines represent 10%- and 90%-quantiles. In the upper panel, returns are simulated from our time-varying risks model. In the bottom panel, returns are simulated from the constant risks model.