Predictable Risks and Predictive Regression in Present-Value Models

Ilaria Piatti  
Saïd Business School, University of Oxford

Fabio Trojani  
University of Geneva and Swiss Finance Institute
Predictable Risks and Predictive Regression in Present-Value Models*

Ilaria Piatti§ and Fabio Trojani §§

July 31, 2017

Abstract

Using a latent variables approach, we estimate the dynamics of dividends and returns in a tractable present-value model with time-varying risks. Expected returns imply a similar return predictability as under homoskedasticity, while expected dividend growth is more persistent and explains a small fraction of future dividends. Stochastically mean reverting dividends and returns are linked to a time-varying predictability, a stochastic decomposition of price-dividend ratio variances and a closed-form decomposition of cash-flow, discount rate and volatility news in an intertemporal CAPM. The estimated model also implies economically plausible time-varying term structures of dividend-return expectations and risks.

* We are grateful to the Editor (Ken Singleton), the Associate Editor and two referees for many useful suggestions that helped us to improve the paper. We thank Jules van Binsbergen, John Cochrane, Julien Cujean, Greg Duffee, Carlo Favero, René García, Lorenzo Garlappi, Benjamin Golez, Christian Gouriéroux, Ralph Koijen, Roman Kozhan, Yan Li, Paolo Purchia, Giovanni Paopolo, Tarun Ramadorai, Oleg Rychkov, Alessandro Sbuelz, Claudio Tebaldi and the participants of the European Finance Association meeting 2011, Stockholm, the Econometric Society European Meeting 2011, Oslo, the Western Finance Association Meeting 2012, Las Vegas, the SFI Asset Pricing Workshop, Lausanne, the II Workshop on Games and Decisions in Reliability and Risk, Belgrate, the Workshop in Time Series and Financial Econometrics at Bocconi University, the 10th Swiss Doctoral Workshop in Finance, the 8th FINRISK Research Day, the 2011 European Summer Symposium in Financial Markets, the Asset Pricing Workshop 2015 at the University of York, and the seminar participants of IE Business School, Madrid, Warwick Business School, and the Bank of England for valuable comments. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK, project A3) and the Swiss Finance Institute (Project “Term Structures and Cross Sections of Asset Risk Premia”). The usual disclaimer applies.

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

§§ University of Geneva and Swiss Finance Institute, Bd du Pont d’Arve, CH-1211 Geneva 4; e-mail: fabio.trojani@alphacruncher.com
1 Introduction

As emphasized by Campbell and Shiller (1988), a variation of the price-dividend ratio reveals essential information about a time-varying expected return or expected dividend growth. While a large part of the literature has focused on return predictability, Cochrane (2008a) emphasizes the importance of jointly studying dividend and return dynamics, in order to incorporate the key information of the present-value relations between price-dividend ratios, expected returns and expected dividends. Recently, Binsbergen and Koijen (2010) and Rytchkov (2012), among others, have followed this insight to characterize empirically the joint properties of returns and dividend growth, based on a preference-free model with latent dividend and return expectations that explicitly incorporates the present-value relations.

A key feature of return and dividend data, which is not modeled by the present-value approaches above, is a time-varying variance-covariance. A time-varying variance-covariance structure of returns and dividends has intuitively important implications for the joint distribution of realized and expected dividends and returns. For instance, under an IID dividend growth and a single-factor stochastic return volatility, Ang and Liu (2007) show that expected returns and the price-dividend ratio are heteroskedastic and potentially nonlinearly related. More generally, in order to satisfy the present-value identity in presence of multivariate time varying risks, expected returns and expected dividend growth can be both heteroskedastic and stochastically correlated with returns and dividends, giving rise to a time-varying dividend and return predictability and to complex dynamics of the term structures of dividend and return risks.

While these features of the joint dividend-return dynamics are largely unexplored in the literature, they have first-order implications for the joint dynamics of expected vs.

---

1 Besides the ample evidence in the literature of conditional return and dividend growth heteroskedasticity, evidence of a comovement of aggregate cash-flow growth and returns is produced by Belo, Collin-Dufresne, and Goldstein (2014), among others.

2 In various continuous-time present-value models with one dimensional stochastic opportunity set, Ang and Liu (2007) explicitly characterize the joint dynamics of dividends, expected returns, stochastic volatility and prices. Their calibration results show that already in such low dimensional present-value models common specifications of expected returns or return volatility imply rich patterns of dividend growth predictability and heteroskedasticity.
realized dividends and returns. To characterize them empirically, we develop a tractable present-value model with time-varying risks, in which we estimate the joint dynamics of dividends and returns using a latent-variables approach that specifies dividend-return expectation, volatility and correlation processes as hidden state variables. Our model estimation is based on postwar US stock market data and shows that the joint dividend-return dynamics in the present-value model with time-varying risks are different from those under constant risks along several key dimensions.

While expected returns have roughly similar persistence and return predictability properties, the expected dividend process under heteroskedasticity is clearly more persistent and explains a lower fraction of future dividends than under homoskedasticity. Heteroskedastic dividend growths and returns feature a stochastic negative correlation with expected dividend growths and expected returns, which is linked to stochastically mean reverting dividend growth and return processes that imply a time-varying degree of dividend and return predictability. In contrast, homoskedastic dividend growths and returns imply static persistence and predictability properties coupled with positively correlated expected and realized dividend growths. These key implications of our model with heteroskedastic returns and dividends are economically plausible. For instance, Lettau and Wachter (2007), among others, show that the negative correlation between expected and realized dividend growth plays an important role in explaining the value premium and the decreasing term structure of zero-coupon equity volatility documented by Binsbergen, Brandt, and Koijen (2012). Similarly, a time-varying degree of return predictability concentrated in bad times is a well-established stylized fact that can be rationalized in equilibrium economies with heterogenous agents; see Cujean and Hasler (2015), among others.

An important property of our model is that the variance decomposition of the price-dividend ratio under heteroskedasticity is different and highly time-varying. On average, we find that the price-dividend ratio variation is dominated by shocks to expected returns. However, the shocks to expected dividend growth can also have a (time-varying) first-order contribution to the price-dividend ratio variation. The time-varying price-dividend ratio variance decomposition in our model is roughly consistent, e.g., with the evidence in Campbell, Giglio, and Polk (2013), who discuss the difference between the 2000-2002 and
2007-2009 stock market downturns. We document that in the early 2000s price variations were led primarily by discount rate shocks, while in the recent financial crisis discount rates played only a minor role and the downturn was mostly the consequence of worsened cash-flow prospects.

The heteroskedastic multivariate dividend-return dynamics in our model gives rise to nontrivial time-varying term structures of dividend-return expectations and risks. We find that the term structure of expected returns stabilizes around a long term expected return of about 6%. It can be both upward and downward sloping, conditional on the level of short-term expected returns, and it can behave quite differently during distinct crisis periods present in our sample. The slope of the term structure of expected dividend growth in the last part of our sample tends to increase during recessions, consistent with the evidence in Binsbergen, Hueskes, Koijen, and Vrugt (2013). For example, short-term expected dividends sharply declined in 2008-2009 after the financial crisis. The term structure of expected dividend growth in those years was upward sloping until a horizon of about 4 years, suggesting that dividends were expected to grow faster in the medium run than in the short run. However, we also find that the term structure of dividend expectations can be virtually flat during various other recessions and crisis periods in the earlier part of our sample.

The term structure of return volatility is downward sloping on average, reflecting a lower equity risk at longer horizons, but it can also be upward sloping in states where the uncertainty about future expected returns is large, as emphasized by Pastor and Stambaugh (2009). These rich term structure dynamics are a natural consequence of the interplay between the stochastic mean reversion of returns in our model and the time-varying uncertainties of return and expected return shocks: Whenever the stochastic return mean reversion is large enough, the term structure is downward sloping.

Finally, we specify an Intertemporal CAPM consistent with the present-value constraints induced by our model with time-varying risks. In this way, we are able to decompose the stochastic discount factor shocks of a representative agent with recursive preferences into the contributions of news about cash flows, discount rates and future dividend-return volatilities and correlations. Based on the estimated model dynamics, we show that various periods of financial distress with large stochastic discount factor
risk are interpretable in terms of the impact of various news about each element of the future variance-covariance matrix of cash flows and returns.

Our approach builds on the 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 Nieuwerburgh (2008), Campbell and Thompson (2008), Rytchkov (2012), Cochrane (2008a,b), Ferreira and Santa-Clara (2011) and Binsbergen and Koijen (2010), among others. We add to this literature a tractable present-value model for the multivariate heteroskedastic dynamics of returns and dividend growth. By estimating our model with time-varying multivariate risks, we obtain a comprehensive characterization of the joint dividend-return dynamics, which is structurally different from the one under constant risks in many economically important aspects. These key preference-independent relations have remained unexplored to a large extent in the literature. An important exception is the work of Ang and Liu (2007), who explicitly characterize the joint dynamics of dividends, expected returns, stochastic volatility and prices in various continuous-time present-value models under a one dimensional stochastic opportunity set. Preference-based models addressing the equilibrium implications of cash-flow and discount-rate news in presence of single-factor time-varying risks have been proposed only more recently in Bansal, Kiku, Shaliastovich, and Yaron (2014) and Campbell, Giglio, Polk, and Turley (2017).

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 and studies the main model implications. Section 5 concludes.

2 Present-Value Model

As shown in Cochrane (2008a), among others, dividend growth and returns are better studied jointly. Following Campbell and Shiller (1988), we introduce a present-value model with time-varying risks for the joint dynamics of aggregate dividends and market
returns. We denote by
\[ r_{t+1} \equiv \log \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right), \] (1)
the log market return, and by
\[ \Delta d_{t+1} \equiv \log \left( \frac{D_{t+1}}{D_t} \right), \] (2)
the log dividend growth. \( \mu_t \equiv E_t[r_{t+1}] \) and \( g_t \equiv E_t[\Delta d_{t+1}] \) are the expected return and dividend growth, conditional on investors’ information at time \( t \), while \( \Sigma_t \) is the conditional variance-covariance matrix of returns and dividend growth.

We specify \( \mu_t, g_t \) and \( \Sigma_t \) as latent processes that model the time-varying second-order structure of returns and dividends:
\[
\begin{pmatrix}
\Delta d_{t+1} \\
r_{t+1}
\end{pmatrix} = \begin{pmatrix}
g_t \\
\mu_t
\end{pmatrix} + \Sigma_t^{1/2} \begin{pmatrix}
\varepsilon^D_{t+1} \\
\varepsilon^r_{t+1}
\end{pmatrix},
\] (3)
where \( (\varepsilon^D_{t+1}, \varepsilon^r_{t+1})' \) is a bivariate Gaussian white noise. Expected returns and dividends follow autoregressive processes:
\[
\begin{align*}
g_{t+1} &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \varepsilon^g_{t+1}, \quad (4) \\
\mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + \varepsilon^\mu_{t+1}, \quad (5)
\end{align*}
\]
with parameters \( \gamma_0, \gamma_1, \delta_0, \delta_1 \). Zero mean shocks \( (\varepsilon^g_{t+1}, \varepsilon^\mu_{t+1})' \) in this expectation dynamics feature a natural time-varying risk structure, which is implied by the present-value constraints on the dynamics of dividends, returns and price-dividend ratios when \( \Sigma_t \) is time-varying.

We specify \( \Sigma_t \) as a persistent process for variance-covariance matrices that ensures a tractability comparable to the case of constant risks. Precisely, \( \Sigma_t \) follows the Wishart Autoregressive process of order one (\( WAR(1) \)) introduced in Gourieroux (2006) and Gourieroux, Jasiak, and Sufana (2009):
\[
\Sigma_{t+1} = \mu_{\Sigma} + M(\Sigma_t - \mu_{\Sigma})M' + \nu_{t+1},
\] (6)

3This is a simple benchmark model, although already quite flexible. A more sophisticated persistence structure of the second moments of returns and dividends could be obtained by using a Wishart process of higher order, i.e. \( WAR(n) \), but we focus on the most parsimonious specification to understand the first-order effects of introducing time-varying risks.
where $M$ is a $2 \times 2$ autoregressive parameter matrix, $\mu_\Sigma$ is the unconditional mean of $\Sigma_t$ and $\nu_{t+1}$ is a $2 \times 2$ IID error term such that $kV + \nu_{t+1}$ is Wishart distributed with $k$ degrees of freedom and scaling matrix $V := \frac{1}{k}(\mu_\Sigma - M \mu_\Sigma M')$.\footnote{This parameterization of the $WAR(1)$ process is equivalent to the more standard parameterization:}

Wishart dynamics (6) yields symmetric positive-definite $\Sigma_t$ realizations for $k > 2$ and is a natural specification of stochastic multivariate risks. It implies a useful degree of flexibility in the variance-covariance dynamics, e.g., by admitting a negative conditional dependence between variances (diagonal elements of $\Sigma_t$) and covariances (out-of-diagonal elements) with unrestricted signs. These characteristics are useful for reproducing the empirical features of return and dividend risks. In contrast to multivariate GARCH-type dynamics, Wishart dynamics (6) implies closed-form affine expressions for the term structures of dividend-return variances and covariances, which simplifies the characterization of volatility news and of the properties of the term structures of risks under our modelling approach.\footnote{The closed-form expressions for the term structures of risks follow from the closed-form expressions of the conditional moments in the $WAR(1)$ model.}

We estimate the model with Pseudo Maximum Likelihood (PML), based on a Kalman filter with a Gaussian pseudo likelihood for the distribution of $(\varepsilon_{D,t+1}, \varepsilon_{r,t+1})'$ and under independence between $(\varepsilon_{D,t+1}, \varepsilon_{r,t+1})'$ and $\nu_{t+1}$ shocks. As we work with yearly frequencies and relatively modest sample sizes in the empirical part of the paper, we refrain from over parameterizing the model with, e.g., some additional specification of volatility-feedbacks. However, note that our model can generate asymmetric feedbacks between first and second conditional moments under the filtered dynamics of our Kalman filter, as discussed in more detail in section F of the Supplemental Appendix.\footnote{See also the VAR representation of the model in Supplemental Appendix E, which shows explicitly how filtered expected returns and dividends depend on all historical observables, i.e., dividend growth and price-dividend ratio, with coefficients that are functions of the whole history of the conditional variance-covariance matrix of returns and dividends.}
2.1 Price-dividend ratio

Let \( pd_t \equiv \log \frac{P_t}{D_t} \) be the log price-dividend ratio. We obtain the expression for the price-dividend ratio in our model using Campbell and Shiller (1988) log linearization:

\[
    r_{t+1} \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t, \tag{7}
\]

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 under dynamics (4)-(6), we obtain a log price-dividend ratio that is an affine function of \( \mu_t \) and \( g_t \). For convenience and in order to obtain easily manageable \( pd_t \) expressions in our Kalman filter, 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 \( \hat{g}_t = g_t - \gamma_0 \)).

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

\[
    pd_t = A - B_1 \hat{\mu}_t + B_2 \hat{g}_t, \tag{8}
\]

with

\[
    A = \frac{\kappa + \gamma_0 - \delta_0}{1 - \rho}, \tag{9}
\]

\[
    B_1 = \frac{1}{1 - \rho \delta_1}, \tag{10}
\]

\[
    B_2 = \frac{1}{1 - \rho \gamma_1}. \tag{11}
\]

The proof is given in Section A of the Supplemental Appendix.

\( pd_t \) is an affine function of expected returns and expected dividend growth. According to intuition, it is decreasing in expected returns and increasing in expected dividend growth. The dependence of \( pd_t \) on \( \hat{\mu}_t \) and \( \hat{g}_t \) in Proposition 1 has the same form as in a model with constant dividend and return risks. Thus, our setting allows a direct comparison between the implications of heteroskedastic expected dividends and returns induced by present-value constraints with time-varying risks and those of present-value models with homoskedastic conditional expectations.

\(^7\)Expression (7) follows 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 (2012)), which is time-varying in our model. In our data the identity is virtually exact and the average approximation error is about 0.2\%. 

8
Note that it is possible to specify an extended model with multivariate volatility feedbacks, using six additional parameters and a conditional mean dynamics (3) that is affine in $\Sigma_t$. This setting induces closed-form $pd$-ratios that are affine in $\mu_t$, $g_t$ and $\Sigma_t$, extending the result in Proposition 1. We explore the relevance of volatility feedbacks in Section F of the Supplemental Appendix, by estimating an extended version of our model that incorporates volatility feedback in expected returns, but we do not obtain evidence of statistically significant volatility feedbacks. Therefore, we focus on models without volatility asymmetries in conditional means. As a consequence, all differences between present-value models with constant and time-varying risks in our study arise directly from the heteroskedasticity of dividends/returns versus expected dividends/returns in the Campbell-Shiller identities, and not from a different functional form of the $pd$-ratio in presence of volatility feedbacks.

The next section addresses the time-varying risk properties of expected returns and dividends in our model.

### 2.2 Time-varying risks in the present-value model

The time-varying risks in dynamics (3) and (6) have direct implications for the conditional risk features of expected returns and dividend growth in equations (4) and (5). Let

$$\tilde{\epsilon}^D_{t+1} = e_1' \Sigma_{t}^{1/2} \begin{pmatrix} \epsilon^D_{t+1} \\ \epsilon^r_{t+1} \end{pmatrix}$$

and

$$\tilde{\epsilon}^r_{t+1} = e_2' \Sigma_{t}^{1/2} \begin{pmatrix} \epsilon^D_{t+1} \\ \epsilon^r_{t+1} \end{pmatrix}$$

be the total shocks to dividends and returns in dynamics (3), where $e_i$ denotes the $i$-th unit vector in $\mathbb{R}^2$. Approximation (7) implies, together with expression (8):

$$\tilde{\epsilon}^r_{t+1} = \tilde{\epsilon}^D_{t+1} + \rho \tilde{\epsilon}^{pd}_{t+1}; \tilde{\epsilon}^{pd}_{t+1} = B_2 \tilde{\epsilon}^g_{t+1} - B_1 \tilde{\epsilon}^\mu_{t+1},$$

so that

$$\frac{1}{\rho} \left( \tilde{\epsilon}^r_{t+1} - \tilde{\epsilon}^D_{t+1} \right) = B_2 \tilde{\epsilon}^g_{t+1} - B_1 \tilde{\epsilon}^\mu_{t+1}.$$
In equation (15), the difference of return and dividend shocks is proportional to a particular linear combination of expected return and expected dividend shocks. Therefore, the present-value relation constraints the second moments of expected returns and dividend growth in a very explicit way. This insight can be exploited to estimate the joint time-varying risk features of returns, dividend growth, expected returns and expected dividend growth. However, as shocks to expected returns and expected dividend growth are not identifiable individually, an identification assumption is needed.

We identify shocks to expected returns and dividend growth with two parameters $p_1$ and $p_2$, which control the weight of return and cash flow shocks on expected return and expected dividend growth shocks:

\[
\varepsilon^g_{t+1} = \frac{1}{\rho B_2} \left( p_1 \tilde{\varepsilon}^r_{t+1} - p_2 \tilde{\varepsilon}^D_{t+1} \right), \\
\varepsilon^\mu_{t+1} = \frac{1}{\rho B_1} \left( (p_1 - 1) \tilde{\varepsilon}^r_{t+1} - (p_2 - 1) \tilde{\varepsilon}^D_{t+1} \right). 
\]

By construction, this parsimonious identification scheme satisfies the present-value constraint (15). In parallel, it is compatible with time-varying conditional second moments of discount rate and cash flow expectations. As we show below, these model features are essential for generating both flexible term structures of expectations and risks and time-varying predictability properties.\(^8\)

Under identification scheme (16)–(17), the variances and covariance of discount rate and cash flow expectation shocks are:

\[
Var_t (\varepsilon^g_{t+1}) = \frac{p_1^2 \Sigma_{22,t} + p_2^2 \Sigma_{11,t} - 2p_1 p_2 \Sigma_{12,t}}{\rho^2 B_2^2}, \\
Var_t (\varepsilon^\mu_{t+1}) = \frac{(p_1 - 1)^2 \Sigma_{22,t} + (p_2 - 1)^2 \Sigma_{11,t} - 2(p_1 - 1)(p_2 - 1) \Sigma_{12,t}}{\rho^2 B_1^2}, \\
Cov_t (\varepsilon^g_{t+1}, \varepsilon^\mu_{t+1}) = \frac{p_1(p_1 - 1) \Sigma_{22,t} + p_2(p_2 - 1) \Sigma_{11,t} - (2p_1 p_2 - p_1 - p_2) \Sigma_{12,t}}{\rho^2 B_2 B_1}.
\]

\(^8\)A more straightforward identification assumption is $\varepsilon^\mu_{t+1} = -\tilde{\varepsilon}^r_{t+1}/(\rho B_1)$ (and $\varepsilon^g_{t+1} = -\tilde{\varepsilon}^D_{t+1}/(\rho B_2)$), i.e., a proportionality between return (dividend) and expected return (expected dividend) shocks. Such proportionality assumptions are typical, e.g., for the state dynamics of most long-run risk models, such as Bansal and Yaron (2004). In our context, these assumptions would imply a perfect correlation between shocks to dividends or returns and shocks to expected dividends or expected returns, which gives rise to restrictive dynamics for the term structures of risk not supported by the data.
where $\Sigma_{ij,t}$ is the $ij$-component of $\Sigma_t$. Therefore, the variance-covariance structure of dividend and return expectations only depends on the variance-covariance structure of dividends and returns, the sensitivity of price-dividend ratios to expected returns and expected dividend growth (parameters $B_1, B_2$), and the loadings of expectation shocks on return and dividend shocks (parameters $p_1, p_2$). The model-implied conditional variance of the $pd$ ratio is:

$$Var_t(pd_{t+1}) = \frac{1}{\rho^2} (\Sigma_{22,t} + \Sigma_{11,t} - 2\Sigma_{12,t}).$$

(21)

Therefore, the fraction of $pd$ variance explained by shocks to expected discount rate and cash flows depends on parameter $B_1, B_2, p_1, p_2$ and the time-varying variance-covariance matrix $\Sigma_t$. Only in the very special case $p_1 = p_2$, the decomposition of the conditional $pd$ variance is constant over time. In this particular case, expectation shocks $\varepsilon^g_{t+1}$ and $\varepsilon^\mu_{t+1}$ are perfectly positively correlated.$^9$

The conditional covariances between expectations and realizations of returns and dividend growth in our model are also time-varying:

$$Cov_t(\varepsilon^\mu_{t+1}, \tilde{\varepsilon}^r_{t+1}) = \frac{(p_1 - 1)\Sigma_{22,t} - (p_2 - 1)\Sigma_{12,t}}{\rho B_1},$$

(22)

$$Cov_t(\varepsilon^g_{t+1}, \tilde{\varepsilon}^D_{t+1}) = \frac{p_1\Sigma_{12,t} - p_2\Sigma_{11,t}}{\rho B_2}.$$  

(23)

As we show in more detail below, this stochastic co-movement allows to incorporate stochastically mean reverting dividend growths or returns and a time-varying degree of dividend and return predictability into our model.

### 2.3 Nested models

Our modelling approach nests several interesting dynamics for dividends and returns as special cases. For instance, a setting with constant risks arises for a $WAR(1)$ model such that $M$ is an identity matrix, as in this case $V$ is a matrix of zeros. Conversely, a model with constant expected dividend growth in equation (4) is parameterized by the

$^9$Figure I in the Supplemental Appendix illustrates the range of possible values of the conditional correlation between $\varepsilon^g_{t+1}$ and $\varepsilon^\mu_{t+1}$ for various parameters $p_1$ and $p_2$, when $\Sigma_t$ is fixed at the sample variance-covariance matrix of returns and dividend growth in our data set. This figure outlines the intrinsic flexibility of this specification, with values of the correlation that range from perfectly negative to perfectly positive.
constraints $\gamma_1 = p_1 = p_2 = 0$, while a setting with constant expected returns in equation (5) emerges for $\delta_1 = 0$ and $p_1 = p_2 = 1$.

Note that the null hypotheses of constant expected dividend growth or constant expected returns are mutually exclusive in our setting. According to the functional form for the price-dividend ratio in Proposition 1, this feature ensures consistency with the empirical evidence of time-varying price-dividend ratios. Moreover, a constant expected dividend growth or expected return does not in general imply in our setting IID dividend growths or returns, because the shocks in the dividend-return dynamics (3) are in general heteroskedastic. Therefore, the assumption of either IID dividend growth or returns requires additional constraints on the variance-covariance dynamics.

A simple way to generate an IID dividend growth in the $WAR(1)$ dynamics is to specify matrices $M$ and $\mu$ to be both diagonal, where the nonzero elements in the first row are 1 and 0 for matrices $M$ and $\mu$, respectively. In this case, the volatility dynamics is single-factor and driven by the return volatility process $\Sigma_{22,t}$ alone. As shown in Gourieroux, Jasiak, and Sufana (2009), $\Sigma_{22,t}$ follows an autoregressive gamma process of order one, which is the discrete-time analog of Heston (1993)'s specification of stochastic volatility. Corollary 3.6 of Ang and Liu (2007) characterizes the present-value implications of IID dividend growth under a Heston (1993) specification of the return volatility. In this setting, expected returns are a nonlinear function of the volatility, and thus heteroskedastic, while the log price-dividend ratio is linear in the volatility. However, at the calibrated model parameters these nonlinearities are weak and a linear specification of log $pd$—ratios as a function of expected returns is accurate.

Our empirical evidence based on a $WAR(1)$ specification of multivariate time-varying risks speaks against the hypothesis of homoskedastic dividend growths or returns. Therefore, we study the joint dividend-return dynamics using a present-value model with multivariate time-varying dividend-return risks.

## 3 Data and Model Estimation

This section describes our data set and the estimation strategy based on a PML estimator with a Kalman filter.
3.1 Data

We obtain the with- and without-dividend monthly returns on the value-weighted portfolio of all NYSE, Amex and Nasdaq stocks from January 1946 until December 2015 from the Center for Research in Security Prices (CRSP). Based on these data, we construct annual series of aggregate dividends and prices, making use of 30-day T-bills to obtain annual series for cash-reinvested log dividend growth.\footnote{CRSP computes quarterly or annual return series under the stock market reinvestment assumption, but Koijen and Van Nieuwerburgh (2011) suggest that market reinvestment can be problematic because it imports some of the properties of returns to cash flows, and the resulting dividend growth series has thus a large volatility and low correlation with other measures of dividend growth.} Data on 30-day T-bill rates are also obtained from CRSP. In order to identify the latent time-varying risk components in our present-value model, we compute the time series of realized annual second moments of market returns and dividend growth, i.e., the squared returns and dividend growth, and their cross-product.

3.2 State space representation and estimation procedure

The redundancy of return shocks in equation (14) implies that the state dynamics of our present-value model are fully described by the dynamics of vector \((\Delta d_{t+1}, pd_{t+1}, \hat{\Sigma}_t, \hat{g}_t, \hat{\mu}_t)\), where \(\hat{\Sigma}_t := vech(\Sigma_t - \mu \Sigma)\) is the demeaned and half-vectorized variance-covariance state.

The hidden state variables in model (3)-(6) are the expected return and dividend growth \(\mu_t, g_t\) and the variance-covariance matrix \(\Sigma_t\). As in the model with constant risks, the observable variables in our setting include the dividend growth \(\Delta d_t\) and the price-dividend ratio \(pd_t\). To identify the latent variance-covariance state from a limited amount of data, we additionally include in the observable variables the realized joint second moments of returns and dividend growth. Indeed, while the market return \(r_{t+1}\) produces redundant information spanned by linear combinations of \(\Delta d_{t+1}\) and \(pd_{t+1}\), the second realized moments of returns and dividends are necessary to identify the time-varying risk structures summarized by hidden state \(\hat{\Sigma}_t\). This is a sharp difference of our setting relative to models with constant risks.

We estimate the conditional dynamics of first and second moments of returns and dividend growth in our present-value model using the following iterative procedure:
(1) Start with an estimation of a constant risk version of the model and obtain an initial estimate of the conditional first moments $g_t^{(0)}$ and $\mu_t^{(0)}$.

(2) Estimate the conditional second moment $\Sigma_t^{(0)}$ given $g_t^{(0)}$ and $\mu_t^{(0)}$, using as measurement the squared demeaned returns and dividends $\left(\Delta d_{t+1} - g_t^{(0)}\right)^2$, $\left(r_{t+1} - \mu_t^{(0)}\right)^2$ and their cross-product $\left(\Delta d_{t+1} - g_t^{(0)}\right) \left(r_{t+1} - \mu_t^{(0)}\right)$.

(3) Estimate the conditional first moments $g_t^{(1)}$ and $\mu_t^{(1)}$ given the filtered $\Sigma_t^{(0)}$ from the previous step, using as measurement the dividend growth and the price-dividend ratio.

(4) Iterate steps (2) and (3) until convergence.

Details on steps (1)–(3) in the above estimation procedure are provided in the following subsections.

### 3.3 Step (1): Constant risk model

As a benchmark and as a starting point for the estimation of our model, we consider a model with constant risks, i.e. with homoskedastic shocks, which is nested in our general specification with time-varying risks. This is naturally achieved by specifying expectation processes $\mu_t$ and $g_t$ that follow the dynamics (4) and (5) in a setting where the conditional covariance matrix of returns and dividends is constant ($\Sigma_t = \Sigma$) and the identification scheme (16)–(17) applies. In this case, $(g_{t+1}, \mu_{t+1})$ is a standard linear autoregressive process with constant risks, similar to those studied in Binsbergen and Koijen (2010) and Rytchkov (2012). In contrast to the identification choices in those papers, which impose a zero correlation between expected and realized dividends, our identification approach allows all joint second moments of discount rate and cash flow expectations, as well as their correlations with realized returns and dividend growth, to be different from zero. The variances and covariance of discount rate and cash flow expectation shocks follow from equations (18)–(20) under a constant covariance matrix $\Sigma$. Similarly, the covariances between expectation shocks and shocks in realizations follow from equations (22)–(23).\(^{11}\)

\(^{11}\)In standard present-value models with constant risks the variance-covariance matrix of shocks $(\tilde{\epsilon}_{d_{t+1}}, \tilde{\epsilon}_g_{t+1}, \tilde{\epsilon}_\mu_{t+1})' \Sigma$ is restricted, because only five out of six elements are identifiable. A simple iden-
The relevant hidden state variables in the model with constant risks are the expected return and dividend growth $\mu_t$ and $g_t$. The observable variables are the dividend growth $\Delta d_t$ and the price-dividend ratio $pd_t$. The model’s transition dynamics is:

$$
\hat{g}_{t+1} = \gamma_1 \hat{g}_t + \varepsilon^g_{t+1}, \tag{24}
$$
$$
\hat{\mu}_{t+1} = \delta_1 \hat{\mu}_t + \varepsilon^\mu_{t+1}. \tag{25}
$$

From Proposition 1, the model’s measurement equation for dividend growth is:

$$
\Delta d_{t+1} = \gamma_0 + \frac{1}{B_2} (pd_t - A + B_1 \hat{\mu}_t) + \tilde{\varepsilon}^D_{t+1}. \tag{26}
$$

From the model’s transition dynamics and Proposition 1, the measurement equation for the log price-dividend ratio is:

$$
pd_{t+1} = A - B_1 \hat{\mu}_{t+1} + B_2 \hat{g}_{t+1} = A - B_1 \delta_1 \hat{\mu}_t + B_2 \gamma_1 \hat{g}_t + \frac{1}{\rho} \left( \tilde{\varepsilon}^r_{t+1} - \tilde{\varepsilon}^D_{t+1} \right) = A(1 - \gamma_1) - B_1 (\delta_1 - \gamma_1) \hat{\mu}_t + \gamma_1 pd_t + \frac{1}{\rho} \left( \tilde{\varepsilon}^r_{t+1} - \tilde{\varepsilon}^D_{t+1} \right). \tag{27}
$$

Therefore, we can reduce the set of transition equations in the present-value model with constant risks to a single equation for $\hat{\mu}_t$.

The resulting state space model is linear and we apply a standard Kalman filter to obtain an exponential quadratic pseudo likelihood for estimating the model parameters $\Xi_0 := (\gamma_0, \delta_0, \gamma_1, \delta_1, \Sigma_{11}, \Sigma_{12}, \Sigma_{22}, p_1, p_2)$, together with the filtered time series of the expectation processes, using pseudo maximum likelihood. For stationarity, parameters $\delta_1$ and $\gamma_1$ are bounded to be less than one in absolute value. Overall, the constant risks version of our present-value model contains 9 parameters estimated by estimator $\hat{\Xi}_0$. Details on the estimation procedure are presented in section C.3 of the Supplemental Appendix.

3.4 Step (2): Estimation of $\Sigma_t$ given $g_t$ and $\mu_t$

The estimated constant risk model provides us with an initial estimate $g_t$ and $\mu_t$ of the relevant expectation processes, which we use to construct an initial time series of realized

\footnote{The correlation assumption sets the correlation between expected and realized dividend growth, $\rho_{gD}$, equal to zero; see e.g. Rytchkov (2012) and Binsbergen and Koijen (2010). We also have only five parameters driving all shock volatilities and correlations in the constant risk version of our model (parameters $\Sigma_{11}, \Sigma_{12}, \Sigma_{22}, p_1, p_2$), but we do not impose a priori any of the shocks to be uncorrelated.}
centred second moments for dividend growth and returns. From equation (3), the squared centred return and dividend growth give rise to the observation equation

\[
\begin{bmatrix}
\Delta d_{t+1} - g_t \\
r_{t+1} - \mu_t
\end{bmatrix} \begin{bmatrix}
\Delta d_{t+1} - g_t \\
r_{t+1} - \mu_t
\end{bmatrix} = \Sigma_t^{1/2} W_{t+1} \Sigma_t^{1/2} = \Sigma_t + \Sigma_t^{1/2} \epsilon_{t+1}^W \Sigma_t^{1/2},
\]

(28)

where

\[
W_{t+1} \equiv \begin{pmatrix}
\epsilon_D^{t+1} \\
\epsilon_r^{t+1}
\end{pmatrix} \begin{pmatrix}
\epsilon_D^{t+1} \\
\epsilon_r^{t+1}
\end{pmatrix}
\]

is IID Wishart distributed with 1 degree of freedom and mean \(I_2\), while \(\epsilon_{t+1}^W\) is IID centred Wishart distributed with 1 degree of freedom and shape parameter \(I_2\).

To write this observation equation in half-vectorized form, we introduce the notations

\[
Y_{t+1} \equiv \begin{pmatrix}
(\Delta d_{t+1} - g_t)^2 \\
(\Delta d_{t+1} - g_t)(r_{t+1} - \mu_t) \\
(r_{t+1} - \mu_t)^2
\end{pmatrix}',
\]

and

\[
\epsilon_Y^{t+1} \equiv vech(\Sigma_t^{1/2} \epsilon_{t+1}^W \Sigma_t^{1/2}) = L_2(\Sigma_t^{1/2} \otimes \Sigma_t^{1/2}) D_2 vech(\epsilon_{t+1}^W),
\]

where \(D_k\) and \(L_k\) are \(k\)-dimensional duplication and elimination matrices, respectively. In this way, we obtain a state space model with observation equation given by

\[
Y_{t+1} = vech(\mu^\Sigma) + \hat{\Sigma}_t + \epsilon_Y^{t+1},
\]

(29)

and transition dynamics given by:

\[
\hat{\Sigma}_{t+1} = S \hat{\Sigma}_t + \epsilon_{t+1}^\Sigma,
\]

(30)

with \(\epsilon_{t+1}^\Sigma := vech(\nu_{t+1})\) and an autoregressive matrix \(S\), defined explicitly in Appendix A.1, that only depends on matrix \(M\).

We estimate the latent covariance state \(\hat{\Sigma}_{t+1}\) and the parameters \(\Delta_0 := (M, k, V)\) using a Kalman filter and maximising a pseudo likelihood. For identification purposes and stationarity, some parameter constraints are necessary. \(M\) is assumed lower triangular, with positive diagonal elements smaller than one. \(V\) is assumed diagonal with positive components and \(k \geq 2\) is an integer. Details on the estimation procedure are presented in section C.1 of the Supplemental Appendix.

The result of this estimation step is a filtered time series of the conditional variance-covariance matrix \(\Sigma_t\) and an estimate of the 6-dimensional vector of parameters \(\hat{\Delta}_0\).
3.5 Step (3): Estimation of $g_t$ and $\mu_t$ given $\Sigma_t$

The third step in our estimation procedure estimates the conditional first moments $g_t$ and $\mu_t$ under a given time series of conditional second moments $\Sigma_t$ in step (2). The hidden state variables in this step are thus the expected return and dividend growth $\mu_t$ and $g_t$. The observable state variables are the dividend growth $\Delta d_t$ and the price-dividend ratio $pd_t$.

The transition dynamics and the measurement equations for dividend growth and the log price-dividend ratio are analogous to those of the model with constant risks in section 3.3, except that the conditional variance-covariance matrix of dividend growth and returns is treated as time-varying but observable, and set equal to the filtered $\Sigma_t$ from step (2). We therefore apply a standard Kalman filter to obtain an exponential quadratic pseudo likelihood and estimate the 6-dimensional vector of parameters $\Xi_1 := (\gamma_0, \delta_0, \gamma_1, \delta_1, p_1, p_2)$ with estimator $\hat{\Xi}_1$. Details of the estimation procedure are provided in section C.2 of the Supplemental Appendix.

Finally, estimated states $\mu_t$ and $g_t$ from step (3) are used to construct a new time series of realized second moments for step (2), and we iterate step (2) and step (3) until convergence of the filtered states and parameters from both steps. Empirically, we find that the convergence of our estimation procedure is relatively fast, with usually less than 10 iterations needed to attain convergence.

4 Model Implications

This section presents our empirical findings and discusses the main implications of time-varying risks for the joint dividend and return dynamics in presence of present-value constraints. In section B of the Supplemental Appendix, we also analyse the robustness of our empirical results to different choices of the cash-flow measure. There, we present estimation results using total payout (dividend plus repurchases) instead of cash dividends, showing that our conclusions are robust to the use of total payout measures.
4.1 Estimation results

Table 1, Panel A, presents the estimation results for our present-value model with time-varying risks. The unconditional expected log return is $\delta_0 = 8.3\%$, while the unconditional expected growth rate of dividends is $\gamma_0 = 5.5\%$. Expected returns feature an autoregressive root $\delta_1 = 0.897$, which is an indication of a persistent expectation process, having an half-life of about 6.8 years. Expected dividend growth is also persistent, but less than expected returns; its autoregressive root $\gamma_1 = 0.675$ implies a half-life of about 2.1 years. Compared with a model with constant risks, which implies a half-life of about 9 and 1 years for expected return and expected dividend growth, respectively, the model with time-varying risks implies a clearly lower heterogeneity in the persistence of dividend and return expectations. This feature has first-order implication for the variance decomposition of price-dividend ratios into the contributions of expected return and expected dividend shocks.

The estimated variance-covariance process implies persistent dividend and return volatilities, as well as persistent dividend-return covariances. To quantify the persistence of the multivariate dynamics $\Sigma_t$ and compare it to those of return and dividend expectations, we write it in vectorized form and compute the eigenvalues of the resulting VAR system, which are $\lambda_1 = 0.875$, $\lambda_2 = 0.662$ and $\lambda_3 = 0.501$, respectively. This implies dividend-return variance-covariance dynamics driven by three persistent state variables with half-lifes of 5.6, 2.1 and 1.4 years, respectively. The low estimated degrees of freedom parameter $k = 2$ indicates a slightly fat tailed distribution for the components of $\Sigma_t$.

Parameter $p_1$, which drives the effect of return shocks on expected dividend growth innovations is negative and small, while $p_2$ is large and positive. At the estimated parameters, shocks in expectations and realizations are related as follows (see again equations

---

12Parameter 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).

13The estimated root of expected returns and expected dividend growth in the model with constant risks, described in section 3.3, is $\delta_1 = 0.923$ and $\gamma_1 = 0.229$, respectively. Detailed estimation results are given in Table 2.
Therefore, positive shocks in realized dividends and returns negatively affect both expectation processes, consistent with a mean reversion effect in both returns and dividend growth.

On the contrary, for the constant risks model the estimated parameters imply no mean reversion in dividend growth and a weaker mean reversion in returns (see also the discussion in section 4.3):

\[
\begin{align*}
\varepsilon_{t+1}^g &= -0.1433\varepsilon_{t+1}^r - 0.825\varepsilon_{t+1}^D, \\
\varepsilon_{t+1}^\mu &= -0.1841\varepsilon_{t+1}^r - 0.1745\varepsilon_{t+1}^D.
\end{align*}
\]

Finally, the confidence intervals in Table 1 indicate a high statistical significance of basically all parameters characterizing the dynamics of expected returns and expected dividend growth, which is already evidence of a time variation of both dividend and return expectations. In contrast, the confidence intervals for the model with constant risks in Table 2 imply no statistical significance of parameters \(\gamma_1, p_1\) and \(p_2\), which generates a challenge for the interpretation of dividend predictability properties in this setting.

### 4.2 Dynamics of expected returns and dividend growth

We present in Figure 1 the time series of estimated expected return and expected dividend growth implied by our present-value model. In each panel, we also plot the filtered values of the expected cash flow growth and discount rate implied by a model with constant risks, as well as the actual value of these variables. We find that the expected return estimated by our present-value model and the one implied by the constant risk model are quite smooth and close to each other. This is natural, as the estimated autoregressive dynamics for expected returns is similar in both models. In contrast, larger differences arise for the estimated expected dividend. Indeed, the expected dividend growth implied by the time-varying risks model is clearly less volatile, with a standard deviation of 2.06% against 2.77% in the constant risks model.
In order to quantify the degree of predictability implied by the present-value model with time-varying risks, we compute the fraction of variability in $r_t$ and $\Delta d_t$ explained by $\mu_{t-1}$ and $g_{t-1}$, respectively, using the following sample $R^2$ goodness-of-fit measures:

$$R^2_{\text{Ret}} = 1 - \frac{\hat{\text{Var}}(r_{t+1} - \mu_t)}{\hat{\text{Var}}(r_{t+1})}; \quad R^2_{\text{Div}} = 1 - \frac{\hat{\text{Var}}(\Delta d_{t+1} - g_t)}{\hat{\text{Var}}(\Delta d_{t+1})},$$

where $\hat{\text{Var}}$ denotes sample variances and $\mu_t$, $g_t$, are the filtered expected return and expected dividend growth in the present-value model. The results in Table 3 show that $R^2_{\text{Ret}} = 8.11\%$ and $R^2_{\text{Div}} = 0.77\%$, i.e., expected returns seem to explain a relatively large fraction of actual returns, while the fraction of explained dividend growth variability is lower. Given the goodness-of-fit measures $R^2_{\text{Ret}} = 8.62\%$, $R^2_{\text{Div}} = 11.54\%$ and the weak statistical significance of a time-varying expected dividend growth in the model with constant risks, the dividend predictability features of this model are clearly less robust with respect to a specification with time-varying risks.

Interestingly, the predictability implications of a constant risk model under the standard identification assumption of uncorrelated dividends and expected dividends (third line in Table 3) are almost identical to those of the constant risk model nested in our time-varying risk specification. However, as we discuss below in more detail, the joint dynamics of realized and expected dividends under these two identification assumptions also feature important differences.

Finally, the evidence produced by the model with time-varying risks is also roughly more aligned to the one of standard predictive regressions of returns and dividend growth on price-dividend ratios, which imply $R^2_{\text{Ret}} = 8.29\%$ and $R^2_{\text{Div}} = 0.10\%$ (fourth line in Table 3). Let $I_t$ denote the econometrician’s information set at time $t$, generated by the history of dividends, returns and price-dividend ratios up to time $t$. Given the estimated parameters, the Kalman filter provides expressions to compute filtered estimates of the unknown latent states $\mu_{t-1}$ and $g_{t-1}$, conditional on $I_{t-1}$.

Let $\delta_1 = 0.930$ and $\gamma_1 = 0.252$, respectively. To derive the implications for the standard model with constant risks, we estimate the model in Binsbergen and Koijen (2010) for the case of cash-reinvested dividends, using data for the sample period 1946-2015. Our parameter estimates are very similar to theirs, which are based on the sample period 1946-2007.
Table 3), even though the model-implied dividend predictability is still higher.\footnote{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. Using the Kalman filter in section C of the Supplemental Appendix, we borrow from 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 and dividend growth rates and it features time-varying coefficients that depend on the whole history of the filtered conditional second moments of returns and dividends, see section E of the Supplemental Appendix.}

Despite the low R-squares for dividend growth, our parameter estimates and bootstrap confidence intervals suggest that both expected returns and expected dividend growth vary over time, e.g., because of the significant point estimates for autoregressive roots $\delta_1$ and $\gamma_1$ in Table 1, Panel A. The statistical significance of these results can be assessed by performing formal hypothesis tests. In our setting, predictability hypotheses can be formulated by means of simple parametric constraints, which can be efficiently tested with a standard likelihood ratio (LR) test, using the statistic

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

where $\Theta$ is the unrestricted parameter space, $\Theta_0$ the restricted set of parameters under the given null hypothesis $H_0$ and $\log L$ the log-likelihood of the model.

As $T \to \infty$, statistic $LR_T$ follows a $\chi^2_r$ distribution with $r$ degrees of freedom, where $r$ is the number of parameter constraints defining the constrained parameter set $\Theta_0$. However, given the limited available sample size, asymptotic theory may provide inaccurate approximation of the finite-sample distribution of the LR statistics. Therefore, we apply the nonparametric bootstrap likelihood ratio tests developed in Piatti and Trojani (2012).\footnote{Piatti and Trojani (2012) show that standard asymptotic tests of present-value models tend to over-reject the null of no predictability and propose nonparametric bootstrap tests with more reliable finite-sample properties. With slight modifications, we can apply their testing method to our framework.}

First, we test the hypothesis of constant return expectation:

$$H_0 : \delta_1 = 0 \quad \text{and} \quad p_1 = p_2 = 1.$$

\footnote{As $T \to \infty$, statistic $LR_T$ follows a $\chi^2_r$ distribution with $r$ degrees of freedom, where $r$ is the number of parameter constraints defining the constrained parameter set $\Theta_0$. However, given the limited available sample size, asymptotic theory may provide inaccurate approximation of the finite-sample distribution of the LR statistics. Therefore, we apply the nonparametric bootstrap likelihood ratio tests developed in Piatti and Trojani (2012).\footnote{Piatti and Trojani (2012) show that standard asymptotic tests of present-value models tend to over-reject the null of no predictability and propose nonparametric bootstrap tests with more reliable finite-sample properties. With slight modifications, we can apply their testing method to our framework.} 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. Using the Kalman filter in section C of the Supplemental Appendix, we borrow from 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 and dividend growth rates and it features time-varying coefficients that depend on the whole history of the filtered conditional second moments of returns and dividends, see section E of the Supplemental Appendix.}
Under this null, all price-dividend ratio variation is due to expected cash flow growth shocks. The value of the likelihood ratio statistic for this null is equal to $LR_T = 39.89$, which corresponds to a p-value of 4.1% in the bootstrap LR test. Therefore, null hypothesis (33) is rejected at a 5% confidence level. Second, we test for a constant expected dividend growth, using the null hypothesis:

$$H_0 : \gamma_1 = 0 \text{ and } p_1 = p_2 = 0 .$$ (34)

Under this null, all variation in the log price-dividend ratio comes from variation in expected returns. The value of the likelihood ratio statistic is now equal to $LR_T = 25.85$, which corresponds to a p-value of 12.5% of the bootstrap LR test. Therefore, null hypothesis (34) cannot be rejected at a 10% significance levels. Using the asymptotic $\chi^2_3$ distribution of the $LR$ statistics, we instead reject both null hypotheses (33) and (34) at a level of 1%.

In summary, the formal bootstrap test results indicate a stronger statistical evidence of return predictability than dividend predictability under time-varying dividend and return risks. As discussed in Piatti and Trojani (2012), the lack of statistical significance of null hypothesis (34) in the bootstrap likelihood ratio test does not necessarily have to be interpreted as evidence that the expected dividend growth is not time-varying. Instead, it may be interpreted as a low power of bootstrap tests of dividend predictability when using relatively short samples of data. Indeed, the estimated expected dividend growth in Figure 1 is clearly time-varying, even though less volatile than the estimated expected dividend growth in a constant risks model, and the simple null hypotheses $\gamma_1 = 0$ and $p_2 = 0$ are individually rejected according to the point estimates and bootstrap standard errors in Table 1, Panel A.\(^{18}\) Therefore, it is important to study the implications of the estimated unconstrained model with time-varying risks, which is consistent with a time variation of both return and dividend growth expectations.

\(^{18}\)As mentioned above, based on the estimation results in Table 2, Panel A, these simple null hypotheses cannot be rejected in the model with constant risks.
4.3 Key time-varying risks properties

The filtered dynamics of the conditional second moments of returns and dividends is presented in Figure 2.\(^{19}\) The conditional volatility of stock returns is large and persistent, with a peak of about 35% during the recent financial crisis. The volatility of dividend growth is smaller and less persistent, but still clearly time-varying, with values between around 3% and 13%. The estimated correlation between returns and dividend growth is usually negative and volatile, with values between -0.7 and 0.1.

The conditional variance-covariance of returns and dividends is linked to the conditional variances and covariances of all shocks in our model, which have the closed-form expressions provided in section 2.2. Using these expressions, Figure 3 reports the estimated dynamic correlations between expected and realized returns and dividend growth, respectively.

The correlation (22) between expected returns and returns is strongly negative on average, with a mean of about \(-0.92\), and it varies substantially over time, with a maximum of about \(-0.67\) and a minimum of about \(-0.99\). Such a time-varying comovement is a natural explanation for the variety of estimates of the correlation between returns and expected returns in the literature, which range approximately between \(-0.95\) and \(-0.5\), depending on the model and sample used.\(^{20}\) As noted in Pastor and Stambaugh (2009), among others, these correlations are a key source of return predictability, because they determine how past returns affect the forecast of future returns. Our estimated negative correlations between returns and expected returns are consistent with the idea that asset prices tend to fall when discount rates rise, as concluded by a many of studies. In parallel, their time variation implies a natural source of time-varying return predictability.

The model-implied expression (22) helps to rationalize this time variation in the context of our present-value model. At the estimated parameters, the correlation between

\(^{19}\)For ease of interpretation, Figure 2 shows the conditional volatilities of returns and dividend growth and their correlation, which in terms of the model state variables are given by \(\sqrt{\Sigma_{22}}\), \(\sqrt{\Sigma_{11}}\) and \(\frac{\Sigma_{12}}{\sqrt{\Sigma_{22}\Sigma_{11}}}\), respectively.

\(^{20}\)Campbell (1991), among others, uses a vector autoregression to estimate a correlation of about \(-0.9\) between return and expected return innovations for the 1952-1988 sample. Campbell and Ammer (1993) also find values around \(-0.9\) using the same sample. For the period 1927 to 1951, Campbell’s results imply values of the same correlation ranging from \(-0.67\) to \(-0.87\) across three different specifications.
returns and expected returns is decreasing in the conditional variance of returns, $\Sigma_{22}$, and the return-dividend covariance $\Sigma_{12}$. The peaks in $\text{corr}_t(\tilde{\epsilon}_{t+1}^H, \tilde{\epsilon}_{t+1}^r)$ in 1952, 1990 and 2006 correspond to years in which the return-dividend covariance $\Sigma_{12}$ was more negative. In contrast, the lowest correlation in 2010, towards the end of the financial crisis, corresponds to a period in which the conditional variance of returns, $\Sigma_{22}$, was at its maximum.

Under a countercyclical stock return variance $\Sigma_{22}$ and a procyclical return-dividend covariance $\Sigma_{12}$ the model tends to generate a countercyclical correlation $\text{corr}_t(\tilde{\epsilon}_{t+1}^H, \tilde{\epsilon}_{t+1}^r)$, which makes return predictability more apparent during economic crises or in periods of financial market distress. These findings are consistent with a concentration of return predictability during bad times and with equilibrium models that endogenously generate a countercyclical return predictability, such as Cujian and Hasler (2015).\footnote{Garcia (2013), among others, concludes that the content of news predicts future returns better during recessions.}

Correlation (23) between expected and realized dividend growth is also negative, with a mean of about $-0.87$, and quite time-varying, ranging from a minimum of $-0.97$ in 1952 to a maximum of $-0.54$ in 2010. At the estimated parameters, this correlation is decreasing in both the return-dividend covariance, $\Sigma_{12}$, and the variance of dividend growth, $\Sigma_{11}$. These findings are economically meaningful. For instance, Lettau and Wachter (2007) show that a negative correlation between realized and expected dividends plays a crucial role in explaining the value premium and the decreasing term structure of zero-coupon equity volatility, documented by Binsbergen, Brandt, and Koijen (2012).\footnote{Lettau and Wachter (2007) specify the stochastic discount factor so that shocks to aggregate dividends are priced. The negative correlation between expected and realized dividend growth leads to lower risk premia for growth stocks, since shocks to expected dividend growth act as a hedge. They calibrate this correlation to $-0.83$, using the consumption-dividend ratio as a proxy for expected cash flow growth, following Lettau and Ludvigson (2005).} Belo, Collin-Dufresne, and Goldstein (2014) also show evidence of a negative serial correlation of dividend growth, which is consistent with a negative correlation between expected and realized dividend growth. In summary, a negative correlation between expected and realized dividends appears as a parameter of first-order importance for a realistic specification of the term structure of dividend risks and the mean reversion properties of
dividend growth.

It is important to note that standard present-value models with constant risks hardly imply a negative covariance between expected and realized dividend growth. Indeed, the standard identification assumption in the literature actually sets this correlation equal to zero; see, e.g., Binsbergen and Koijen (2010) and Rytchkov (2012). Similarly, the constant risk model estimated under our more flexible identification assumptions (16)–(17) implies a large correlation between expected and realized dividends of about 0.74. Table 4 displays a direct comparison of the average estimated volatilities and correlations of all shocks in our present-value model with time-varying risks (Panel A), together with those in the estimated constant risks version of the model (Panel B), and those in the constant risks model under the standard identification assumption (Panel C).

We find that while the predictability implications of the two constant risk specifications under the filtered dynamics are almost identical (see again Table 3), the joint distributions of the innovations in the two models are extremely different and imply very different dividend growth expectation dynamics under the corresponding model probabilities. Figure 4 reports the time series of these conditional expectations, which are estimated from the smoothed states in our Kalman filter, and shows that the expected dividend growth under the constant risks model with the standard identification assumption virtually overlaps with the observed dividend growth.23 In other words, the standard identification $\rho_{gD} = 0$ in the literature implies a virtually deterministic dividend growth process. This implausible implication of this model is a direct consequence of the extremely low estimated volatility $\sigma_D$ reported in Table 4, which implies that virtually all of the unconditional volatility of dividends is generated by the predictable component $g$. A second economically implausible implication emerging from the findings in Panel C of Table 4 is that the correlations between realized dividend growth shocks and innovations in realized and expected returns are both estimated as quite large in absolute value.

The constant risks model nested by our time-varying risks specification solves some of the issues resulting from the standard identification with, e.g., more plausible volatilities of shocks to realized and expected dividends and less extreme correlations between

23As the main focus is here on model-implied expectations, we estimate them from filtered states conditional on information $I_t$ rather than information $I_{t-1}$; see again Footnote 14.
dividend shocks and shocks to realized or expected returns in Panel B of Table 4. These features imply more plausible estimated expected dividend growth dynamics in Figure 4. However, this model also implies large positive correlations between expected dividend shocks and shocks to realized dividends or returns, which are difficult to reconcile with the empirical evidence discussed above.

The average correlations implied by our model, in Panel A of Table 4, show a qualitatively similar decomposition between expected and unexpected changes in dividends and returns as in the constant risks model, but also a more realistic correlation between expected and realized dividend growth. Moreover, the fact that in our model these volatilities and correlations are naturally time-varying provides an additional flexibility in modelling the link between price-dividend ratios, returns and cash flows. The time-varying comovement between price-dividend ratios, returns and dividends directly follows from the model’s present-value constraints:

\[
\begin{align*}
\text{corr}_t(\varepsilon_{pd}^{t+1}, \tilde{r}^{t+1}) &= \frac{\Sigma_{22,t} - \Sigma_{12,t}}{\sqrt{\Sigma_{22,t} (\Sigma_{22,t} + \Sigma_{11,t} - 2\Sigma_{12,t})}}, \\
\text{corr}_t(\varepsilon_{pd}^{t+1}, \tilde{D}^{t+1}) &= \frac{\Sigma_{12,t} - \Sigma_{11,t}}{\sqrt{\Sigma_{11,t} (\Sigma_{22,t} + \Sigma_{11,t} - 2\Sigma_{12,t})}}.
\end{align*}
\]

At the filtered variance covariance states, the correlation between price-dividend ratios and returns is on average 0.95. The correlation between price-dividend ratios and dividend growth is on average of -0.55 and highly time-varying, ranging between -0.87 and -0.11. These time-varying variance covariance features naturally imply a time-varying degree of predictability in predictive regression of returns or dividends on lagged price-dividend ratios.

### 4.4 Dynamic price-dividend ratio decomposition

The predictability implications of our model can be intuitively understood also in terms of the conditional variance decomposition of price-dividend ratios. Under 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 \(-B_1 = -7.851\) (\(B_2 = 2.908\)). Such a large loading of expected returns implies that price-dividend ratio variations mostly reflect discount rate shocks, obfuscating the predictive power for
actual dividend growth. Since the expected return component is difficult to estimate from actual returns, due to a low signal-to-noise ratio, isolating it from aggregate price-dividend ratios in a model-free way is a difficult task. This creates the well-known EIV problem in standard predictive regressions, see also Binsbergen and Koijen (2009).

The conditional variance of the price-dividend ratio in our model reads explicitly:

$$\text{Var}_t(p_{d,t+1}) = B_1^2 \text{Var}_t(e_{t+1}^{\mu}) + B_2^2 \text{Var}_t(e_{t+1}^{g}) - 2B_1B_2 \text{Cov}_t(e_{t+1}^{\mu}, e_{t+1}^{g})$$  (37)

where $\text{Var}_t(e_{t+1}^{\mu})$, $\text{Var}_t(e_{t+1}^{g})$ and $\text{Cov}_t(e_{t+1}^{\mu}, e_{t+1}^{g})$ are affine functions of $\Sigma_t$ (see again equations (18)-(20)) that depend on parameters $p_1$ and $p_2$. Therefore, the fraction of conditional price-dividend ratio variation due to discount rate shocks, expectation shocks in dividend growth and the co-variation of dividend and return expectations is time-varying. Table 5 reports the average $pd$ variance decompositions for the estimated model, as well as their range of values. Figure 5 additionally shows the time series of this decomposition. While on average changes in expected returns have the largest effect on the price-dividend ratio, the fraction of variance explained is highly volatile. The effect of the covariance between expected returns and dividends in the decomposition is also negative and quite large.

4.5 Time-varying term structures of expectations

By iterating the Campbell-Shiller approximation (7) forward, we obtain:

$$p_{d,t} \approx \frac{\kappa}{1 - \rho} + \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} \gamma_{t+j}. \quad (38)$$

24Section D of the Supplemental Appendix provides analytic expressions for the model-implied asymptotic bias in standard predictive regression coefficients. These expressions allow us to evaluate the relative importance of EIV and small sample bias (Stambaugh (1999)), using simulations from the model. We find that, for the dividend regression, the EIV and small sample biases go in opposite directions and the EIV bias dominates. Therefore, the typical small sample bias correction (Stambaugh (1999)) produces even more biased point estimates in dividend growth predictive regressions.

25The dominating role of discount rate shocks emerges also from the variance decomposition in the model with constant risks (see Table 5). However, in this model the percentage variation due to the covariance term is positive and much smaller in absolute value. The conditional variance decomposition for the constant risks model follows from equation (37), but with constant conditional variances and covariance of expected returns and dividend growth.
As in Cochrane (2008a), we define the long-run return and dividend growth as the weighted sum of yearly returns and dividend growth based on the right-hand side of this expression. This definition ensures that the total long-horizon return and dividend growth exactly measure for $n \rightarrow \infty$ the $pd$–ratio variation induced by returns and dividend growth in the Campbell-Shiller identity. Moreover, it allows us to directly embed in Section 4.7 the term structures of long-horizon dividend-return expectations and risks of our model in an intertemporal CAPM setting, which decomposes stochastic discount factor shocks into news about cash flows, news about discount rates, and news about future dividend-return volatilities and correlations.\(^{26}\)

By applying recursively equations (3)-(5) and taking conditional expectations, we then obtain the annualized expected $n$-year expected return and dividend growth in closed form:

\[
\mu_t^{(n)} = \frac{1}{n}E_t \left[ \sum_{j=1}^{n} \rho^{j-1} r_{t+j} \right] = \frac{1 - \rho^n}{n(1 - \rho)} \delta_0 + \frac{1 - (\rho \delta_1)^n}{n(1 - \rho \delta_1)} \hat{\mu}_t, \tag{39}
\]

\[
g_t^{(n)} = \frac{1}{n}E_t \left[ \sum_{j=1}^{n} \rho^{j-1} \Delta d_{t+j} \right] = \frac{1 - \rho^n}{n(1 - \rho)} \gamma_0 + \frac{1 - (\rho \gamma_1)^n}{n(1 - \rho \gamma_1)} \hat{g}_t. \tag{40}
\]

The left panel of Figure 6 plots the time series of the term structure of return expectations $\mu_t^{(n)}$ in equation (39). We find that this term structure is quite volatile, mainly at short horizons. It stabilizes around a long-term expected market return of approximately 6% and it can be both downward or upward sloping, when annual expected returns are sufficiently large or low, respectively, which is consistent with the mean reversion properties of returns at the estimated parameters.

Interestingly, the term structure of market returns can behave quite differently during distinct crisis periods present in our sample. For instance, while the term structure of expected returns in Figure 7 is quite flat at the end of 1973 and 2011, after the first oil crisis and the recent financial crisis, it was clearly downward and upward sloping at

\(^{26}\)Cochrane (2008a)’s definition implies a higher sensitivity of long horizon expected returns and dividend growth to deviations of annual expected returns and dividend growth from their long term means. However, as we document in Section G of the Supplemental Appendix, the key properties of the term structures of expectations and risks in our estimated model are entirely analogous when using weighted or unweighted long-horizon returns.
the end of 1990 and 2000, in concomitance with the oil price shock and the burst of the
dot-com bubble, respectively.

The term structure of dividend growth expectations \( g_t^{(n)} \) in the left panel of Figure 6
is time-varying and usually downward sloping, as a consequence of the usually positive
annual expected dividend growth and the mean reversion of dividends under the estimated
model parameters. The term structure can be also increasing or hump-shaped, with a
peak at horizons between 5 and 10 years, when the annual expected dividend growth is
particularly low and future dividend growths are expected to mean revert. To illustrate,
note that short-term expected dividends sharply declined in the years following the recent
financial crisis. The term structure of expected dividend growth in those years was upward
sloping until an horizon of about 7 years, suggesting that dividends were expected to
grow faster in the medium run than in the short run. The hump in the term structure
approximately captures the expected recovery time after the crisis. At the end of the
sample period in 2015, about six years after the financial crisis, the term structure was
almost flat.

The dynamic properties of the term structure of dividend expectations in our model
are broadly consistent with those of the term structures of expectations obtained by more
direct approaches. Using a data set of dividend derivatives with maturities up to 10 years,
Binsbergen, Hueskes, Koijen, and Vrugt (2013) find that the slope of the term structure of
growth is countercyclical. Precisely, they show that the 5-year expected dividend growth
rate is higher (lower) than the 2-year expected growth rate during recessions (expansions).
Due to data availability, the sample in Binsbergen, Hueskes, Koijen, and Vrugt (2013)
starts in 2003 and includes only one crisis period. Our approach allows us to estimate the
term structure of dividend growth expectations for a longer sample that includes several
recessions and various crisis events.

The time series of the level and the slope of the term structure of dividend expectations
implied by our model are plotted in Figure 8. We find that while the slope of the term
structure of dividend expectations can increase during recessions, it can also behave
very differently during other crisis periods. For instance, in the last twenty years of our
sample the slope of the term structure was highest immediately after the two recessions
corresponding to the recent financial crisis and the burst of the Dot-Com bubble. These
large slopes were largely the consequence or sharply reduced (negative) 2-year expected dividends growths. However, we also find that the recessions following the oil crisis in the early 1970s and 1980s did not have a sizable impact on short-term expected dividend growth, giving rise to an almost flat term structure of dividend growth expectations and sharply reduced short-term expected returns. Return expectations sharply decreased in 2000 as well, leading to upward sloping term structures of long horizon expected returns, while the recent recession following the financial crisis is characterized by an almost flat term structure of expected returns, despite the term structure of dividend growth expectation is highly upward sloping in both cases.

This separation between cash-flow and discount rate crises is consistent with the evidence in, e.g., Campbell, Giglio, and Polk (2013), who discuss the difference between the 2000-2002 and 2007-2009 stock market downturns. We find that in the early 2000s stock price variations were led primarily by discount rate shocks. In contrast, in the late 2000s discount rates played only a minor role and stock price variations mainly reflected the worsened cash-flow prospects. This time varying importance of returns and dividend growth expectations in explaining price movements is evident also in Figure 9, where we report the time series of the time-varying fraction of the $pd\text{-ratio}$ variance driven by discount rates and cash flow growth in our model, obtained by allocating half of the effect of the covariance in equation (37) to expected returns and to dividend growth, respectively. We find that while discount rates always drive most of the variation in $pd$, the fraction of variance explained is highly time varying and cash flow growth news acquire a substantial role in several cases, including the recent financial crisis and the early 90s.

In summary, the time-varying risk features and the term structures of dividend and return expectations estimated by our model are helpful to understand the nature of stock market turmoils and the resulting optimal behaviour of long-term investors.

4.6 Time-varying term structures of risks

In our model, the conditional variance-covariance matrix of dividends and returns is time-varying and dependent on the horizon, giving rise to a term structure of dividend and
return risks, which we study in this section.\footnote{The properties of the term structure of return risk have been studied by several authors. Siegel (2008) reports that unconditional (sample) variances of returns 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 (2012) show that from the perspective of an investor subject to parameter uncertainty and imperfect predictors stocks can be more risky over longer horizons. Using Bayesian Model Averaging to account for model uncertainty, Diris (2011) finds that stocks are at least as risky in the long-run as in the short-run.}

4.6.1 Return risk

By applying recursively equations (3)-(5) and taking conditional variances, we obtain the closed-form annualized conditional variance of a $n$-year return:

$$\frac{1}{n} \text{Var}_t \left[ \sum_{j=1}^{n} \rho^{j-1} r_{t+j} \right] = \frac{1}{n} \sum_{j=1}^{n-1} \rho^{2j} \left( \frac{1 - (\rho \delta_1)^{n-j}}{1 - \rho \delta_1} \right)^2 \text{Var}_t(\varepsilon_{t+j}^\mu) + \frac{1}{n} \sum_{j=1}^{n-1} \rho^{2(j-1)} \text{Var}_t(\varepsilon_{t+j}^r) + \frac{2}{n} \sum_{j=1}^{n-1} \rho^{2j-1} \frac{1 - (\rho \delta_1)^{n-j}}{1 - \rho \delta_1} \text{Cov}_t(\varepsilon_{t+j}^\mu, \varepsilon_{t+j}^r), \quad (41)$$

where $\text{Var}_t(\varepsilon_{t+j}^\mu)$, $\text{Var}_t(\varepsilon_{t+j}^r)$ and $\text{Cov}_t(\varepsilon_{t+j}^\mu, \varepsilon_{t+j}^r)$ are affine functions of the variance-covariance state $\Sigma_t$, given explicitly in Appendix A.2. Consistently with the literature, the term structure of market risk measures the relation between the annualized volatility of cumulative weighted returns and the investment horizon.\footnote{Binsbergen, Brandt, and Koijen (2012) study the returns of zero-coupon dividend strips and their term structure of volatility. In contrast, we study the term structure of the annualized volatility of a claim to all future dividend cash flows.} Since shocks to expected returns and dividends are heteroskedastic and dynamically co-moving in our model, equation (41) implies a time-varying term structure, with dynamics summarized in the left panel of Figure 10. We find that the term structure of market risk is often downward sloping, but can also be hump-shaped with a peak at around 5 years maturity and increasing at longer horizons. These features are illustrated in more detail by Figure 11, where the term structures in years 2000 and 2010 are downward sloping, but the term structure is increasing in 1967 and hump-shaped in 1990.

A characteristic of our model is that the potentially increasing pattern of market risk between short and medium maturities completely arises from the time-varying uncertainty
of future expected returns, which is independent of the existence of imperfect predictors (Pastor and Stambaugh (2012)) or the specification of an explicit concern for model uncertainty (Diris (2011)). To understand why, we can split the conditional variance of returns in its three components in equation (41). The first term $\text{Var}_t(\epsilon_{t+j})$ reflects the uncertainty about future expected returns. The second term $\text{Var}_t(\tilde{\epsilon}_{t+j})$ captures the risk of future return shocks, while the last term $\text{Cov}_t(\epsilon_{t+j}, \tilde{\epsilon}_{t+j})$ reflects the mean reversion generated by the negative correlation of realized and expected returns.

Figure 12 presents the three components of the term structure of market risk in three different years, characterized by different levels of short term return volatility. Consistent with intuition, the return mean reversion tends to produce a decreasing term structure of market risk. Therefore, this component is linked to a strongly negative term structure effect, which is however partly offset by the impact of the other components. The uncertainty about future expected returns has a large positive and increasing effect, which, as highlighted by Pastor and Stambaugh (2012), is often underestimated or neglected. Its relative contribution is positively linked to the degree of predictability in returns, which is large for long horizons. Finally, the term structure effect of return shock risk is positive and can be increasing or decreasing with the horizon.

4.6.2 Dividend risk

The annualized conditional variance of $n$-year dividend growth is given by:

$$
\frac{1}{n} \text{Var}_t \left[ \sum_{j=1}^{n} \rho^{j-1} \Delta d_{t+j} \right] = \frac{1}{n} \sum_{j=1}^{n-1} \rho^{2j} \left( \frac{1 - (\rho \gamma_1)^{n-j}}{1 - \rho \gamma_1} \right)^2 \text{Var}_t(\epsilon_{t+j}) + \frac{1}{n} \sum_{j=1}^{n} \rho^{2(j-1)} \text{Var}_t(\tilde{\epsilon}_{t+j}) + 2 \frac{1}{n} \sum_{j=1}^{n-1} \rho^{2j-1} \frac{1 - (\rho \gamma_1)^{n-j}}{1 - \rho \gamma_1} \text{Cov}_t(\epsilon_{t+j}, \tilde{\epsilon}_{t+j}),
$$

where $\text{Var}_t(\epsilon_{t+j})$, $\text{Var}_t(\tilde{\epsilon}_{t+j})$ and $\text{Cov}_t(\epsilon_{t+j}, \tilde{\epsilon}_{t+j})$ are affine functions of the variance-covariance state $\Sigma_t$, given explicitly in Appendix A.2. The right panel of Figure 10 shows that this term structure is upward sloping, and sometimes U-shaped at short maturities. These features are illustrated in more detail by Figure 11, where the term structures in years 1967 and 2000 and 2010 are increasing, but the term structure in year 1990 is U-shaped.

Figure 13 plots the term structure of dividend risk and its components at three dif-
ferent points in time. The first component (blue line) reflects the uncertainty $\text{Var}_t(\varepsilon_{t+j}^g)$ about future expected dividend growth. It is positive and increasing with the horizon. The second component (red line) is related to the risk $\text{Var}_t(\varepsilon_{t+j}^D)$ of future dividend growth shocks. This component can be decreasing or hump shaped but it is usually relatively small. The third component, which reflects the covariance $\text{Cov}_t(\varepsilon_{t+j}^g, \varepsilon_{t+j}^D)$ between realized and expected dividend growth shocks, is negative and strongly decreasing. The shape of the overall term structure of dividend risk is therefore mainly characterised by the opposing effects of the mean reversion and the uncertainty about future expected dividend growth. For short horizons, up to about five years, the mean reversion effect tends to dominate, generating the U-shape in the term structure.

### 4.7 An intertemporal CAPM

A recent important literature decomposes stochastic discount factor (SDF) innovations into heteroskedastic cash flow and discount rate shocks, in order to understand the importance of volatility fluctuations for asset prices and the macroeconomy.\(^{29}\) This literature relies on various single-factor specifications of stochastic volatility in the log SDF of a representative agent with recursive preferences:

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi} \Delta d_{t+1} + (\theta - 1) r_{t+1} ,$$

where $\theta = (1-\gamma)/(1-1/\psi)$, $\delta$ is a subjective discount factor, $\gamma$ is a risk aversion coefficient and $\psi$ the intertemporal elasticity of substitution. In this specification, $\Delta d_{t+1}$ and $r_{t+1}$ are a potentially heteroskedastic log consumption growth and log return on wealth, which can give rise to a heteroskedastic SDF. Under joint normality of $m_{t+1}$ and $r_{t+1}$, a standard log linearization yields the following SDF decomposition in terms of news to cash flows, future discount rates and future risks:

$$N_{m,t+1} \equiv m_{t+1} - E_t[m_{t+1}] = -\gamma N_{CF,t+1} + N_{DR,t+1} + \frac{1}{2} N_{V,t+1} ,$$

where

$$V_t \equiv \text{Var}_t(m_{t+1} + r_{t+1}) = \text{Var}_t(\theta r_{t+1} - (\theta/\psi) \Delta d_{t+1}) ,$$

\(^{29}\)See, e.g., Campbell, Giglio, Polk, and Turley (2017) and Bansal, Kiku, Shaliastovich, and Yaron (2014).
and

\[ N_{CF,t+1} \equiv (E_{t+1} - E_t) \left( \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1} \right), \quad (46) \]
\[ N_{DR,t+1} \equiv (E_{t+1} - E_t) \left( \sum_{j=1}^{\infty} \rho^j r_{t+j+1} \right), \quad (47) \]
\[ N_{V,t+1} \equiv (E_{t+1} - E_t) \left( \sum_{j=1}^{\infty} \rho^j V_{t+j} \right), \quad (48) \]

with \( \rho \) the log-linearization parameter related to the average consumption-wealth ratio.

The SDF specification is completed by specifying a functional form for risk \( V_t \). For instance, Campbell, Giglio, Polk, and Turley (2017) assume proportionality between \( V_t \) and the variance of market returns, while Bansal, Kiku, Shaliastovich, and Yaron (2014) specify \( V_t \) as a linear function of the variance of consumption growth.

Note that in order to ensure consistency of the distribution of returns and consumption growth with the underlying present-value constraints, the specification of \( V_t \) cannot in general be disconnected from the specification of the news components \( N_{CF,t+1} \) and \( N_{DR,t+1} \). Our framework in Section 2 supports an approach with multivariate risks, under a present-value model that imposes the Campbell and Shiller (1988) identity (7) on the joint dynamics of \( (\Delta d_{t+1}, r_{t+1})' \) and \( V_t \). Indeed, if variables \( (\Delta d_{t+1}, r_{t+1})' \) and \( \Sigma_t \) follow a state dynamics of the form given in Section 2, we have:

\[ V_t = \theta^2 \left( \Sigma_{22,t} + \frac{1}{\psi^2} \Sigma_{11,t} - 2 \frac{1}{\psi} \Sigma_{12,t} \right), \quad (49) \]

together with following closed-form expression for cash flow and discount rate news, which are directly inferred from formulas (39)–(40) for the term structures of long-horizon dividend growth and returns:

\[ N_{CF,t+1} = \Delta d_{t+1} - \gamma_0 + \rho \hat{g}_{t+1} - \hat{g}_t \left( \frac{1}{1 - \rho \gamma_1} \right); \quad N_{DR,t+1} = \rho \frac{\hat{\mu}_{t+1} - \delta_1 \hat{\mu}_t}{1 - \rho \delta_1}. \quad (50) \]

Therefore, in our model \( N_{CF,t+1} \) and \( N_{DR,t+1} \) are heteroskedastic, stochastically correlated and stochastically co-moving with consumption growth and returns. Risk news are also available in closed-form, from formulas (41)-(42) for the term structures of long-horizon dividend and return risks and from the closed-form affine conditional moments in the
WAR(1) process:

\[
N_{V,t+1} = \theta^2 \sum_{j=1}^{\infty} \rho^j (E_{t+1} - E_t) \left( \Sigma_{22,t+j} + \frac{1}{\psi^2} \Sigma_{11,t+j} - 2 \frac{1}{\psi} \Sigma_{12,t+j} \right)
\]

\[
= \theta^2 C_\psi \sum_{j=1}^{\infty} \rho^j (E_{t+1} - E_t) \Sigma_{t+j},
\]  

(51)

where \(C_\psi := [1/\psi^2 - 2/\psi 1]\). From the explicit expression for \(E_t(\Sigma_{t+j})\) in Appendix A.2, we obtain:

\[
(E_{t+1} - E_t) \Sigma_{t+j} = vech \left[ M^{j-1} (\Sigma_{t+1} - M \Sigma_t M' - kV)(M^{j-1})' \right]
\]

\[
= L_2 (M \otimes M)^{j-1} D_2 \Sigma_{t+1}. 
\]  

(52)

Therefore:

\[
N_{V,t+1} = \theta^2 C_\psi \sum_{j=1}^{\infty} \rho^j L_2 (M \otimes M)^{j-1} D_2 \Sigma_{t+1}
\]

\[
= \theta^2 \rho C_\psi L_2 (I_4 - \rho (M \otimes M))^{-1} D_2 \Sigma_{t+1}. 
\]  

(53)

It follows that \(N_{V,t+1}\) is also heteroskedastic and given by a simple affine function of the stochastic variance-covariance matrix \(\Sigma_t\) of consumption growth and returns. Importantly, news to risk depend on both news about the future volatility of dividends or returns and news about their future covariance. In summary, we obtain a tractable intertemporal CAPM with priced multivariate heteroskedasticity, which is consistent with the present-value constraints between consumption growth, the return on wealth and the wealth-consumption ratio.

To illustrate more concretely the SDF news decomposition in our model, we report in Figure 14 the time series of filtered news components under the estimated model parameters of Section 4.1, where we proxy for simplicity consumption growth by dividend growth and the return on wealth by the market return. The top panel displays the news to cash flows \(N_{CF,t+1}\) and discount rates \(N_{DR,t+1}\), while the bottom panel shows the time series of news to the different variance components, i.e.,

\[
N_{V_{11},t+1} \equiv \rho e_1' L_2 (I_4 - \rho (M \otimes M))^{-1} D_2 \Sigma_{t+1},
\]

This solution requires the condition \(\max |eig(\rho (M \otimes M))| < 1\). This condition is satisfied, because \(|\rho| < 1\) and because stationarity of the WAR(1) process requires \(M\) to have eigenvalues (its diagonal elements, since \(M\) is assumed to be triangular) less than 1 in absolute value.
\[
N_{V_{12},t+1} \equiv \rho e'_2 L_2 (I_4 - \rho (M \otimes M))^{-1} D_2 \varepsilon^\Sigma_{t+1},
\]
\[
N_{V_{22},t+1} \equiv \rho e'_3 L_2 (I_4 - \rho (M \otimes M))^{-1} D_2 \varepsilon^\Sigma_{t+1},
\]

where \(e_j\) is the \(j\)-th unit vector in \(\mathbb{R}^3\). In general, we see that while news to cash flows, discount rates and market volatility exhibit similar average orders of magnitude, cash-flow volatility news and dividend-return covariance news are clearly smaller.

While the time series of news components in Figure 14 are independent on the preference parameters in the SDF, the decomposition of SDF news \(N_{m,t+1}\) and total risk news \(N_{V,t+1}\) depend on the risk aversion \(\gamma\) and the intertemporal elasticity of substitution \(\psi\). For illustration purposes, we set a risk aversion \(\gamma = 2\) and two possible values of \(\psi\), above and below one. The top panels of Figure 15 display the decomposition of SDF shocks into cash flows, discount rates and total risk news, computed using the estimated parameters and filtered states. The bottom panels show the decomposition of \(N_{V,t+1}\) into the news components directly related to \(\Sigma_{11}\), \(\Sigma_{12}\) and \(\Sigma_{22}\). The intertemporal elasticity of substitution is \(\psi = 1.5\) in the left panels and \(\psi = 0.5\) in the right panels.

Under the given parameter choice, the SDF news decomposition for elasticities of substitution above one is dominated by total risk news \(N_{V,t+1}\), which is a consequence of the larger effects of news in dividend-return volatilities and covariances on \(N_{V,t+1}\) as \(\theta\) increases. In parallel, the decomposition of \(N_{V,t+1}\) into news about dividend-return variances and covariances is dominated by news to market volatility. For a lower elasticity parameter \(\psi = 0.5\), we obtain both a lower contribution of total risk news to SDF shocks and a slightly lower weight of market volatility news in the decomposition of total risk news. These effects are again a direct consequence of the closed-form dependence of \(N_{V,t+1}\) on parameter \(\psi\).

In summary, various parameter choices for the elasticity of substitution give rise to structurally different compositions of total risk news. In parallel, they also strongly influence the contribution of total risk news to SDF shocks. For standard choices of risk aversion and elasticities of substitution above one, our estimated present-value model with time-varying risks yields a preponderant role of market volatility news for understanding SDF shocks.
5 Conclusion

We characterize the dynamics of dividend growth and returns using a tractable present-value model with multivariate time-varying risks, which is estimated with a latent variables approach that treats dividend-return expectations and risks as unobservable. Our estimation based on postwar US stock market data shows that the joint dividend-return dynamics under time-varying risks are different from those under constant risks along several key dimensions.

Expected returns have roughly similar persistence and return predictability properties, but the expected dividend process under heteroskedasticity is clearly more persistent and explains only a small fraction of future dividends. Through the model’s present-value constraints, the heteroskedastic dividend growth and return have a stochastic negative correlation with the expected dividend growth and the expected return, respectively. Therefore, dividend growth and return processes in our model are stochastically mean reverting and linked to a time-varying degree of dividend and return predictability. In contrast, the homoskedastic present-value setting yields static persistence and predictability properties, coupled with positively correlated expected and realized dividend growths.

Another distinguishing feature of the dividend-return dynamics in our model is that the variance decomposition of the price-dividend ratio is time-varying. We find that while on average the price-dividend ratio variation is dominated by shocks to expected returns, the shocks to expected dividend growth also have a time-varying first-order contribution to the price-dividend ratio variation. These features are helpful to systematically differentiate periods of financial market turmoil primarily associated with discount rate or cash-flow shocks.

The heteroskedastic multivariate dividend-return dynamics in our model gives rise to economically plausible time-varying term structures of dividend-return expectations and risks. We find that the term structure of expected returns stabilizes around a long term expected return of about 6%. It can be both upward and downward sloping and it can behave quite differently during distinct crisis periods present in our sample. The slope of the term structure of expected dividend growth in the last part of our sample tends to increase during recessions, consistent with the evidence in Binsbergen, Hueskes,
Koijen, and Vrugt (2013). However, the term structure can also be virtually flat during various other recessions and crisis periods in the earlier part of our sample. Finally, the term structure of return volatility is downward sloping on average, but it can also be upward sloping in states of large uncertainty about future expected returns, consistently with the intuition developed in Pastor and Stambaugh (2009). These rich dynamics are a natural consequence of the interplay between the stochastic mean reversion of returns in our model and the time-varying uncertainties of return and expected return shocks.

The joint dividend-return dynamics estimated in our model can provide useful guidelines for the specification of preference-based asset pricing models with heteroskedastic and stochastically correlated cash flows and discount rates. Incorporating such dynamic co-movement features into preference-based macro asset pricing models with time-varying risks, such as Bansal, Kiku, Shaliastovich, and Yaron (2014) and Campbell, Giglio, Polk, and Turley (2017), is an interesting avenue for future research. We explore the implications of multivariate time-varying risks in such settings, by means of an Intertemporal CAPM supported by the present-value constraints of our present-value model with time-varying risks. In this setting, we show how various periods of financial distress can be interpreted in terms of the impact of various news corresponding to each element of the future variance-covariance matrix of cash flows and returns.

Finally, our modelling approach can be extended to explicitly incorporate tractable specifications of multivariate leverage effects, based on the continuous-time Wishart dynamics introduced in Buraschi, Porchia, and Trojani (2010). In this setting, closed-form characterizations of the joint dynamics of dividends and returns, following Ang and Liu (2007)’s methodology, may produce additional insights into the implications of present-value relations under time-varying risks.
References


A Present-value model

A.1 Main notation

The state variables of the model are:

\[ \hat{\mu}_t = \mu_t - \delta_0, \]
\[ \hat{g}_t = g_t - \gamma_0, \]
\[ \hat{\Sigma}_t = vech(\Sigma_t - \mu^\Sigma), \]

where \( \mu^\Sigma \) is the unconditional mean of stationary variance-covariance process \( \Sigma_t \), which is such that

\[ vech(\mu^\Sigma) = [I_3 - L_2(M \otimes M)D_2]^{-1}kL_2vec(V), \]

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 \( 2 \times 2 \) matrix \( A \):

\[ D_2vec(A) = vec(A), \quad L_2vec(A) = vech(A), \]

where \( vec \) denotes vectorization and \( vech \) half-vectorization.

The dynamics of the state variables are obtained from (4)-(6) as follows:

\[ \hat{g}_{t+1} = \gamma_0 \hat{g}_t + \varepsilon^g_{t+1}, \]
\[ \hat{\mu}_{t+1} = \delta_0 \hat{\mu}_t + \varepsilon^\mu_{t+1}, \]
\[ \hat{\Sigma}_{t+1} = S\hat{\Sigma}_t + \varepsilon^\Sigma_{t+1}, \]

where \( S = L_2(M \otimes M)D_2 \).

In terms of these demeaned states, the dynamics of realized returns and dividend growth in equation (3) is the following:

\[ \Delta d_{t+1} = \gamma_0 + \hat{g}_t + \varepsilon^D_{t+1}, \]
\[ r_{t+1} = \delta_0 + \hat{\mu}_t + \varepsilon^r_{t+1}, \]

where

\[ \varepsilon^D_{t+1} = e'_1\Sigma^{-1/2}_{t} \begin{pmatrix} \varepsilon^D_{t+1} \\ \varepsilon^r_{t+1} \end{pmatrix}, \]
and

\[
\tilde{\varepsilon}_{t+1} = \varepsilon_{t+1}^{1/2} \left( \begin{array}{l} \varepsilon_{t+1}^D \\ \varepsilon_{t+1}^r \end{array} \right).
\]

Since \( \varepsilon_{t+1}^g \) and \( \varepsilon_{t+1}^\mu \) are linear combinations of the other shocks (see equations (16) and (17)), to complete the specification of the model we only need to specify the conditional covariance matrix of

\[
\begin{pmatrix}
\varepsilon_{t+1}^D \\
\varepsilon_{t+1}^r \\
\varepsilon_{t+1}^\Sigma
\end{pmatrix},
\]

which is given by:

\[
Q_t = \begin{bmatrix}
\Sigma_t & 0_{2\times 3} \\
0_{3\times 2} & Var_t(\varepsilon_{t+1}^\Sigma)
\end{bmatrix},
\]

(54)

where \( Var_t(\varepsilon_{t+1}^\Sigma) \) is given by:

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

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

### A.2 Term structure of conditional variances

The conditional variance of model-implied \( n \)-year returns and dividend growth are the following:

\[
\begin{align*}
Var_t \left[ \sum_{j=1}^{n} \rho^{j-1} r_{t+j} \right] &= \sum_{j=1}^{n-1} \rho^{2j} \left( \frac{1 - (\rho \delta_1)^{n-j}}{1 - \rho \delta_1} \right)^2 Var_t(\varepsilon_{t+j}^\mu) + \sum_{j=1}^{n} \rho^{2(j-1)} Var_t(\varepsilon_{t+j}^r)
+ 2 \sum_{j=1}^{n-1} \rho^{2j-1} \frac{1 - (\rho \delta_1)^{n-j}}{1 - \rho \delta_1} \text{Cov}_t(\varepsilon_{t+j}^\mu, \varepsilon_{t+j}^r), \\
Var_t \left[ \sum_{j=1}^{n} \rho^{j-1} \Delta d_{t+j} \right] &= \sum_{j=1}^{n-1} \rho^{2j} \left( \frac{1 - (\rho \gamma_1)^{n-j}}{1 - \rho \gamma_1} \right)^2 Var_t(\varepsilon_{t+j}^g) + \sum_{j=1}^{n} \rho^{2(j-1)} Var_t(\varepsilon_{t+j}^D)
+ 2 \sum_{j=1}^{n-1} \rho^{2j-1} \frac{1 - (\rho \gamma_1)^{n-j}}{1 - \rho \gamma_1} \text{Cov}_t(\varepsilon_{t+j}^g, \varepsilon_{t+j}^D),
\end{align*}
\]

where

\[
\begin{align*}
Var_t(\varepsilon_{t+j}^\mu) &= \frac{1}{\rho^2 B_t^2} \left[ (p_2 - 1)^2 - 2(p_1 - 1)(p_2 - 1) (p_1 - 1)^2 \right] \text{vech} E_t(\Sigma_{t+j-1}^1), \\
Var_t(\varepsilon_{t+j}^r) &= [0 \ 0 \ 1] \text{vech} E_t(\Sigma_{t+j-1}^1),
\end{align*}
\]
\[
\text{Cov}_t(\varepsilon_{t+j}^n, \tilde{\varepsilon}_{t+j}^r) = \frac{1}{\rho B_1} [0 \ 1 - p_2 \ p_1 - 1] \text{vech} E_t(\Sigma_{t+j-1}),
\]
\[
\text{Var}_t(\varepsilon_{t+j}^g) = \frac{1}{\rho^2 B_2} \left[ p_2^2 - 2p_1 p_2 \ p_1^2 \right] \text{vech} E_t(\Sigma_{t+j-1}),
\]
\[
\text{Var}_t(\tilde{\varepsilon}_{t+j}^D) = [1 \ 0 \ 0] \text{vech} E_t(\Sigma_{t+j-1}),
\]
\[
\text{Cov}_t(\varepsilon_{t+j}^g, \tilde{\varepsilon}_{t+j}^D) = \frac{1}{\rho B_2} [-p_2 \ p_1 \ 0] \text{vech} E_t(\Sigma_{t+j-1}),
\]

and
\[
E_t(\Sigma_{t+j}) = M^j \Sigma_t (M^j)' + kV(j),
\]
\[
V(j) = V + VM' + \ldots + M^{j-1}V(M^{j-1})',
\]

Note that non-contemporaneous correlations between return and expected return shocks are equal to zero and that the conditional variance of long-run returns is an affine function of the variance-covariance state \(\Sigma_t\).
B Tables and Figures

Table 1: Estimation results for our present-value model. The model is estimated in two steps, by quasi maximum-likelihood using yearly data from 1946 to 2015. Panel A presents estimates of the coefficients of the underlying processes, with bootstrap standard errors in parenthesis. Panel B reports resulting coefficients of the present-value model in equation (8).

<table>
<thead>
<tr>
<th>Panel A: Quasi maximum-likelihood estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\gamma_0)</td>
</tr>
<tr>
<td>0.055</td>
</tr>
<tr>
<td>(0.023)</td>
</tr>
<tr>
<td>(M_{11})</td>
</tr>
<tr>
<td>0.708</td>
</tr>
<tr>
<td>(0.208)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Implied present-value parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rho)</td>
</tr>
<tr>
<td>0.972</td>
</tr>
</tbody>
</table>
Table 2: Estimation results for the constant risks model nested in our time-varying risks specification.

<table>
<thead>
<tr>
<th>Panel A: Quasi maximum-likelihood estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
</tr>
<tr>
<td>0.061</td>
</tr>
<tr>
<td>(0.024)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Sigma_{21}$</th>
<th>$\Sigma_{22}$</th>
<th>$p_1$</th>
<th>$p_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0009</td>
<td>0.0207</td>
<td>0.142</td>
<td>-0.321</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.007)</td>
<td>(0.696)</td>
<td>(2.232)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Implied present-value parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
</tr>
<tr>
<td>0.9714</td>
</tr>
</tbody>
</table>

Table 3: Sample R-squared values of returns and dividend growth, computed using equation (31), for our time-varying risks model (first row), the nested constant risks model (second row) and the standard constant risk model (third row). The fourth row gives results for a standard OLS predictive regression of observed returns and dividend growth on price-dividend ratio. All models are estimated using yearly data from 1946 to 2015.

<table>
<thead>
<tr>
<th>R-squared values (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{\text{Ret}}^2$</td>
</tr>
<tr>
<td>Time-varying risks model</td>
</tr>
<tr>
<td>Nested constant risks model</td>
</tr>
<tr>
<td>Standard constant risks model</td>
</tr>
<tr>
<td>OLS</td>
</tr>
</tbody>
</table>
Table 4: Average estimated volatilities and correlations of the shocks in our present-value model with time varying risks (Panel A), in the estimated constant risks version of the model (Panel B), and in the standard constant risks model (Panel C).

<table>
<thead>
<tr>
<th></th>
<th>Panel A: Time-varying risks model</th>
<th>Panel B: Nested constant risks model</th>
<th>Panel C: Standard constant risks model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \tilde{\varepsilon}_D )</td>
<td>( \tilde{\varepsilon}_r )</td>
<td>( \tilde{\varepsilon}_g )</td>
</tr>
<tr>
<td>( \tilde{\varepsilon}_D )</td>
<td>0.0551</td>
<td>-0.2735</td>
<td>-0.8749</td>
</tr>
<tr>
<td>( \tilde{\varepsilon}_r )</td>
<td>0.1511</td>
<td>-0.1907</td>
<td>-0.9234</td>
</tr>
<tr>
<td>( \tilde{\varepsilon}_g )</td>
<td>0.0443</td>
<td>0.5299</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\varepsilon}_\mu )</td>
<td>0.0264</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Conditional variance decomposition of the price-dividend ratio for our time-varying risks model and in the constant risks model.

<table>
<thead>
<tr>
<th></th>
<th>Discount Rates</th>
<th>Div.Growth</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-Varying Risks</td>
<td>138.96%</td>
<td>59.20%</td>
<td>-98.16%</td>
</tr>
<tr>
<td>([\text{min} \quad \text{max}])</td>
<td>[46.32% 211.62%]</td>
<td>[13.06% 136.24%]</td>
<td>[-186.30% -5.01%]</td>
</tr>
<tr>
<td>Constant Risks</td>
<td>88.86%</td>
<td>4.04%</td>
<td>7.10%</td>
</tr>
</tbody>
</table>
Figure 1: Expected vs Realized yearly returns and dividend growth. These graphs show the model-implied (filtered) series (red lines) of expected returns $\mu_t$ (first panel) and expected dividend growth $g_t$ (second panel), as well as the realized (blue lines) return, $r_{t+1}$ and log dividend growth, $\Delta d_{t+1}$, respectively. The two panels also show the expectations implied by a version of the model with constant risks. Shaded areas correspond to NBER recessions.
Figure 2: Filtered standard deviations of yearly dividend growth, returns, and their correlation. Shaded areas correspond to NBER recessions.
Figure 3: Conditional correlation between shocks in expected and unexpected returns (upper panel), $corr_t(\varepsilon_{t+1}^{\mu}, \tilde{\varepsilon}_{t+1}^{\tau})$, and conditional correlation between shocks in expected and unexpected dividend growth (lower panel), $corr_t(\varepsilon_{t+1}^{g}, \tilde{\varepsilon}_{t+1}^{D})$. Shaded areas correspond to NBER recessions.
Figure 4: Expected vs Realized yearly returns and dividend growth, using the *a posteriori* estimate of the latent expectation states in the Kalman filter, i.e. $\mu_{t-1}$ and $g_{t-1}$ given the information at time $t$, $I_t$. These graphs show the model-implied series (red lines) of expected returns $\mu_t$ (first panel) and expected dividend growth $g_t$ (second panel), as well as the realized (blue lines) return, $r_{t+1}$ and log dividend growth, $\Delta d_{t+1}$, respectively. The two panels also show the expectations implied by the constant risks model nested in our time-varying risks specification (yellow lines) and by a standard constant risks model (purple lines).
Figure 5: Time series of log price-dividend ratio variance decomposition. The blue line denotes the discount rate component, the red line shows the effect of cash flow shocks, while the magenta line is the effect of the covariance between expected returns and cash flows. Shaded areas correspond to NBER recessions.

Figure 6: Dynamics of the term structure of the conditional per-period expected long-horizon return ($\mu_t^{(n)}$, left panel) and dividend growth ($g_t^{(n)}$, right panel), from equations (39) and (40), respectively, computed using estimated parameters and filtered state. We consider horizons of 1 to 20 years.
Figure 7: Term structure of the conditional per-period expected long-horizon return (left axes, $\mu_t^{(n)}$) and dividend growth (right axes, $g_t^{(n)}$), from equations (39) and (40), in years 1973, 1990, 2000 and 2011.
Figure 8: Level (upper panel) and slope (lower panel) of the term structure of dividend growth expectation. The level is measured as the 2-year expected growth rate, standardized, while the slope is measured by the difference between the 5- and 2-year expected dividend growth, also standardized. Expected growth rates are computed using estimated parameters and filtered state, for our time-varying risks model (blue line) and the model with constant risks (red line). Shaded areas corresponds to NBER recessions.
Figure 9: Time series of the fraction of the variation in \( pd \) ratio driven by discount rates and cash flow growth variation.

Figure 10: Dynamics of the term structure of the conditional per-period long-horizon return (left panel) and dividend growth (right panel) volatility, from equations (41) and (42), respectively, computed using estimated parameters and filtered state. We consider horizons of 1 to 20 years.
Figure 11: Term structure of the conditional per-period volatilities of long-horizon return (left axes) and dividend growth (right axes), from equations (41) and (42), in years 1967, 1990, 2000 and 2010.
**Figure 12:** 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 1 to 20 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 13:** Decomposition of the term structure of the conditional per-period variance of long-horizon dividend growth, computed using estimated parameters and filtered states. The blue line denotes the component of the variance that is due to uncertainty about future expected cash flow growth, the red line denotes the component due to future dividend growth shocks, while the green line denotes the mean reversion component. The black dashed line denotes the total conditional variance, for horizons of 1 to 20 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 14: The top panel displays the news to cash flows ($N_{CF,t+1}$) and discount rates ($N_{DR,t+1}$) implied by the intertemporal CAPM in Section 4.7, computed using estimated parameters and filtered states (see equation (50)). The bottom panel shows the news to the different variance components ($N_{V_{11},t+1}$, $N_{V_{12},t+1}$ and $N_{V_{22},t+1}$). Shaded areas corresponds to NBER recessions.
Figure 15: The top panels display the decomposition of SDF shocks (blue lines) into cash flows (red lines), discount rates (green dashed lines) and variance (magenta dash dotted lines) news implied by the intertemporal CAPM in Section 4.7, computed using estimated parameters and filtered states. The bottom panels show the decomposition of the total news to the SDF variance into the components due to $\Sigma_{11}$, $\Sigma_{12}$ and $\Sigma_{22}$, respectively. The risk aversion parameter is set to $\gamma = 2$ and the intertemporal elasticity of substitution is equal to $\psi = 1.5$ in the left panels and $\psi = 0.5$ in the right panels.
A Proof of Proposition 1: Price-dividend ratio

In this Appendix we present the detailed derivation of equation (8) in the text. From Campbell-Shiller approximation we have

\[ pd_t \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}. \] (1)

By iterating this equation we find:

\[ pd_t \simeq \kappa + \rho(\kappa + \rho pd_{t+2} + \Delta d_{t+2} - r_{t+2}) + \Delta d_{t+1} - r_{t+1} \]
\[ \quad = \sum_{j=0}^{\infty} \rho^j \kappa + \rho^\infty pd_\infty + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}) \]
\[ \quad = \frac{\kappa}{1 - \rho} + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}) \]

(2)

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 \( t \):

\[ pd_t \simeq \frac{\kappa}{1 - \rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[\Delta d_{t+j} - r_{t+j}] \]
Iterating the dynamics of \( \hat{\mu}_{t+1} \) and \( \hat{g}_{t+1} \) and taking conditional expectation we find

\[
E_t[\hat{\mu}_{t+j}] = \delta_j \hat{\mu}_t
\]

and

\[
E_t[\hat{g}_{t+j}] = \gamma_j \hat{g}_t.
\]

Therefore,

\[
pd_t \simeq \frac{\kappa}{1 - \rho} + \sum_{j=0}^{\infty} \rho^j \left[ \gamma_0 + \gamma_j \hat{g}_t - \delta_0 - \delta_j \hat{\mu}_t \right]
= \frac{\kappa}{1 - \rho} + \frac{\gamma_0 - \delta_0}{1 - \rho \gamma_1} + \frac{\hat{g}_t}{1 - \rho \gamma_1} - \frac{\hat{\mu}_t}{1 - \rho \delta_1}
= A + B_2 \hat{g}_t - B_1 \hat{\mu}_t. \tag{4}
\]

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

\[
A = \frac{\kappa + \gamma_0 - \delta_0}{1 - \rho},
B_1 = \frac{1}{1 - \rho \delta_1},
B_2 = \frac{1}{1 - \rho \gamma_1}.
\]

B Measuring cash flows: dividends or payout?

This Appendix discusses the robustness of our empirical results to different choices of the cash-flow measure.

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

Boudoukh, Michaely, Richardson, and Roberts (2007), among others,\(^1\) find that total and net payout yields have a stronger predictive power for market returns than the dividend yield. It is therefore interesting to look at the effects of such alternative cash-flow measures for the estimation of our present-value model.

Annual time series of repurchases and issuances from 1946 to 2003 are obtained from the dataset constructed by Boudoukh, Michaely, Richardson, and Roberts (2007).\(^2\) Figure II shows the dynamics of yearly cash-flow growth (upper panel) and valuation ratios (lower panel) using different measures of cash-flow: dividend (blue line), total payout (dividend plus repurchases, red line) or net payout (dividend plus repurchases minus issuances, green line).\(^3\) While dividend and total payout share similar patterns, issuances seem to be more related to returns, likely because of strategic firm behaviour. Therefore, we focus our robustness checks on the total payout as an alternative for cash dividends. The second column of Table I reports the results of the estimation of our present-value model using total payout (dividend plus repurchases) as an alternative measure of dividends. The model is estimated by quasi maximum-likelihood in two steps, as explained in the main text. In the first step we use yearly data from 1946 to 2003 on realized second moments of returns and cash flow growth, while in the second step we use yearly data on log total payout growth rates and log price-dividend ratio (adjusted for repurchases). The parameter estimates are qualitatively similar to those shown in the main text (and in the

---


\(^2\)The series are drawn from Michael Roberts’ website: http://finance.wharton.upenn.edu/~mrrobert/.

\(^3\)Returns, as well as the Campbell-Shiller approximation, should also be adjusted for repurchases. However to perform these adjustments, we would need information on the time-varying number of repurchased shares. Therefore, for simplicity we abstract from the issue of time-varying capitalization.
first column of Table I). The persistence of expected cash flow growth and expected returns are almost identical, giving rise to similar predictability implications. In particular, the R-squared for dividend growth is still economically small and statistically insignificant. Parameter $p_1$, which drives the effect of return shocks on expected dividend growth innovations is positive instead of negative but still small, and the main time-varying risk features of the model are qualitatively very similar to those presented in the paper.

C Kalman Filter

In this Appendix we describe the estimation procedure of the model in Section 2 of the paper. As described in Section 3 of the main text, the model is estimated in two steps: in the first step we obtain an estimate of the conditional covariance state $\Sigma_t$ given a preliminary estimate of the the conditional first moments (obtained from the estimate of a constant risk version of the model initially and then from Step 2 in the iteration). In the second step we estimate the conditional first moments $\mu_t$ and $g_t$ given the filtered $\Sigma_t$ from the previous step. These two steps are iterated until convergence of the estimated states and parameters. The following two subsections describe the details of the estimation algorithm for the two steps, and subsection C.3 discusses the estimation algorithm for the constant risk model.

C.1 Estimation of $\Sigma_t$

We estimate the latent covariance state $\hat{\Sigma}_t$ and the parameters $\Delta := (M, k, V)$ using a Kalman filter and maximising a pseudo likelihood. For identification purposes, some parameter constraints are necessary. $M$ is assumed lower triangular, with positive diagonal elements smaller than one. $V$ is assumed diagonal with positive components and $k \geq 2$ is an integer.

For the filter, we first define an expanded state vector by the concatenation of the
original state variable $\hat{\Sigma}_t$ and the process and observation noise random variables:

$$X_{1,t} = \begin{pmatrix} \hat{\Sigma}_{t-1} \\ \varepsilon_t^\Sigma \\ \varepsilon_t^Y \end{pmatrix},$$

which satisfy:

$$X_{1,t+1} = F_1 X_{1,t} + \Gamma_1 \varepsilon_{t+1}^X,$$

where

$$\varepsilon_{t+1}^X = \begin{pmatrix} \varepsilon_{t+1}^\Sigma \\ \varepsilon_{t+1}^Y \end{pmatrix},$$

with conditional variance $Q_t$:

$$Q_{1,t} = \begin{bmatrix} V_t (\varepsilon_{t+1}^\Sigma) & 0_{3 \times 3} \\ 0_{3 \times 3} & V_t (\varepsilon_{t+1}^Y) \end{bmatrix},$$

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

Moreover,

$$F_1 = \begin{bmatrix} S & I_3 & 0_{3 \times 3} \\ 0_{6 \times 9} & I_6 \end{bmatrix},$$

and $\Gamma_1 = \begin{bmatrix} 0_{3 \times 6} \\ I_6 \end{bmatrix}$.

The 3-dimensional measurement equation is given by:

$$Y_{1,t+1} = vech(\mu^\Sigma) + \hat{\Sigma}_t + \varepsilon_{t+1}^Y,$$

where

$$Y_{1,t+1} \equiv [(\Delta d_{t+1} - g_t)^2 (\Delta d_{t+1} - g_t)(r_{t+1} - \mu_t) (r_{t+1} - \mu_t)^2]^T,$$

$\hat{\Sigma}_t = vech(\Sigma_t - \mu^\Sigma)$ is the half-vectorized and centred conditional covariance, and

$$\varepsilon_{t+1}^Y = L_2(\Sigma_t^{1/2} \otimes \Sigma_t^{1/2}) D_2 vech(\varepsilon_{t+1}^W),$$

The measurement is thus of the form:

$$Y_{1,t} = M_0 + M_1 X_{1,t},$$
where

\[ M_0 = \text{vech}(\mu^\Sigma), \quad M_1 = [I_3 \quad 0_{3 \times 3} \quad I_3]. \]

The steps of the filter algorithm are the following:

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

\[
X_{0,0} = 0_{9 \times 1}, \\
P_{0,0} = E(X_tX_t') = \begin{bmatrix} V(\hat{\Sigma}) & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & V(\varepsilon^\Sigma) & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & V(\varepsilon^Y) \end{bmatrix},
\]

where

\[
V(\hat{\Sigma}) = (I_3 - S)^{-1}L_2(I_4 + K_{2,2})k(V \otimes V)L_2', \\
V(\varepsilon^\Sigma) = L_2(I_4 + K_{2,2})k(V \otimes V)L_2'.
\]

- The time-update equations are

\[
X_{t,t-1} = F_1X_{t-1,t-1}, \\
P_{t,t-1} = F_1P_{t-1,t-1}F_1' + \Gamma_1Q_{t-1}\Gamma_1'.
\]

- The prediction error \( \eta_t \) and the variance-covariance matrix of the measurement equations are then:

\[
\eta_t = Y_t - M_0 - M_1X_{t,t-1}, \\
S_t = M_1P_{t,t-1}M_1'.
\]

- Update filtering:

\[
K_t = P_{t,t-1}M_1'S_t^{-1}, \\
X_{t,t} = X_{t,t-1} + K_t\eta_t, \\
P_{t,t} = (I - K_tM_1)P_{t,t-1},
\]

where \( K_t \) is called \textit{Kalman gain}. 

6
To estimate model parameters, $\Delta$, we define the pseudo log-likelihood for each time $t$ as

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

where $\eta_t$ 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 pseudo log-likelihood of the data series:

$$\Delta \equiv \arg \max_{\Delta} \mathcal{L}(\Delta, \{Y_{1,t}\}_{t=1}^T),$$

with

$$\mathcal{L}(\Delta, \{Y_{1,t}\}_{t=1}^T) = \sum_{t=1}^T l_t(\Delta),$$

where $T$ denotes the number of periods in the sample of estimation.

The result of this first estimation is a filtered time series of the conditional variance-covariance matrix $\Sigma_t$ and an estimate of the parameters $\hat{\Delta}$.

### C.2 Estimation of $\mu_t$ and $g_t$

The following step is an estimation of the conditional first moments for given conditional second moment $\Sigma_t$.

The state variable in this step is the expected return $\mu_t$ (expected dividend growth $g_t$ can then be derived from the estimated parameters, $\mu_t$ and $pd$). Observable variables are dividend growth $\Delta d_t$ and the price-dividend ratio $pd_t$.

The transition dynamics is given by:

$$\hat{\mu}_{t+1} = \delta_1 \hat{\mu}_t + \varepsilon_{\hat{\mu}_{t+1}}.$$

The measurement equation for dividend growth is as follows:

$$\Delta d_{t+1} = \gamma_0 + \hat{\gamma}_t + \tilde{\varepsilon}^D_{t+1}$$

$$= \gamma_0 + \frac{1}{B_2} (pd_t - A + B_1 \hat{\mu}_t) + \tilde{\varepsilon}^D_{t+1}. \quad (13)$$

The measurement equation for the log price-dividend ratio is given by:

$$pd_{t+1} = A - B_1 \delta_1 \hat{\mu}_t + B_2 \hat{\gamma}_t + \frac{1}{\rho} (\tilde{\varepsilon}^r_{t+1} - \tilde{\varepsilon}^D_{t+1}). \quad (14)$$
We then apply a standard Kalman filter to obtain an exponential quadratic pseudo likelihood and estimate the model parameters \( \Xi := (\gamma_0, \delta_0, \gamma_1, \delta_1, p_1, p_2) \) using pseudo maximum likelihood. For identification purposes, parameters \( \delta_1 \) and \( \gamma_1 \) are bounded to be less than one in absolute value.

For the Kalman filter estimation, we first define an expanded state vector by the concatenation of the original state variable and the noise random variables:

\[
X_{2,t} = \begin{pmatrix}
\hat{\mu}_{t-1} \\
\hat{\varepsilon}_{t}^D \\
\hat{\varepsilon}_{t}^r
\end{pmatrix},
\]

which satisfy:

\[
X_{2,t+1} = F_2 X_{2,t} + \Gamma_2 \varepsilon_{t+1}^X,
\]

where

\[
\varepsilon_{t+1}^X = \begin{pmatrix}
\varepsilon_{t+1}^D \\
\varepsilon_{t+1}^r
\end{pmatrix},
\]

with conditional variance \( \Sigma_t \), which is observed (from the previous step). Moreover,

\[
F_2 = \begin{bmatrix}
\delta_1 & \frac{1-p_2}{\rho B_1} & \frac{p_1-1}{\rho B_1} \\
0_{2 \times 3} & \frac{1}{\rho B_1}
\end{bmatrix}, \quad \Gamma_2 = \begin{bmatrix}
0_{1 \times 2} \\
I_2
\end{bmatrix}.
\]

The 2-dimensional measurement equation,

\[
Y_{2,t} = \begin{pmatrix}
\Delta d_t \\
p d_t
\end{pmatrix},
\]

is of the form

\[
Y_{2,t} = M_0 Y_{2,t-1} + M_1 X_{2,t},
\]

where

\[
M_0 = \begin{bmatrix}
\gamma_0 - \frac{A}{B_2} \\
(1 - \gamma_1) A
\end{bmatrix}, \quad M_1 = \begin{bmatrix}
0 & \frac{1}{B_2} \\
0 & \gamma_1
\end{bmatrix},
\]

and

\[
M_2 = \begin{bmatrix}
\frac{B_1}{B_2} & 1 & 0 \\
B_1 (\gamma_1 - \delta_1) & -\frac{1}{\rho} & \frac{1}{\rho}
\end{bmatrix}.
\]

The steps of the filter algorithm are the following:
• Initialize with the unconditional mean and covariance of the expanded state:

\[
X_{0,0} = 0_{3 \times 1},
\]

\[
P_{0,0} = E(X_tX_t') = \begin{bmatrix}
\frac{\sigma^2_\mu}{(1-\delta_1)^2} & 0_{1 \times 2} \\
0_{2 \times 1} & \Sigma
\end{bmatrix},
\]

where \( \Sigma \) is the unconditional mean of the filtered \( \Sigma_t \) and \( \sigma^2_\mu \) is given by:

\[
\sigma^2_\mu = \frac{1}{\rho^2 B^2_2} [(p_1 - 1)^2 \Sigma_{22} + (p_2 - 1)^2 \Sigma_{11} - 2(p_1 - 1)(p_2 - 1)\Sigma_{12}].
\]

• The time-update equations are

\[
X_{t,t-1} = F_2X_{t-1,t-1},
\]

\[
P_{t,t-1} = F_2P_{t-1,t-1}F_2' + \Gamma_2\Sigma_{t-1}\Gamma_2'.
\]

• The prediction error \( \eta_t \) and the variance-covariance matrix of the measurement equations are then:

\[
\eta_t = Y_t - M_0 - M_1Y_{t-1} - M_2X_{t,t-1},
\]

\[
S_t = M_2P_{t,t-1}M_2'.
\]

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

• Update filtering:

\[
K_t = P_{t,t-1}M_2'S_t^{-1},
\]

\[
X_{t,t} = X_{t,t-1} + K_t\eta_t,
\]

\[
P_{t,t} = (I - K_tM_2)P_{t,t-1}.
\]

To estimate model parameters, \( \Xi \), we define the pseudo log-likelihood for each time \( t \) as

\[
l_t(\Xi) = -\frac{1}{2} \log |S_t| - \frac{1}{2} \eta_t'S_t^{-1}\eta_t, \tag{15}
\]

where \( \eta_t \) 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 pseudo log-likelihood of the data series:

\[
\Xi \equiv \operatorname{arg \ max}_\Xi \mathcal{L}(\Xi, \{Y_{2,t}\}_{t=1}^T), \tag{16}
\]
with
\[
\mathcal{L}(\Xi, \{Y_{2,t}\}_{t=1}^{T}) = \sum_{t=1}^{T} l_t(\Xi).  
\]  

(17)

C.3 Constant risks model

The relevant state variable in the constant risk version of our present-value model is the expected return \( \mu_t \) and observable variables are dividend growth \( \Delta d_t \) and the price-dividend ratio \( pd_t \). The transition and measurement dynamics is described in Section 3.3 of the main text.

The resulting state space model is fully linear and we can apply a standard Kalman filter to obtain an exponential quadratic pseudo likelihood and estimate the model parameters \( \Xi_0 := (\gamma_0, \delta_0, \gamma_1, \delta_1, \Sigma_{11}, \Sigma_{12}, \Sigma_{22}, p_1, p_2) \) using pseudo maximum likelihood.

We first define an expanded state vector by the concatenation of the original state variable and the noise random variables:

\[
X_t = \begin{pmatrix}
\hat{\mu}_{t-1} \\
\tilde{\varepsilon}_D^t \\
\tilde{\varepsilon}_r^t
\end{pmatrix},
\]

which satisfy:

\[
X_{t+1} = FX_t + \Gamma \tilde{\varepsilon}^X_{t+1},
\]

where

\[
\tilde{\varepsilon}^X_{t+1} = \begin{pmatrix}
\tilde{\varepsilon}_D^{t+1} \\
\tilde{\varepsilon}_r^{t+1}
\end{pmatrix},
\]

with constant variance \( \Sigma \). Moreover,

\[
F = \begin{bmatrix}
\delta_1 & \frac{1-p_2}{p_1} & \frac{p_1-1}{p_2} \\
0_{2 \times 3}
\end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0_{1 \times 2} \\ I_2 \end{bmatrix}.
\]

The 2-dimensional measurement equation,

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

is of the form

\[
Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t,
\]
where

\[ M_0 = \begin{bmatrix} \frac{\gamma_0 - \frac{A}{B^2}}{(1 - \gamma_1)A} \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & \frac{1}{B^2} \\ 0 & \gamma_1 \end{bmatrix}, \]

and

\[ M_2 = \begin{bmatrix} \frac{B_1}{B_2} & 1 & 0 \\ B_1(\gamma_1 - \delta_1) & -\frac{1}{\rho} & \frac{1}{\rho} \end{bmatrix}. \]

The steps of the filter algorithm are the following:

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

\[
X_{0,0} = 0_{3 \times 1}, \\
P_{0,0} = E(X_tX_t') = \begin{bmatrix} \frac{\sigma^2}{(1 - \delta_1)^2} & 0_{1 \times 2} \\ 0_{2 \times 1} & \Sigma \end{bmatrix},
\]

where \( \sigma^2 \) is given by:

\[
\sigma^2 = \frac{1}{\rho^2B_2^2}[(p_1 - 1)^2\Sigma_{22} + (p_2 - 1)^2\Sigma_{11} - 2(p_1 - 1)(p_2 - 1)\Sigma_{12}].
\]

- The time-update equations are

\[
X_{t,t-1} = FX_{t-1,t-1}, \\
P_{t,t-1} = FP_{t-1,t-1}F' + \Gamma \Sigma \Gamma'.
\]

- The prediction error \( \eta_t \) and the variance-covariance matrix of the measurement equations are then:

\[
\eta_t = Y_t - M_0 - M_1Y_{t-1} - M_2X_{t,t-1}, \\
S_t = M_2P_{t,t-1}M_2',
\]

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

- Update filtering:

\[
K_t = P_{t,t-1}M_2'S_t^{-1}, \\
X_{t,t} = X_{t,t-1} + K_t\eta_t, \\
P_{t,t} = (I - K_tM_2)P_{t,t-1}
\]
To estimate model parameters, $\Xi_0$, we define the pseudo log-likelihood for each time $t$ as

$$l_t(\Xi_0) = -\frac{1}{2} \log |S_t| - \frac{1}{2} \eta_t' S_t^{-1} \eta_t,$$

where $\eta_t$ 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 pseudo log-likelihood of the data series:

$$\Xi_0 \equiv \arg \max_{\Xi_0} \mathcal{L}(\Xi_0, \{Y_t\}_{t=1}^T),$$

with

$$\mathcal{L}(\Xi_0, \{Y_t\}_{t=1}^T) = \sum_{t=1}^{T} l_t(\Xi_0),$$

where $T$ denotes the number of time periods in the sample of estimation.

**D  Asymptotic bias in standard predictive regressions**

We have argued in Section 4.2 of the main text 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. Indeed, the true model for aggregate stock returns is:

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

but we wrongly assume the following model to hold:

$$r_{t+1} = a_r + b_r pd_t + \tilde{\varepsilon}_{t+1},$$

where $pd_t = A - B_1 \hat{\mu}_t + B_2 \hat{g}_t$, and we try to estimate the true parameter $b_r = -1/B_1$ from (22). The p-limit of the OLS slope coefficient is the following:4

$$\hat{b}_r \rightarrow \frac{\text{Cov}(pd_t, r_{t+1})}{\text{Var}(pd_t)},$$

where

$$\text{Cov}(pd_t, r_{t+1}) = \text{Cov}(A - B_1 \hat{\mu}_t + B_2 \hat{g}_t, \delta_0 + \hat{\mu}_t + \tilde{\varepsilon}_{t+1})$$

$$= -B_1 \text{Var}(\hat{\mu}_t) + B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t)$$

$$\text{Var}(pd_t) = B_1^2 \text{Var}(\hat{\mu}_t) + B_2^2 \text{Var}(\hat{g}_t) - 2B_1B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t)$$

4Note that here we denote with $b_r$ the OLS estimate of the slope coefficient $b_r$ in (22).
so that
\[ \hat{b}_r \rightarrow \frac{1}{-B_1 + \frac{B_2^{\rho} \text{Var}(\hat{g}_t) - B_1 B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t)}{B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t) - B_1 \text{Var}(\hat{\mu}_t)}} \] (24)
and the unconditional variances and covariance of demeaned expected return and dividend growth are the following:
\[
\begin{align*}
\text{Var}(\hat{\mu}_t) &= \left((p_2 - 1)^2 - 2(p_1 - 1)(p_2 - 1)\right) \frac{(p_1 - 1)^2}{\rho^2 B_1^2 (1 - \delta_1^2)} \text{vech}(\mu^\Sigma), \\
\text{Var}(\hat{g}_t) &= \left[p_2^2 - 2p_1p_2 + p_1^2\right] \frac{\text{vech}(\mu^\Sigma)}{\rho^2 B_2^2 (1 - \gamma_1^2)}, \\
\text{Cov}(\hat{g}_t, \hat{\mu}_t) &= \left[p_2(p_2 - 1) - (2p_1p_2 - p_1 - p_2)\right] \frac{p_1(p_1 - 1)}{\rho^2 B_1 B_2 (1 - \delta_1)(1 - \gamma_1)} \text{vech}(\mu^\Sigma).
\end{align*}
\]
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 smaller in absolute value than the true one, 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:
\[ \Delta d_{t+1} = \gamma_0 + \hat{g}_t + \hat{\epsilon}_D, \] (25)
while the wrong model is:
\[ \Delta d_{t+1} = a_D + b_D pd_t + \hat{\epsilon}_D, \] (26)
and we try to estimate the true parameter \( b_D = 1/B_2 \) from (26). The p-limit of the OLS slope is the following:
\[ \hat{b}_D \rightarrow \frac{\text{Cov}(pd_t, \Delta d_{t+1})}{\text{Var}(pd_t)}, \] (27)
where
\[
\begin{align*}
\text{Cov}(pd_t, \Delta d_{t+1}) &= \text{Cov}(A - B_1 \hat{\mu}_t + B_2 \hat{g}_t, \gamma_0 + \hat{g}_t + \hat{\epsilon}_D) \\
&= B_2 \text{Var}(\hat{g}_t) - B_1 \text{Cov}(\hat{g}_t, \hat{\mu}_t)
\end{align*}
\]
so that
\[ \hat{b}_D \rightarrow \frac{1}{B_2 + \frac{B_2^{\rho} \text{Var}(\hat{g}_t) - B_1 B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t)}{B_2 \text{Cov}(\hat{g}_t, \hat{\mu}_t) - B_1 \text{Var}(\hat{\mu}_t)}} \] (28)
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, more significant than the one for standard return regressions.
E Derivation of VAR

This Appendix provides the approximate expressions for the observable VAR counterparts implied by our time-varying risks state-space model. Using the Kalman filter in Appendix C.2 recursively, we can express the filtered state in terms of historical observables, i.e. dividend growth and price-dividend ratio.

\[
X_{t,t} = X_{t,t-1} + K_t(Y_t - M_0 - M_1Y_{t-1} - M_2X_{t,t-1})
\]

\[
= (I - K_tM_2)X_{t,t-1} + K_t(Y_t - M_0 - M_1Y_{t-1})
\]

\[
= \ldots
\]

\[
= K_t(Y_t - M_0 - M_1Y_{t-1}) + \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} (I - K_{t-j}M_2) F K_{t-i}(Y_{t-i} - M_0 - M_1Y_{t-1-i}).
\]

The first element of the vector \( X_t \) is \( \mu_{t-1} \), thus \( \mu_{t-1,t-1} \) is the first element of

\[
X_{t,t-1} = FX_{t-1,t-1}
\]

\[
= FK_{t-1}(Y_{t-1} - M_0 - M_1Y_{t-2}) + F \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} (I - K_{t-j}M_2) F K_{t-1-i}(Y_{t-1-i} - M_0 - M_1Y_{t-2-i}).
\]

The filtered value of expected dividend growth, \( \hat{g}_{t-1,t-1} \) is obtained exploiting present-value relation:

\[
\hat{g}_{t-1,t-1} = \frac{1}{B_2}(pd_{t-1} - A + B_1\hat{\mu}_{t-1,t-1}).
\]

We define \( \varepsilon_t^* = r_t - \delta_0 - \hat{\mu}_{t-1,t-1} \) and obtain

\[
r_t = \delta_0 + \varepsilon_t^* X_{t,t-1} + \varepsilon_t^* = \delta_0 + \varepsilon_t^* F K_{t-1}(Y_{t-1} - M_0 - M_1Y_{t-2}) + \varepsilon_t^*
\]

\[
\sum_{i=1}^{\infty} \prod_{j=0}^{i-1} (I - K_{t-j}M_2) F K_{t-1-i}(Y_{t-1-i} - M_0 - M_1Y_{t-2-i}) + \varepsilon_t^* = a_0^r + \sum_{i=0}^{\infty} a_{1i}^r \Delta d_{t-1-i} + \sum_{i=0}^{\infty} a_{2i}^r pd_{t-1-i} + \varepsilon_t^*,
\]

where \( a_0^r(\Sigma_{t-1}) \), \( a_{1i}^r(\Sigma_{t-i-1}) \) and \( a_{2i}^r(\Sigma_{t-i-1}) \) are time-varying coefficients that depend on the history of the variance covariance state up to time \( t-1 \).

For dividends, we define \( \varepsilon_t^{D*} = \Delta d_t - \gamma_0 - \hat{g}_{t-1,t-1} \) and obtain

\[
\Delta d_t = \gamma_0 + \frac{1}{B_2}(pd_{t-1} - A + B_1\hat{\mu}_{t-1,t-1}) + \varepsilon_t^{D*}
\]
where $a^D_0(\Sigma_{t-1})$, $a^D_{1i}(\Sigma_{t-i-1})$ and $a^D_{2i}(\Sigma_{t-i-1})$ are time-varying coefficients that depend on the history of the variance covariance state up to time $t-1$.

These time-varying weights, for both the return and dividend growth expressions, could explain why predictability regressions give different results according to the sample considered.

## F Risk-return tradeoff and Sharpe ratio

While understanding the dynamic relation between market return and market risk is a central topic in financial economics, the empirical evidence is largely ambiguous and typically dependent on the model choice or the instruments used to specify the conditional information set.

### F.1 Benchmark model and filtered states

In our model shocks to return volatility and expected returns are assumed to be conditionally uncorrelated. For simplicity, we do not include explicitly any leverage or volatility feedback effects. However, the filtered dynamics of expected returns and dividends are naturally related. The VAR representation of the model in Supplemental Appendix E shows explicitly how expected returns are a function of all historical observables, i.e. dividend growth and price-dividend ratio, with coefficients that depend on the history of the conditional variance-covariance matrix of returns and dividends.

We can therefore use the filtered time series of expected returns and return conditional variances estimated by our model to study the dynamics of the resulting risk-return tradeoff. This approach allows us to study the co-movement of market return mean...
and variance in a parsimonious but flexible framework, without relying on additional assumptions about exogenous instruments.

The time series of expected returns and return volatility in Figure III are largely but imperfectly correlated, which is a first indication of a time-varying risk-return tradeoff. Following Lettau and Ludvigson (2010), a regression of filtered variable $\hat{\mu}_t$ on $\Sigma_{22,t}$ gives an indication about the sign of the unconditional risk-return relation:

$$\hat{\mu}_t = \alpha_0 + \alpha_1 \Sigma_{22,t} + \varepsilon_t. \quad (31)$$

The negative regression coefficient $\hat{\alpha}_1$ in the first panel of Table II provides a first evidence of a negative unconditional risk-return tradeoff. Following Whitelaw (1994) and Brandt and Kang (2004), Lettau and Ludvigson (2010) stress the importance of including lagged means and volatilities in model (31), when studying conditional risk-return relations:

$$\hat{\mu}_t = \alpha_0 + \alpha_1 \Sigma_{22,t} + \alpha_2 \hat{\mu}_{t-1} + \alpha_3 \Sigma_{22,t-1} + \varepsilon_t. \quad (32)$$

In the second panel of Table II, the parameter estimates $\hat{\alpha}_2$ and $\hat{\alpha}_3$ for lagged expected return and lagged return variance are statistically significant and add substantially to the explanatory power of the regression. In particular, even though the lagged expected return explains a large fraction of the return variation, due to the persistence of expected returns, a large lagged conditional variance significantly predicts low future expected returns, with a t-stat of about $-3.87\%$. Lastly, the parameter $\hat{\alpha}_1$ measuring the risk-return relation is still significant and turns positive, once lead-lag interactions are taken into account.

Our finding of a significant negative unconditional risk-return relation and a positive conditional risk-return tradeoff is consistent with the evidence in Lettau and Ludvigson (2010), who estimate a mean-variance dynamics for returns based on a conditioning information set generated by a family of additional exogenous instruments. We obtain this finding using exclusively the joint information generated by returns, dividend growth and price dividend-ratios, while explicitly incorporating the model’s present-value constraints with a latent variables approach.\textsuperscript{5}

\textsuperscript{5}In order to avoid relying on exogenous instruments, Brandt and Kang (2004) follow a latent variables approach that does not incorporate present-value constraints and dividend growth information. Using a
The joint dynamics of expected returns and return volatilities in our model also imply a time-varying price of market volatility. Conditional Sharpe ratios are defined as the ratio of conditional excess expected returns and volatility, which requires assumptions on the riskless interest rate $r^{f}_t$ for their computation:

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

We compute the time series of Sharpe ratios $SR_t$ in our model, assuming $r^{f}_t$ is equal to zero (see Figure IV). Consistently with the evidence in the literature, we find that the Sharpe ratios estimated by our model are often countercyclical and quite volatile, with a standard deviation of about 0.49. In contrast, the conditional Sharpe ratio implied by the model with constant risks is both less countercyclical and not as volatile, with a standard deviation of about 0.29. For comparison, the conditional Sharpe ratio estimated by vector autoregressions for mean and variance has a volatility ranging from about 0.45 to about 0.7, depending on the instruments used to model the conditioning informations set; see Lettau and Ludvigson (2010), among others.\footnote{In such models, return mean and variance are estimated using linear regressions of the form:} In summary, the Sharpe ratio dynamics in our model is highly volatile and countercyclical as suggested by Lettau and Ludvigson (2010) and Lustig and Verdelhan (2012).

\section*{F.2 Model extension with risk-in-mean effect}

Our benchmark model can be extended to introduce explicit dependencies on $\Sigma_t$ in the conditional expectation dynamics. Potentially, this would imply three additional parameters for both equation (4) and (5) in the main text. In this Appendix we show the model first-order linear Gaussian processes, they specify an heteroskedastic univariate dynamics for log returns and obtain a positive (negative) unconditional (conditional) risk-return relation.\footnote{In such models, return mean and variance are estimated using linear regressions of the form:}

$$r_{t+1} = \beta'_r Z_t + \varepsilon^r_{t+1},$$ and

$$RV^{r}_{t+1} = \beta'_v Z_t + \varepsilon^{v}_{t+1},$$ respectively, where $Z_t$ is a vector of predetermined conditioning variables. For instance, in Lettau and Ludvigson (2010) vector $Z_t$ contains $cay_t$ and two lags of the realized variance in the equation for volatility, while it contains $cay_t$ and the risk free rate in the equation for return means.
structure for the case of a risk-in-mean effect in returns only, which is more relevant for the risk-return tradeoff literature, and we provide the explicit expressions for the $pd$ ratio in this case.

Equation (5) in the main text can be extended as follows:

$$
\mu_{t+1} = \delta_0 + \delta_1 (\mu_t - \delta_0) + Tr \left( \Lambda (\Sigma_t - \mu^2) \right) + \varepsilon^\mu_{t+1},
$$

(33)

where $Tr$ denotes the trace operator and $\Lambda$ is a $2 \times 2$ symmetric matrix. In terms of demeaned states, dynamics (33) can be written as:

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

(34)

where $N = D_2 vec(\Lambda)$ is a 3-dimensional vector.

A conditional mean specification that is affine in $\Sigma_t$ is still tractable and induces a closed-form $pd-$ratio that is affine in $\mu_t$, $g_t$ and $\Sigma_t$, extending the result in Proposition 1 in the main text. More specifically,

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

(35)

where $A$, $B_1$ and $B_2$ are as in the benchmark model and

$$
B_3 = N' \left[ (\rho S^2 - (1 + \rho \delta_1)S + \delta_1 I_3)^{-1} + B_1 (S - \delta_1 I_3)^{-1} \right].
$$

The addition of three parameters to estimate is problematic given the limited amount of data at annual frequency, but we can estimate special cases of this extended model. For example, we can restrict the first two elements of the $N$ vector to be equal to zero, leaving the expected return depend only on the return variance. When estimating this version of the model we find parameter estimates very similar to the benchmark model and an extremely small and statistically insignificant parameter for the volatility feedback, i.e. $N = [0 \ 0 \ -0.014]'$, with a 95% bootstrap confidence interval given by $[-0.082 \ 0.337]$. Therefore, we focus on the simpler model without volatility feedbacks in the paper. We feel that the separation between conditional first and second moments for annual observation frequencies also makes the comparison between benchmark models with constant risks and models with time-varying risks more transparent, as differences

---

7Detailed estimation results are available from the authors on request.
essentially directly arise from the heteroskedasticity of dividends/returns versus expected dividends/returns induced by the Campbell-Shiller identities, and not from a different functional form of the \( \rho d \)-ratio in presence of volatility feedbacks, for which we do not find a statistically significant evidence at annual frequencies.

G  Term Structures of Unweighted Returns and Dividends

Our definition of long horizon returns and dividend growth underlying equations (39)-(42) of the main text is borrowed from Cochrane (2008) and ensures that the total expected long horizon return and dividend growth exactly measure for \( n \rightarrow \infty \) the \( \rho d \)-ratio variation induced by returns and dividend growth in the Campbell-Shiller identity. In our data, \( \hat{\rho} = 0.972 \) (see Table 1 of the main text). Therefore, we plot the various term structures for a parameter \( \rho \) that is nearly 1 in the definition of long-horizon returns and dividend growth. A consequence of Cochrane’s definition for \( \rho = 0.972 \) is a larger sensitivity of long horizon returns and dividend growth to deviations of annual expected return \( \mu_t \) and dividend growth \( g_t \) from their long term means \( \delta_0 \) and \( \gamma_0 \), respectively, i.e., quantities \( \hat{\mu}_t \) and \( \hat{g}_t \) in equations (39)-(40) of the main text. The term structures based on weighted quantities are typically below those based on unweighted quantities for horizons above one year. They are also more downward sloping and less convex (less upward sloping and more concave) when annual expected quantities are above (below) the long term mean. However, at the estimated parameters the key properties of the term structure dynamics of long horizon expectations and risks under either definition (such as, e.g., the dynamics and cyclicity of the slope of the term structure) are similar, as well as the term structures of risk. To show this explicitly we provide in this Appendix the closed-form solutions for the term structures of unweighted returns and dividend expectations and variances, and we plot these term structures at the calibrated parameters.

19
G.1 Time-varying term structures of expectations

We define here long-horizon returns and dividend growth as the simple sum of annual log returns and dividend growth. By applying recursively equations (3)-(5) in the main text and taking conditional expectations, we obtain explicit expressions for the model-implied $n$-year return and dividend growth:

$$
\mu_t^{(n)} := \frac{1}{n} E_t \left[ \sum_{j=1}^{n} r_{t+j} \right] = \delta_0 + \frac{1 - \delta^n_1}{n(1 - \delta_1)} \hat{\mu}_t,
$$

(36)

$$
g_t^{(n)} := \frac{1}{n} E_t \left[ \sum_{j=1}^{n} \Delta d_{t+j} \right] = \gamma_0 + \frac{1 - \gamma^n_1}{n(1 - \gamma_1)} \hat{g}_t.
$$

(37)

The left panel of Figure V plots the time series of the term structure of return expectations $\mu_t^{(n)}$ in equation (36) while the term structure of dividend growth expectations $g_t^{(n)}$ is in the left panel of Figure V. Figure VI displays the two term structures in four selected years. Figures V and VI are equivalent to Figures 6 and 7 in the main text of the paper but for the case of unweighted long-horizon returns and dividends. The term structures based on weighted quantities are typically more downward sloping and less convex (less upward sloping and more concave) when annual expected quantities are above (below) the long term mean. However, at the estimated parameters the key properties of the term structure dynamics of long horizon expected return and dividend growth under either definition (such as, e.g., the dynamics and cyclicality of the slope of the term structure) are similar.

G.2 Time-varying term structures of risks

By applying recursively equations (3)-(5) in the main text and taking conditional vari-
ances, we obtain the closed-form annualized conditional variance of a $n$-year return:

$$
\frac{1}{n} Var_t \left[ \sum_{j=1}^{n} r_{t+j} \right] = \frac{1}{n} \sum_{j=1}^{n-1} \left( \frac{1 - \delta^{n-j}_1}{1 - \delta_1} \right)^2 Var_t(\varepsilon_{t+j}^\mu) + \frac{1}{n} \sum_{j=1}^{n} Var_t(\tilde{\varepsilon}_{t+j}^\mu) + 2 \frac{2}{n} \sum_{j=1}^{n-1} \frac{1 - \delta^{n-j}_1}{1 - \delta_1} Cov_t(\varepsilon_{t+j}^\mu, \tilde{\varepsilon}_{t+j}^\mu),
$$

(38)

while the annualized conditional variance of $n$-year dividend growth is given by:

$$
\frac{1}{n} Var_t \left[ \sum_{j=1}^{n} \Delta d_{t+j} \right] = \frac{1}{n} \sum_{j=1}^{n-1} \left( \frac{1 - \gamma^{n-j}_1}{1 - \gamma_1} \right)^2 Var_t(\varepsilon_{t+j}^g) + \frac{1}{n} \sum_{j=1}^{n} Var_t(\tilde{\varepsilon}_{t+j}^g)
$$
\[ + \frac{2}{n} \sum_{j=1}^{n-1} \frac{1}{1-\gamma_1} \text{Cov}_t(\varepsilon_{t+j}^\mu, \tilde{\varepsilon}_{t+j}^D), \]  

(39)

where \( \text{Var}_t(\varepsilon_{t+j}^\mu), \text{Var}_t(\tilde{\varepsilon}_{t+j}^r) \) and \( \text{Cov}_t(\varepsilon_{t+j}^\mu, \tilde{\varepsilon}_{t+j}^r) \) are affine functions of the variance-covariance state \( \Sigma_t \), given explicitly in Appendix A.2 of the main text. Equations (38) and (39) imply time-varying term structures, with dynamics summarized in Figure VII, while Figure VIII shows the term structures in four selected years, as in Figures 10 and 11 of the main text.

Figure IX presents the three components of the term structure of market risk in three different years, characterized by different levels of short term return volatility, as in Figure 12 of the text. Figure X does the same for the term structure of dividend risk, as in Figure 13 of the main text.

In general, the term structures of risks are qualitatively very similar for weighted or unweighted long horizon returns and dividend growth.
H Tables and Figures

Table I: Estimation results for our time-varying risks present-value model with an alternative measure of cash flow growth. The first column refers to the estimation results discussed in the paper, based on CRSP value-weighted index from 1946 to 2015. The second column reports the results using total payout (dividend plus repurchases) as an alternative measure of dividends. The model is estimated by quasi maximum-likelihood using yearly data from 1946 to 2003.

<table>
<thead>
<tr>
<th>Sample period</th>
<th>CRSP</th>
<th>Payout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1946-2015</td>
<td>1946-2003</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.055</td>
<td>0.065</td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>0.083</td>
<td>0.108</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.675</td>
<td>0.653</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.897</td>
<td>0.889</td>
</tr>
<tr>
<td>$p_1$</td>
<td>-0.405</td>
<td>0.069</td>
</tr>
<tr>
<td>$p_2$</td>
<td>2.332</td>
<td>2.967</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>0.708</td>
<td>0.339</td>
</tr>
<tr>
<td>$M_{21}$</td>
<td>-0.494</td>
<td>0.499</td>
</tr>
<tr>
<td>$M_{22}$</td>
<td>0.936</td>
<td>0.929</td>
</tr>
<tr>
<td>$k$</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$V_{11}$</td>
<td>0.0021</td>
<td>0.0032</td>
</tr>
<tr>
<td>$V_{22}$</td>
<td>0.0010</td>
<td>0.0004</td>
</tr>
<tr>
<td>$R^2_{ret}$</td>
<td>8.11</td>
<td>15.80</td>
</tr>
<tr>
<td>$R^2_{div}$</td>
<td>0.77</td>
<td>2.75</td>
</tr>
</tbody>
</table>
Table II: Conditional and unconditional risk-return tradeoff.

The first panel reports results from the regression of filtered expected return $\mu_t$ on the conditional variance $\Sigma_{22,t}$. The second panel also considers lags in the regression (see equation (32)). The numbers in parenthesis are Newey-West corrected t-statistics.

<table>
<thead>
<tr>
<th>Constant</th>
<th>$\Sigma_{22,t}$</th>
<th>$\hat{\mu}_{t-1}$</th>
<th>$\Sigma_{22,t-1}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1118</td>
<td>-0.7399</td>
<td></td>
<td></td>
<td>13.25%</td>
</tr>
<tr>
<td>(7.4330)</td>
<td>(-4.8219)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0100</td>
<td>0.5509</td>
<td>0.9144</td>
<td>-0.6831</td>
<td>84.50%</td>
</tr>
<tr>
<td>(1.2952)</td>
<td>(3.6026)</td>
<td>(16.8102)</td>
<td>(-3.8696)</td>
<td></td>
</tr>
</tbody>
</table>
**Table III**: Estimation results for the standard constant risks model, with identification assumption \( \rho_{gD} = 0 \). The model is estimated by maximum-likelihood, using yearly data from 1946 to 2015 on log dividend growth rates and log price-dividend ratio. Panel A presents estimates of the coefficients of the underlying processes. Panel B reports resulting coefficients of the present-value decomposition. Bootstrapped standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Panel A: Quasi maximum-likelihood estimates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_0 )</td>
<td>( \delta_0 )</td>
</tr>
<tr>
<td>0.060</td>
<td>0.087</td>
</tr>
<tr>
<td>(0.018)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>( \sigma_g )</td>
<td>( \sigma_\mu )</td>
</tr>
<tr>
<td>0.0669</td>
<td>0.0145</td>
</tr>
<tr>
<td>(0.0213)</td>
<td>(0.0291)</td>
</tr>
<tr>
<td>Panel B: Implied present-value parameters</td>
<td></td>
</tr>
<tr>
<td>( \rho )</td>
<td>( A )</td>
</tr>
<tr>
<td>0.9734</td>
<td>3.5984</td>
</tr>
</tbody>
</table>

24
Figure I: Conditional correlation of the expectation processes, $\text{Corr}_t(\varepsilon_{t+1}^g, \varepsilon_{t+1}^\mu)$, as a function of the parameters $p_1$ and $p_2$, fixing $\Sigma_t$ at the unconditional mean of the realized variance-covariance matrix of returns and dividend growth.
**Figure II:** Yearly cash-flow growth (upper panel) and ratio of price over cash-flow (lower panel) from 1946 to 2003, using different measures of cash-flow: dividend (blue line), total payout (dividend plus repurchases, red line) or net payout (dividend plus repurchases minus issuances, green line).
**Figure III:** Risk-Return tradeoff. Filtered values of conditional expected returns, $\mu_t$ (blue line, left axis) against conditional variance of returns, $\Sigma_{22,t}$ (red line, right axis).
**Figure IV:** 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. The dashed red line is obtained in the same way, but for a version of the model with constant risks.
Figure V: Dynamics of the term structure of the conditional per-period expected long-horizon return ($\mu_t^{(n)}$, left panel) and dividend growth ($g_t^{(n)}$, right panel), from equations (36) and (37), respectively, computed using estimated parameters and filtered state. We consider horizons of 1 to 20 years.
Figure VI: Term structure of the conditional per-period expected long-horizon return (left axes, $\mu_t^{(n)}$) and dividend growth (right axes, $g_t^{(n)}$), from equations (36) and (37), in years 1973, 1990, 2000 and 2011.
Figure VII: Dynamics of the term structure of the conditional per-period long-horizon return (left panel) and dividend growth (right panel) volatility, from equations (38) and (39), respectively, computed using estimated parameters and filtered state. We consider horizons of 1 to 20 years.
Figure VIII: Term structure of the conditional per-period volatilities of long-horizon return (left axes) and dividend growth (right axes), from equations (38) and (39), in years 1967, 1990, 2000 and 2010.
**Figure IX:** 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 1 to 20 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 X:** Decomposition of the term structure of the conditional per-period variance of long-horizon dividend growth, computed using estimated parameters and filtered states. The blue line denotes the component of the variance that is due to uncertainty about future expected cash flow growth, the red line denotes the component due to future dividend growth shocks, while the green line denotes the mean reversion component. The black dashed line denotes the total conditional variance, for horizons of 1 to 20 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.
References


