Online Appendix: The Term Structure of Equity Risk Premia*

Ravi Bansal,† Shane Miller,‡ Dongho Song,§ and Amir Yaron¶

Abstract

We estimate a regime-switching model for the equity term structure with Bayesian methods. Our approach accounts for the data sample being unrepresentative of the population distribution of regimes. We find that (i) the term structure of expected equity dividend strip returns is downward sloping in recessions and upward sloping in expansions, and (ii) the unconditional slope of the term structure of expected equity returns is positive. Our estimation shows that the sample unrepresentativeness induces a downward bias in the estimate of the equity term structure slope. We present a regime-switching consumption-based asset-pricing model that matches the empirical findings.

JEL classification codes: D51, E21, G12, G13

Keywords: Asset pricing, business cycle phases, dividend strips, equity term structure, regime switching

---

*First version: October 2017. This paper was previously circulated under the title "Is the Term Structure of Equity Risk Premia Upward Sloping?" This research was supported by Rodney White Center and Jacob Levy Center. We thank a major financial institution for supplying us the data, and Mete Kilic for providing excellent research assistance. We also thank seminar participants at 2017 Macro-Finance Society Meeting in Chicago, Arizona State University, Duke University, European Central Bank, Goethe University, HEC-Montreal-McGill, London Business School, London School of Economics, Stockholm School of Economics, University of California Berkeley, and University of Michigan (Ann Arbor) for their comments. The views expressed herein are those of the authors and not necessarily those of Bank of Israel.

†Corresponding Author. Fuqua School of Business, Duke University, and NBER. 100 Fuqua Drive, Durham, NC 27708. ravi.bansal@duke.edu.

‡Stephen M. Ross School of Business, Univeristy of Michigan, Ann Arbor. 701 Tappan Ave, Ann Arbor, MI 48109. shanehm@umich.edu.

§Carey Business School, Johns Hopkins University. 100 International Drive, Baltimore, MD 21202. dongho.song@jhu.edu.

¶Wharton School, University of Pennsylvania, NBER, and Bank of Israel. 2300 Steinberg-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104. yarona@wharton.upenn.edu.

Appendix A. Estimating Regime-Switching Dividend Growth Dynamics

Appendix A.1. State space representation

Consider the state-space model

$$\begin{aligned} Y_t &= D(S_{t-n}) + Z\alpha_t, \\ \alpha_t &= T\alpha_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma), \end{aligned} \tag{A-1}$$

where

$$\begin{aligned} Y_t &= \begin{bmatrix} y_{1,t} \\ y_{2,t-n} \end{bmatrix}, \quad D(S_{t-n}) = \begin{bmatrix} \mu_1(S_{t-n}) \\ \mu_2(S_{t-n}) \end{bmatrix}, \quad Z = \begin{bmatrix} \frac{1}{n} & \dots & \frac{1}{n} & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}, \\ \alpha_t &= \begin{bmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{1,t-n+1} \\ \epsilon_{2,t-n} \end{bmatrix}, \quad T = \begin{bmatrix} \mathbf{0}_{1\times n} & 0 \\ \mathbf{I}_{(n-1)\times(n-1)} & \mathbf{0}_{(n-1)\times 1} \\ \mathbf{0}_{1\times n} & 0 \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{1,t-n+1} \\ \epsilon_{2,t-n} \end{bmatrix}. \end{aligned} \tag{A-2}$$

Assume that the discrete state can be $S_t \in \{1, ..., K\}$. Note that S_t and α_t are latent variables.

Appendix A.2. Evaluation of the likelihood function

See Chapter 5 of Kim and Nelson (1999) for detailed descriptions. Given $\alpha^i_{t-1|t-1}$, $P^i_{t-1|t-1}$, for $i, j, \in \{1, ..., K\}$

Forecasting

$$\begin{aligned}
\alpha^{(i,j)}_{t|t-1} &= T\alpha^{i}_{t-1|t-1} \\
P^{(i,j)}_{t|t-1} &= TP^{i}_{t-1|t-1}T' + \Sigma \\
e^{(i,j)}_{t|t-1} &= Y_t - D(S_{t-n} = j) - Z\alpha^{(i,j)}_{t|t-1} \\
F^{(i,j)}_{t|t-1} &= ZP^{(i,j)}_{t|t-1}Z'.
\end{aligned}$$

Updating

$$\begin{aligned}
\alpha^{(i,j)}_{t|t} &= \alpha^{(i,j)}_{t|t-1} + \left(P^{(i,j)}_{t|t-1}Z'\right)\left(F^{(i,j)}_{t|t-1}\right)^{-1} e^{(i,j)}_{t|t-1} \\
P^{(i,j)}_{t|t} &= P^{(i,j)}_{t|t-1} - \left(P^{(i,j)}_{t|t-1}Z'\right)\left(F^{(i,j)}_{t|t-1}\right)^{-1}\left(ZP^{(i,j)}_{t|t-1}\right).
\end{aligned}$$

Each iteration of the Kalman filter produces a K-fold increase in the number of cases to consider. It is necessary to introduce some approximations to make the above Kalman filter operable. The key is to collapse terms in the right way at the right time. Therefore, it remains to reduce the $K \times K$ posteriors $\alpha^{(i,j)}_{t|t}, P^{(i,j)}_{t|t}$ into K posteriors $\alpha^{j}_{t|t}, P^{j}_{t|t}$. Note that

$$\begin{aligned}
E\big(\alpha_t|S_{t-n} = j, Y_t\big) &= \frac{\sum_{i=1}^{K} Pr(S_{t-n-1} = i, S_{t-n} = j|Y_t)E\big(\alpha_t|S_{t-n} = j, S_{t-n-1} = i, Y_t\big)}{Pr(S_{t-n} = j|Y_t)} \\
&= \sum_{i=1}^{K} \Delta^{(i,j)}_t E\big(\alpha_t|S_{t-n-1} = i, S_{t-n} = j, Y_t\big), \\
&\quad \Delta^{(i,j)}_t = \frac{Pr(S_{t-n-1} = i, S_{t-n} = j|Y_t)}{Pr(S_{t-n} = j|Y_t)} \\
\alpha^{j}_{t|t} &= \sum_{i=1}^{K} \Delta^{(i,j)}_t \alpha^{(i,j)}_{t|t}.
\end{aligned}$$

The variance of α_t conditional on $S_{t-n} = j, Y_t$ could be derived in the following way:

$$\begin{aligned} P^j_{t|t} &= E\Big((\alpha_t - \alpha^j_{t|t})(\alpha_t - \alpha^j_{t|t})' | S_{t-n} = j, Y_t\Big) \\ &= \sum_{i=1}^{K} \Delta_t^{(i,j)} E\Big((\alpha_t - \alpha^j_{t|t})(\alpha_t - \alpha^j_{t|t})' | S_{t-n} = j, S_{t-n-1} = i, Y_t\Big) \\ &= \sum_{i=1}^{K} \Delta_t^{(i,j)} \Big[P^{(i,j)}_{t|t} + (\alpha^j_{t|t} - \alpha^{(i,j)}_{t|t})(\alpha^j_{t|t} - \alpha^{(i,j)}_{t|t})'\Big]. \end{aligned}$$

Merging

$$\begin{aligned} \alpha^j_{t|t} &= \frac{\sum_{i=1}^{K} Pr(S_{t-n-1} = i, S_{t-n} = j | Y_t)}{Pr(S_{t-n} = j | Y_t)} \Big(\alpha^{(i,j)}_{t|t}\Big) \\ P^j_{t|t} &= \frac{\sum_{i=1}^{K} Pr(S_{t-n-1} = i, S_{t-n} = j | Y_t)}{Pr(S_{t-n} = j | Y_t)} \Big(P^{(i,j)}_{t|t} + \big(\alpha^j_{t|t} - \alpha^{(i,j)}_{t|t}\big)\big(\alpha^j_{t|t} - \alpha^{(i,j)}_{t|t}\big)'\Big). \end{aligned}$$

Finally, the likelihood density of observation Y_t is given by

Likelihood

$$l(Y_t | Y_{1:t-1}) = \sum_{j=1}^{K} \sum_{i=1}^{K} f(Y_t | S_{t-n-1} = i, S_{t-n} = j, Y_{1:t-1}) Pr(S_{t-n-1} = i, S_{t-n} = j | Y_{t-1})$$

$$f(Y_t | S_{t-n-1} = i, S_{t-n} = j, Y_{1:t-1}) = (2\pi)^{-\frac{nv}{2}} \det\Big(F^{(i,j)}_{t|t-1}\Big)^{-\frac{1}{2}} \exp\Big[-\frac{1}{2}\big(e^{(i,j)}_{t|t-1}\big)'\Big(F^{(i,j)}_{t|t-1}\Big)^{-1}\big(e^{(i,j)}_{t|t-1}\big)\Big].$$

Appendix B. Solving the Long-Run Risks Model

This section provides approximate analytical solutions for the equilibrium asset prices.

Appendix B.1. Exogenous dynamics

The joint dynamics of consumption and dividend growth are

$$\begin{aligned}
\Delta c_{t+1} &= \mu(S_{t+1}) + x_{t+1} + \sigma_c \eta_{c,t+1}, \quad \eta_{c,t+1} \sim N(0,1), \\
\Delta d_{t+1} &= \bar{\mu} + \phi(\Delta c_{t+1} - \bar{\mu}) + \sigma_d \eta_{d,t+1}, \quad \eta_{d,t+1} \sim N(0,1), \\
x_{t+1} &= \rho x_t + \sigma_x(S_{t+1})\epsilon_{t+1},
\end{aligned} \tag{A-1}$$

where $\bar{\mu}$ is the unconditional mean of consumption growth and x_t is the persistent component of consumption growth. Agents observe the current regime, S_t, and make forecast of future regime, S_{t+1}, based on the transition matrix below

$$\mathbb{P} = \begin{bmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{bmatrix}. \tag{A-2}$$

Appendix B.2. Derivation of the approximate analytical solutions

The Euler equation for the economy is

$$1 = E_t \left[\exp\left(m_{t+1} + r_{c,t+1}\right)\right] \tag{A-3}$$

where the log stochastic discount factor is

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \tag{A-4}$$

$z_{c,t}$ is the log price to consumption ratio and $r_{c,t+1}$ is the log return on the consumption claim

$$r_{c,t+1} = \kappa_{0,c} + \kappa_{1,c} z_{c,t+1} - z_{c,t} + \Delta c_{t+1}. \tag{A-5}$$

Derivation of (A-3) follows Bansal and Zhou (2002), which make repeated use of the law of iterated expectations and log-linearization.

$$\begin{aligned} 1 &= E\Big(E\left[\exp\left(m_{t+1} + r_{c,t+1}\right) \mid S_{t+1}\right] \mid S_t \Big) \qquad \text{(A-6)} \\ &= \sum_{j=1}^{2} \mathbb{P}_{ij} E\Big(\exp\left(m_{t+1} + r_{c,t+1}\right) \mid S_{t+1} = j, S_t = i \Big) \\ 0 &= \sum_{j=1}^{2} \mathbb{P}_{ij} \Big(E\big[m_{t+1} + r_{c,t+1} \mid S_{t+1} = j, S_t = i\big] + \frac{1}{2} Var\big[m_{t+1} + r_{c,t+1} \mid S_{t+1} = j, S_t = i\big] \Big). \end{aligned}$$

The first line uses the law of iterated expectations; the second line uses the definition of Markov chain; and the third line relies on the log-normality assumption and applies log-linearization (i.e., $\exp(B) - 1 \approx B$).

Appendix B.3. Real consumption claim

Conjecture that the log price to consumption ratio follows

$$z_{c,t}(S_t) = A_{0,c}(S_t) + A_{1,c}(S_t) x_t. \tag{A-7}$$

From (A-1), (A-5), and (A-7), we can express the return on consumption claim by

$$\begin{aligned} r_{c,t+1} &= \kappa_{0,c} + \kappa_{1,c} A_{0,c}(S_{t+1}) - A_{0,c}(S_t) + \mu(S_{t+1}) \qquad \text{(A-8)} \\ &+ (\kappa_{1,c} A_{1,c}(S_{t+1})\rho - A_{1,c}(S_t) + \rho) x_t + (\kappa_{1,c} A_{1,c}(S_{t+1}) + 1)\sigma_x(S_{t+1})\epsilon_{t+1} + \sigma_c \eta_{c,t+1}. \end{aligned}$$

Using (A-8), we can re-express the log SDF (A-4) by

$$
\begin{aligned}
m_{t+1} &= \theta \ln \delta + (\theta - 1)\big(\kappa_{0,c} - A_{0,c}(S_t) + \kappa_{1,c} A_{0,c}(S_{t+1})\big) - \gamma \mu(S_{t+1}) \\
&+ \big((\theta - 1)(\kappa_{1,c} A_{1,c}(S_{t+1})\rho - A_{1,c}(S_t)) - \gamma\rho\big) x_t \\
&+ \big((\theta - 1)\kappa_{1,c} A_{1,c}(S_{t+1}) - \gamma\big)\sigma_x(S_{t+1})\epsilon_{t+1} - \gamma\sigma_c \eta_{c,t+1}.
\end{aligned} \tag{A-9}
$$

The solutions for $A_{0,c}$ and $A_{1,c}$ that describe the dynamics of the price-consumption ratio are determined from (A-6) which are

$$
\begin{bmatrix} A_{1,c}(1) \\ A_{1,c}(2) \end{bmatrix} = \begin{bmatrix} \mathbb{I}_2 - \rho\kappa_{1,c}\mathbb{P} \end{bmatrix}^{-1} \begin{bmatrix} (1 - \frac{1}{\psi})\rho \\ (1 - \frac{1}{\psi})\rho \end{bmatrix} \tag{A-10}
$$

$$
\begin{aligned}
\begin{bmatrix} A_{0,c}(1) \\ A_{0,c}(2) \end{bmatrix} &= \begin{bmatrix} \mathbb{I}_2 - \kappa_{1,c}\mathbb{P} \end{bmatrix}^{-1} \mathbb{P} \times \\
& \begin{bmatrix} \ln\delta + \kappa_{0,c} + (1 - \frac{1}{\psi})\mu(1) + \frac{\theta}{2}(1 - \frac{1}{\psi})^2\sigma_c^2 + \frac{\theta}{2}\big(\kappa_{1,c}A_{1,c}(1) + (1 - \frac{1}{\psi})\big)^2\sigma_x(1)^2 \\ \ln\delta + \kappa_{0,c} + (1 - \frac{1}{\psi})\mu(2) + \frac{\theta}{2}(1 - \frac{1}{\psi})^2\sigma_c^2 + \frac{\theta}{2}\big(\kappa_{1,c}A_{1,c}(2) + (1 - \frac{1}{\psi})\big)^2\sigma_x(2)^2 \end{bmatrix}.
\end{aligned} \tag{A-11}
$$

Appendix B.4. Linearization parameters

Let $\bar{p}_j = \frac{1-p_l}{2-p_l-p_j}$. The linearization parameters are determined endogenously by the following system of equations

$$
\begin{aligned}
\bar{z}_c &= \sum_{j=1}^{2} \bar{p}_j A_{0,c}(j) \\
\kappa_{1,c} &= \frac{\exp(\bar{z}_c)}{1 + \exp(\bar{z}_c)} \\
\kappa_{0,c} &= \log(1 + \exp(\bar{z}_c)) - \kappa_{1,c}\bar{z}_c.
\end{aligned}
$$

The solution is determined numerically by iteration until reaching a fixed point of $\bar{z}_c$.

Appendix B.5. Real bond prices

Conjecture that $b_{n,t}$ depends on the regime S_t and x_t,

$$b_{n,t} = b_{n,0}(S_t) + b_{n,1}(S_t)x_t. \tag{A-12}$$

Exploit the law of iterated expectations

$$b_{n,t} \quad = \quad \ln E_t\Big(E[\exp(m_{t+1} + b_{n-1,t+1})|S_{t+1}]\Big)$$

and log-linearization to solve for $b_{n,t}$

$$b_{n,t} \quad \approx \quad \sum_{j=1}^{2} \mathbb{P}_{ij}\Big(E[m_{t+1} + b_{n-1,t+1}|S_{t+1}] + \frac{1}{2}Var[m_{t+1} + b_{n-1,t+1}|S_{t+1}]\Big).$$

The solution to (A-12) is

$$\begin{aligned}
\begin{bmatrix} b_{n,1}(1) \\ b_{n,1}(2) \end{bmatrix} &= \mathbb{P}\begin{bmatrix} b_{n-1,1}(1)\rho + (\theta-1)\kappa_{1,c}A_{1,c}(1)\rho \\ b_{n-1,1}(2)\rho + (\theta-1)\kappa_{1,c}A_{1,c}(2)\rho \end{bmatrix} - \begin{bmatrix} (\theta-1)A_{1,c}(1) + \gamma\rho \\ (\theta-1)A_{1,c}(2) + \gamma\rho \end{bmatrix} \qquad \text{(A-13)} \\
\begin{bmatrix} b_{n,0}(1) \\ b_{n,0}(2) \end{bmatrix} &= \begin{bmatrix} \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(1) + \frac{\gamma^2}{2}\sigma_c^2 \\ \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(2) + \frac{\gamma^2}{2}\sigma_c^2 \end{bmatrix} \\
&+ \mathbb{P}\times\begin{bmatrix} b_{n-1,0}(1) + (\theta-1)\kappa_{1,c}A_{0,c}(1) - \gamma\mu(1) \\ b_{n-1,0}(2) + (\theta-1)\kappa_{1,c}A_{0,c}(2) - \gamma\mu(2) \end{bmatrix} \\
&+ \mathbb{P}\times\begin{bmatrix} \frac{1}{2}\big((\theta-1)\kappa_{1,c}A_{1,c}(1) - \gamma + b_{n-1,1}(1)\big)^2\sigma_x^2(1) \\ \frac{1}{2}\big((\theta-1)\kappa_{1,c}A_{1,c}(2) - \gamma + b_{n-1,1}(2)\big)^2\sigma_x^2(2) \end{bmatrix}
\end{aligned}$$

with the initial condition $b_{0,0}(i) = 0$ and $b_{0,1}(i) = 0$ for $i \in \{1,2\}$. The real yield of the maturity n-period bond is $y^r_{n,t} = -\frac{1}{n}b_{n,t}$.

Appendix B.6. Price to dividend ratio of zero coupon equity

Conjecture that the log price to dividend ratio of zero coupon equity $z_{n,t}$ depends on the regime S_t and persistent component x_t,

$$z_{n,t} = z_{n,0}(S_t) + z_{n,1}(S_t)x_t. \tag{A-14}$$

Exploit the law of iterated expectations

$$Z_{n,t} = E_t\left(E[M_{t+1} Z_{n-1,t+1} \frac{D_{t+1}}{D_t} | S_{t+1}] \right)$$

Take log

$$z_{n,t} = \ln E_t\left(E[\exp(m_{t+1} + z_{n-1,t+1} + \Delta d_{t+1}) | S_{t+1}] \right)$$

and log-linearization to solve for $z_{n,t}$

$$z_{n,t} \approx \sum_{j=1}^{2} \mathbb{P}_{ij}\left(E[m_{t+1} + z_{n-1,t+1} + \Delta d_{t+1} | S_{t+1}] + \frac{1}{2} Var[m_{t+1} + z_{n-1,t+1} + \Delta d_{t+1} | S_{t+1}] \right).$$

The solution is

$$\begin{bmatrix} z_{n,1}(1) \\ z_{n,1}(2) \end{bmatrix} = \mathbb{P} \begin{bmatrix} z_{n-1,1}(1)\rho + (\theta-1)\kappa_{1,c}A_{1,c}(1)\rho \\ z_{n-1,1}(2)\rho + (\theta-1)\kappa_{1,c}A_{1,c}(2)\rho \end{bmatrix} - \begin{bmatrix} (\theta-1)A_{1,c}(1) - (\phi-\gamma)\rho \\ (\theta-1)A_{1,c}(2) - (\phi-\gamma)\rho \end{bmatrix} \tag{A-15}$$

$$\begin{aligned} \begin{bmatrix} z_{n,0}(1) \\ z_{n,0}(2) \end{bmatrix} &= \begin{bmatrix} (1-\phi)\bar{\mu} + \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(1) + \frac{1}{2}(\phi-\gamma)^2\sigma_c^2 + \frac{1}{2}\sigma_d^2 \\ (1-\phi)\bar{\mu} + \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(2) + \frac{1}{2}(\phi-\gamma)^2\sigma_c^2 + \frac{1}{2}\sigma_d^2 \end{bmatrix} \\ &+ \mathbb{P} \times \begin{bmatrix} z_{n-1,0}(1) + (\theta-1)\kappa_{1,c}A_{0,c}(1) + (\phi-\gamma)\mu(1) \\ z_{n-1,0}(2) + (\theta-1)\kappa_{1,c}A_{0,c}(2) + (\phi-\gamma)\mu(2) \end{bmatrix} \\ &+ \mathbb{P} \times \begin{bmatrix} \frac{1}{2}\big((\theta-1)\kappa_{1,c}A_{1,c}(1) + (\phi-\gamma) + z_{n-1,1}(1)\big)^2\sigma_x^2(1) \\ \frac{1}{2}\big((\theta-1)\kappa_{1,c}A_{1,c}(2) + (\phi-\gamma) + z_{n-1,1}(2)\big)^2\sigma_x^2(2) \end{bmatrix} \end{aligned}$$

with the initial condition $z_{0,0}(i) = 0$ and $z_{0,1}(i) = 0$ for $i \in \{1, 2\}$.

Appendix B.7. m-holding-period and hold-to-maturity expected return

The price of zero coupon equity is $P_{n,t} = Z_{n,t}D_t$. Define the m-holding period return of the n-maturity equity is

$$R_{n,t+m} = \frac{Z_{n-m,t+m}}{Z_{n,t}} \frac{D_{t+m}}{D_t}. \tag{A-16}$$

The corresponding log expected return is defined by

$$E_t[r_{n,t+m}] = \frac{1}{m} E_t\big(z_{n-m,t+m} - z_{n,t} + \sum_{i=1}^{m} \Delta d_{t+i}\big) \tag{A-17}$$

To compute the excess return, we subtract the real rate of the same maturity

$$E_t[r_{n,t+m}] - y^r_{m,t}. \tag{A-18}$$

We consider two cases

- $m \neq n$: This is the m-holding-period expected excess return of the n-maturity equity.

$$\begin{aligned} E_t[g_{d,t+m}] &= \frac{1}{m} E_t\big(\sum_{i=1}^{m} \Delta d_{t+i}\big) \\ e_{n,m,t} &= \frac{1}{m} E_t\big(z_{n-m,t+m} - z_{n,t}\big) \\ E_t[r_{n,t+m}] &= e_{n,m,t} + E_t[g_{d,t+m}] \\ E_t[rx_{n,t+m}] &= E_t[r_{n,t+m}] - y^r_{m,t}. \end{aligned} \tag{A-19}$$

- $m = n$: This is the hold-to-maturity expected excess return of the n-maturity equity. Define

$$\begin{aligned} E_t[g_{d,t+n}] &= \frac{1}{n} E_t\big(\sum_{i=1}^{n} \Delta d_{t+i}\big) \\ e_{n,t} &= \frac{1}{n} E_t\big(-z_{n,t}\big) \\ E_t[r_{t+n}] &= e_{n,t} + E_t[g_{d,t+n}] \\ E_t[rx_{t+n}] &= E_t[r_{t+n}] - y^r_{n,t}. \end{aligned} \tag{A-20}$$

Appendix B.8. Computing moments

The cumulative sum of log dividend growth rates are

$$\begin{aligned} \sum_{i=1}^{n} \Delta d_{t+i} = n(1-\phi)\bar{\mu} + \phi\big(\mu(S_{t+1}) + ... + \mu(S_{t+n})\big) + \phi\rho\left(\frac{1-\rho^n}{1-\rho}\right)x_t \\ + \phi\left(\frac{1-\rho^n}{1-\rho}\right)\sigma_x(S_{t+1})\epsilon_{t+1} + ... + \phi\left(\frac{1-\rho}{1-\rho}\right)\sigma_x(S_{t+n})\epsilon_{t+n} \\ + \phi\sigma_c(\eta_{c,t+1} + ... + \eta_{c,t+n}) + \sigma_d(\eta_{d,t+1} + ... + \eta_{d,t+n}). \end{aligned} \tag{A-21}$$

For ease of exposition, we introduce the following notations

$$\boldsymbol{\mu} = [\mu(1), \mu(2)]', \quad \boldsymbol{\sigma^2_x} = [\sigma_x(1)^2, \sigma_x(2)^2]'.$$

The first two moments of the average log dividend growth rates for the case of $S_t = k$ are

$$\begin{aligned} E_t[g_{d,t+n}] &= \frac{1}{n}E_t[\sum_{i=1}^{n}\Delta d_{t+i}] = \frac{1}{n}\boldsymbol{\mu_G}(k) \\ V_t[g_{d,t+n}] &= \frac{1}{n^2}V_t[\sum_{i=1}^{n}\Delta d_{t+i}] = \frac{1}{n^2}\boldsymbol{\sigma^2_G}(k) \end{aligned} \tag{A-22}$$

where

$$\begin{aligned} \boldsymbol{\mu_G} &= \begin{bmatrix} n(1-\phi)\bar{\mu} + \phi\rho\left(\frac{1-\rho^n}{1-\rho}\right)x_t \\ n(1-\phi)\bar{\mu} + \phi\rho\left(\frac{1-\rho^n}{1-\rho}\right)x_t \end{bmatrix} + \sum_{j=1}^{n}\phi\mathbb{P}^j\boldsymbol{\mu} \\ \boldsymbol{\sigma^2_G} &\approx \begin{bmatrix} n(\phi^2\sigma_c^2 + \sigma_d^2) \\ n(\phi^2\sigma_c^2 + \sigma_d^2) \end{bmatrix} + \phi^2\sum_{j=1}^{n}\left(\frac{1-\rho^{n+1-j}}{1-\rho}\right)^2\mathbb{P}^j\boldsymbol{\sigma^2_x}. \end{aligned} \tag{A-23}$$

We acknowledge that the expression for $\boldsymbol{\sigma^2_G}$ is not exact because we are ignoring the variance component associated with uncertainty about $\mu(S_{t+j})$.

The above expressions allow us to calculate the Sharpe ratio

$$SR_{n,t} = \frac{e_{n,t} + E_t[g_{d,t+n}] - y^r_{n,t}}{\sqrt{V_t[g_{d,t+n}]}} \tag{A-24}$$

In the main text, we report the case of $x_t = 0$ for ease of illustration, e.g., $E_t g_{d,t+n}|_{x_t=0}$ and $V_t g_{d,t+n}|_{x_t=0}$.

Appendix B.9. Market return and equity premium

We derive the market return via Campbell-Shiller approximation

$$\begin{aligned} r_{d,t+1} \quad &= \quad \kappa_{0,d} + \kappa_{1,d} A_{0,d}(S_{t+1}) - A_{0,d}(S_t) + (1-\phi)\bar{\mu} + \phi\mu(S_{t+1}) \\ &\quad + \big(\phi\rho + \kappa_{1,d} A_{1,d}(S_{t+1})\rho - A_{1,d}(S_t)\big)x_t \\ &\quad + (\phi + \kappa_{1,d} A_{1,d}(S_{t+1}))\sigma_x(S_{t+1})\epsilon_{t+1} + \phi\sigma_c\eta_{c,t+1} + \sigma_d\eta_{d,t+1} \end{aligned} \tag{A-25}$$

where the log price-dividend ratio is given by

$$z_t = A_{0,d}(S_t) + A_{1,d}(S_t)x_t. \tag{A-26}$$

We solve for the coefficients

$$\begin{bmatrix} A_{1,d}(1) \\ A_{1,d}(2) \end{bmatrix} = \begin{bmatrix} \mathbb{I}_2 - \rho\kappa_{1,d}\mathbb{P} \end{bmatrix}^{-1} \left(\begin{bmatrix} (\phi-\gamma)\rho - (\theta-1)A_{1,c}(1) \\ (\phi-\gamma)\rho - (\theta-1)A_{1,c}(2) \end{bmatrix} + \mathbb{P} \begin{bmatrix} (\theta-1)\kappa_{1,c}A_{1,c}(1)\rho \\ (\theta-1)\kappa_{1,c}A_{1,c}(2)\rho \end{bmatrix} \right) \tag{A-27}$$

$$\begin{bmatrix} A_{0,d}(1) \\ A_{0,d}(2) \end{bmatrix} = \begin{bmatrix} \mathbb{I}_2 - \kappa_{1,d}\mathbb{P} \end{bmatrix}^{-1} \times$$

$$\left(\begin{bmatrix} (1-\phi)\bar{\mu} + \kappa_{0,d} + \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(1) + 0.5(\phi-\gamma)^2\sigma_c^2 + 0.5\sigma_d^2 \\ (1-\phi)\bar{\mu} + \kappa_{0,d} + \theta\ln\delta + (\theta-1)\kappa_{0,c} - (\theta-1)A_{0,c}(2) + 0.5(\phi-\gamma)^2\sigma_c^2 + 0.5\sigma_d^2 \end{bmatrix} \right.$$

$$\left. \mathbb{P} \begin{bmatrix} (\phi-\gamma)\mu(1) + (\theta-1)\kappa_{1,c}A_{0,c}(1) + 0.5((\theta-1)\kappa_{1,c}A_{1,c}(1) + \kappa_{1,d}A_{1,d}(1) + \phi-\gamma)^2\sigma_x(1)^2 \\ (\phi-\gamma)\mu(2) + (\theta-1)\kappa_{1,c}A_{0,c}(2) + 0.5((\theta-1)\kappa_{1,c}A_{1,c}(2) + \kappa_{1,d}A_{1,d}(2) + \phi-\gamma)^2\sigma_x(2)^2 \end{bmatrix} \right).$$

The market equity premium $E_t[r_{d,t+1}] - y^r_{1,t} + \frac{1}{2}V_t[r_{d,t+1}]$ is

$$-Cov_t(r_{d,t+1}, m_{t+1}) = \phi\gamma\sigma_c^2 + \mathbb{P} \times \begin{bmatrix} (\phi + \kappa_{1,d}A_{1,d}(1))((\theta - 1)\kappa_{1,c}A_{1,c}(1) - \gamma)\sigma_x(1)^2 \\ (\phi + \kappa_{1,d}A_{1,d}(2))((\theta - 1)\kappa_{1,c}A_{1,c}(2) - \gamma)\sigma_x(2)^2 \end{bmatrix}. \tag{A-28}$$

The conditional variance of the market return is

$$V_t[r_{d,t+1}] \approx \begin{bmatrix} \phi^2\sigma_c^2 + \sigma_d^2 \\ \phi^2\sigma_c^2 + \sigma_d^2 \end{bmatrix} + \mathbb{P} \times \begin{bmatrix} (\phi + \kappa_{1,d}A_{1,d}(1))^2\sigma_x(1)^2 \\ (\phi + \kappa_{1,d}A_{1,d}(2))^2\sigma_x(2)^2 \end{bmatrix}. \tag{A-29}$$

The market Sharpe ratio is

$$SR_t = \frac{E_t[r_{d,t+1}] - y^r_{1,t}}{\sqrt{V_t[r_{d,t+1}]}}. \tag{A-30}$$

Here, we are not accounting for $\frac{1}{2}V_t[r_{d,t+1}]$ in the numerator.

Appendix B.10. Calibration

Table D-1 provides the configuration of investors' preferences and the parameters that govern the dynamics of consumption and dividend growth rates. The model is calibrated on a monthly decision interval.

Appendix B.11. Model-implied moments

Based on the calibration with a severe-recession specification, the model-implied arithmetic equity premium is 6.6% and the risk-free rate is around 2.9%. The standard deviation of log dividend-price ratio and market return are 8.7% and 18.0%, respectively. The standard deviation of the 1-year and 10-year equity yields are 5.9% and 0.9%, respectively. The population conditional moments for the recession state are more extreme as the state becomes riskier.

Appendix B.12. Model-implied moments based on a calibration with a mild recession specification

We change three parameters associated with the recession state in Table D-1. Specifically, the new calibrated values are $\mu(2) = 0.0010$, $\sigma_x(2) = 0.0064$, $p_2 = 0.98$, respectively. The ergodic probability of the recession state is 14.9%. Put together, these calibrated values imply a milder but more frequent recession in the economy. With this calibration, the results are qualitatively similar to the severe recession calibration.

Appendix C. Forecasting Dividend Growth Rates

We expect that the dynamics of expected dividend growth rates, and consequently, expected returns, see (12), would be quite different conditional on recession and expansion. To show this, we develop a model for forecasting dividend growth rates both in-sample and out-of-sample, as discussed below.

Appendix C.1. VAR-based dividend forecasts

Let $x_{A,t}$ be a vector of monthly variables that predicts dividend growth. We consider an annual first order VAR dynamics for the predictor vector

$$x_{A,t+12} = \mu_A + \Gamma_A x_{A,t} + \varepsilon_{A,t+12}. \tag{A-1}$$

This is because we are interested in the annual horizon forecasts. There are two ways of estimating the coefficients in (A-1): the first is via the direct projection method and the second is estimate a monthly first order VAR model[1]

$$x_{A,t+1} = \mu_m + \Gamma_m x_{A,t} + \varepsilon_{m,t+1} \tag{A-2}$$

and obtain

$$\mu_A \equiv \left(\sum_{i=0}^{11} \Gamma_m^i \right) \mu, \quad \Gamma_A \equiv \Gamma_m^{12}, \quad \varepsilon_{A,t+12} \equiv \sum_{i=0}^{11} \Gamma_m^{12-i} \varepsilon_{m,t+i}.$$

Regressing dividend growth on lagged predictor vector gives the estimates for ψ_0 and ψ_1

$$g_{d,t+12} = \psi_0 + \psi_1 x_{A,t} + \varepsilon_{d,t+12}. \tag{A-3}$$

To recap, $g_{d,t+12} = \ln(\frac{D_{t+12}}{D_t})$ where D_t is the 12-month trailing sum dividends.

[1]Note that Binsbergen, Hueskes, Koijen, and Vrugt (2013) follow the second approach.

For ease of exposition, we stack (A-1) and (A-3) together and express them in an annual first order VAR model as

$$\begin{bmatrix} x_{A,t+12} \\ g_{d,t+12} \end{bmatrix} = \begin{bmatrix} \mu_A \\ \psi_0 \end{bmatrix} + \begin{bmatrix} \Gamma_A & 0 \\ \psi_1 & 0 \end{bmatrix} \begin{bmatrix} x_{A,t} \\ g_{d,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{A,t+12} \\ \varepsilon_{d,t+12} \end{bmatrix}. \tag{A-4}$$

From (A-4), we derive the conditional expectation of the annual dividend growth n years ahead as

$$E_t[g_{d,t+12n}] = \psi_0 + \psi_1\left(\left[\sum_{i=0}^{n-2}\Gamma_A^i\right]\mu_A + \Gamma_A^{(n-1)}x_t\right). \tag{A-5}$$

The $12n$-month-ahead dividend growth shocks and their cumulative shocks are

$$\begin{aligned} g_{d,t+12n} - E_t[g_{d,t+12n}] &= \psi_1\left(\sum_{i=0}^{n-2}\Gamma_A^i\varepsilon_{A,t+12(n-1-i)}\right) + \varepsilon_{d,t+12n} \qquad \text{(A-6)} \\ \sum_{n=1}^{m}\Big(g_{d,t+12n} - E_t(g_{d,t+12n})\Big) &= \psi_1\left(\sum_{j=0}^{m-2}(I-\Gamma_A)^{-1}(I-\Gamma_A^{m-1-j})\varepsilon_{t+12(j+1)}\right) + \sum_{n=1}^{m}\varepsilon_{d,t+12n}. \end{aligned}$$

It is straightforward to compute $V_t[g_{d,t+12n}]$ and $V_t\big[\sum_{n=1}^{m} g_{d,t+12n}\big]$ from (A-6).

Appendix C.2. Bayesian inference

The sample in which we have equity yields is quite short. However, data on dividend growth rates are available much before. We have seen from Figure A-1 that we can potentially rely on the historical data to learn about the future dividend growth dynamics to the extent that dynamics have not changed substantially over time. In this section, we formally show how one could optimally use prior information (extracted from historical data) and improve forecasts.

Posterior. The first-order vector autoregression (A-4) can be always re-written as

$$y_t = \Phi x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma) \tag{A-7}$$

where $y_t = \left[x'_{A,t}, g_{d,t}\right]'$ and $x_t = \left[1, x'_{A,t-12}\right]'$. Define $Y = [y_{13}, ..., y_T]'$, $X = [x_{13}, ..., x_T]'$, and $\varepsilon = [\varepsilon_{13}, ..., \varepsilon_T]'$. Taking the initial 12 observations as given, if the prior is

$$\Phi|\Sigma \sim MN\big(\underline{\Phi}, \Sigma \otimes (\underline{V}_\Phi \xi)\big), \quad \Sigma \sim IW(\Psi, d) \tag{A-8}$$

then the posterior can be expressed as

$$\Phi|\Sigma \sim MN\big(\overline{\Phi}, \Sigma \otimes \overline{V}_\Phi\big), \quad \overline{\Phi} = \left(X'X + (\underline{V}_\Phi \xi)^{-1}\right)^{-1}\left(X'Y + (\underline{V}_\Phi \xi)^{-1}\underline{\Phi}\right) \tag{A-9}$$

because of the conjugacy.[2] Here, ξ is a scalar parameter controlling the tightness of the prior information.

The elicitation of prior. Suppose that we can divide the sample into the pre-sample, estimation sample, and prediction sample. We set the prior mean $\underline{\Phi}$ equal to the pre-sample OLS estimate, from the 1980-2004 sample for the U.S. and the 1995-2004 sample for Europe and Japan. Here, prior becomes more informative when $\xi \to 0$. In the limit, posterior equals the pre-sample OLS estimate, i.e., prior. In contrast, when $\xi = \infty$, then it is easy to see that $\overline{\Phi} = \hat{\Phi} = (X'X)^{-1}(X'Y)$, i.e., an OLS estimate from the estimation sample. In this case, prior does not play any role. We can optimize the scaling parameter by choosing the value that maximizes the marginal likelihood function, $\hat{\xi} = \text{argmax } p(Y|\xi)$. The closed form of the marginal likelihood function is available in the appendix. We refer to Giannone, Lenza, and Primiceri (2015) for a detailed discussion.

[2]Since we are mainly interested in the conditional expectation, we omit the expression for Σ. The readers are referred to Giannone, Lenza, and Primiceri (2015).

Appendix C.3. Model selection and estimation

The VAR expression in (A-4) can describe the approach of Binsbergen, Hueskes, Koijen, and Vrugt (2013) where the predictor vector x_A comprises the 2-year and 5-year forward equity yields. We refer to this three-variable (dividend growth plus two predictors) VAR, identical to that of BHKV, as the Short Sample Predictor (SSP) approach for simplicity. We propose a different three-variable VAR where the predictor vector x_A comprises the 5y-1y nominal bond yield spread and dividend to earnings ratio, which is referred to as the Long Sample Predictor (LSP) approach. In addition to improved forecast accuracy, the LSP will allow us to conduct out-of-sample forecasting because these predictors have a longer history than equity yields.

Ideally, we would like to conduct both in-sample and out-of-sample forecast exercises for the SSP and LSP approaches. Unfortunately, we cannot conduct out-of-sample forecast exercise for the SSP approach since their predictor variables, the 2-year and 5-year forward equity yields, are only available from 2004:M12 according to our data vendor. The VAR coefficients in the SSP approach cannot be recursively estimated unless the prediction sample is substantially shortened. This is not ideal given the data availability. The reason we desire to use recursive updating and out-of-sample forecasting using the LSP is to establish that our results hold in real-time forecasts. This study is the first to show the term structure of expected growth and returns in real time.

To minimize confusion, we define two estimation strategies. The in-sample estimation is carried out with data from 2004:M12 to 2017:M2 using the maximum available data. We can formally conduct model selection and compare the in-sample forecasting performance of the SSP and LSP approaches. When we generate in-sample forecasts, we allow for look-ahead bias by including all the data in the estimation at once. Here, we briefly describe the model selection result. The forecast results are discussed shortly. We choose the LSP approach over the SSP approach based on the model selection via the marginal likelihood maximization. For U.S., the log marginal likelihood values are 430 versus 374; for Europe, they are 704.4 versus 58.5; for Japan, 596 versus 135, all in favor of LSP over SSP approach.

To define the out-of-sample estimation period, we first set the prediction sample to 2005:M1 to 2013:M2. The initial out-of-sample estimation starts from 2001:M1 to 2004:M12. When we move the forecast origin from 2005:M1 to 2013:M2, the posterior VAR coefficients are also updated as we recursively increase the sample. In doing so, we optimize the scaling parameter ξ that controls the tightness of the prior.

We use data from 1979:M12 to 2000:M12 (U.S.) and from 1994:M12 to 2000:M12 (Europe and Japan) to elicit prior information. It is possible for the LSP approach because the 5y-1y nominal bond yield spread and dividend to earnings ratio are available. Since elicitation of prior information is not possible for the SSP approach, we set the scaling parameter to $\xi = \infty$ so that (whichever specified) prior does not play any role and the posterior mean is identical to the OLS estimate. Thus, we can rely on the same expression (A-9) to generate posterior forecasts for both the SSP and LSP approaches.

Appendix C.4. Forecast results

The dividend growth rate forecasts are generated up to 5-year-out to maximize the data availability.[3] Table D-2 summarizes the root mean squared errors (RMSEs) of the dividend growth rate forecasts for the three markets. Let us focus on Panel A: U.S. and compare the RMSEs from the out-of-sample LSP (o.o.s.) with those from the in-sample SSP approach. The results are surprising given that the LSP approach (o.o.s.) is at a clear informational disadvantage compared with the SSP approach. Except at the 1-year horizon, we find that the RMSEs for the respective horizons are much smaller than those from the SSP approach. What is interesting is the magnitude of the RMSEs, whether they are based on in-sample or

[3]One could generate up to 7-year-out horizon which results in shortening the prediction sample to 2005:M1-2011:M2 instead of 2005:M1-2013:M2.

out-of-sample forecasts, are similar for the LSP approach. This evidence strongly suggests the superior forecast performance of our VAR approach and the usefulness of extracting information embedded in the historical data in the form of priors.

The RMSEs from the out-of-sample forecasts are uniformly larger (roughly by a factor of two) than those of the in-sample SSP forecasts for Europe and Japan. Again, the large RMSE and apparent changes in dividend dynamics relative to the pre-sample suggest that these dividend growth events were unexpected. For in-sample forecasts, our model continues to produce superior estimates to those in Binsbergen, Hueskes, Koijen, and Vrugt (2013) and dominates in marginal likelihood.

Once the expected dividend growth rates are generated, given the forward equity yields and real rates, we are able to compute the expected return (see (12)), excess return (see (13)), and hold to maturity Sharpe Ratio (see (15)) as well. We construct the real rate proxy by subtracting the average inflation from the nominal yields.[4] To be conservative and to save space, we only show the out-of-sample forecast results for the U.S. market. We provide the in-sample forecast results for the three markets in the appendix in Tables D-5, D-6, and D-7.

Table D-3 provides the annualized average expected dividend growth rates $E_t[g_{t+n}]$ ("Exp. growth"); the expected discount rate $E_t[r_{t+n}]$ ("Exp. return"); the expected excess return $E_t rx_{t+n}$ ("Premium"); the Sharpe ratio $SR_{n,t}$ ("Sharpe ratio"); the forward equity yields $e^f_{n,t}$ ("STRIPS"); and the nominal bond yields $y_{n,t}$ ("YLD"). We provide the corresponding averages of the entire prediction sample and the averages conditional on whether the forward equity yield spread between 5-year and 1-year is positive or negative. This is

[4]Given the relatively small variation of inflation rates, especially relative to the large movements in real growth and discount rates, by horizon within the recession and expansion subsamples, it is highly unlikely that compensation for inflation risk substantively has any bearing on our measure of risk premia. For U.S., we later replace with the Treasury Inflation-Protected Securities (TIPS) to confirm the robustness of our results.

because we believe that the negative spread of forward equity yields between 5-year and 1-year closely tracks the recession dates.[5] Remember that these are the averages of real time out-of-sample forecasts. We provide the evidence for the remaining regions in the appendix. We refrain from using the recession indicators to forecast as they are determined ex post. In contrast, the equity yields and LSP predictors are available to investors in real time.

We summarize the main findings as follows. The slope of the expected dividend growth is negative (positive) during expansions (recessions). The slopes of the expected return, excess return, and sharpe ratio are positive (negative) during expansion (recession). The slopes of the entire period averages of expected return, expected excess return, and sharpe ratio are positive. Both the in and out-of-sample hold to maturity Sharpe Ratios are either upward-sloping (U.S. and Japan) or flat (Europe in-sample), however these statistics have large standard errors relative to risk premium estimates.

To check the statistical significance of the findings, Table D-4 provides the 90% credible intervals associated with the selective forecasts: expected return and growth. It is interesting to observe that all slopes of the conditional moments are statistically significantly different from zero at the 90% confidence level. The findings are largely robust to the out-of-sample and in-sample forecast results. The in-sample forecast results based on the SSP approach deliver qualitatively similar message, which is provided in Table D-8. Finally, we note that while the estimates of hold to maturity Sharpe Ratios are substantially noisier, the Sharpe Ratio slope is significantly positive in the U.S., where the estimates are best and the recession frequency is in line with the long run mean, at 0.39 for 5y-1y in-sample and 0.51 out-of-sample.

We find that the sign of the conditional slopes for the European and Japanese markets

[5] Figure A-3 plots the two series, which appear to be highly correlated. The correlation between the equity yield spread and recession indicator is around 65% for the U.S. and Europe and 40% for Japan, respectively. The seemingly low correlation is because we are computing correlation with an indicator variable.

are broadly consistent with the U.S. market with expected returns sloping downward in recession and upward in expansions. Caution is required in interpreting the sample average as an unconditional mean because the balance of recessions in the short sample is not representative for these regions. If the sample overrepresents recessions, as is the case in Europe and Japan, the behavior of the sample average expected growth and dividend discount rate slopes will be biased towards their recession means.

Appendix C.5. Bayesian linear regression

Without loss of generality, we can express any linear dynamics by

$$y_t = \Phi x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma). \tag{A-10}$$

For ease of exposition, define $Y = [y_p, ..., y_T]'$, $X = [x_p, ..., x_T]'$, and $\varepsilon = [\varepsilon_p, ..., \varepsilon_T]'$. Assume that the initial p observations are available. Because of the conjugacy if the prior is

$$\Phi|\Sigma \sim MN\big(\underline{\Phi}, \Sigma \otimes (\underline{V}_\Phi \xi)\big), \quad \Sigma \sim IW(\Psi, d) \tag{A-11}$$

then the posterior can be expressed as $\Phi|\Sigma \sim MN\big(\overline{\Phi}, \Sigma \otimes \overline{V}_\Phi\big)$ where

$$\overline{\Phi} \;=\; \Big(X'X + (\underline{V}_\Phi \xi)^{-1}\Big)^{-1}\Big(X'Y + (\underline{V}_\Phi \xi)^{-1}\underline{\Phi}\Big), \quad \overline{V}_\Phi = \Big(X'X + (\underline{V}_\Phi \xi)^{-1}\Big)^{-1}.$$

We follow the exposition in Giannone, Lenza, and Primiceri (2015). ξ is a scalar parameter controlling the tightness of the prior information. For instance, prior becomes more informative when $\xi \to 0$. In contrast, when $\xi = \infty$, then it is easy to see that $\overline{\Phi} = \hat{\Phi}$, i.e., an OLS estimate. We can choose ξ that maximizes the marginal likelihood function (A-12), which is available in closed form

$$p(Y|\xi) = \left(\frac{1}{\pi}\right)^{\frac{n(T-p)}{2}} \frac{\Gamma_n(\frac{T-p+d}{2})}{\Gamma_n(\frac{d}{2})} |\underline{V}_\Phi \xi|^{-\frac{n}{2}} |\Psi|^{\frac{d}{2}} \left| X'X + (\underline{V}_\Phi \xi)^{-1} \right|^{-\frac{n}{2}} \quad \text{(A-12)}$$
$$\left| \Psi + \hat{\varepsilon}'\hat{\varepsilon} + (\hat{\Phi} - \underline{\Phi})'(\underline{V}_\Phi \xi)^{-1}(\hat{\Phi} - \underline{\Phi}) \right|^{-\frac{T-p+d}{2}}.$$

We refer to Giannone, Lenza, and Primiceri (2015) for a detailed description.

Appendix D. Tables

Table D-1. Calibration

The steady state probabilities for the expansion and recession states are $(1-p_2)/(2-p_1-p_2) = 0.9605$ and $(1-p_1)/(2-p_1-p_2) = 0.0395$, respectively. The steady state consumption growth mean is $\bar{\mu} = (1-p_2)/(2-p_1-p_2)\mu(1) + (1-p_1)/(2-p_1-p_2)\mu(2) = 0.0019$.

Parameters					
$\mu(1)$	0.0020	ρ	0.60	δ	0.998
$\mu(2)$	-0.0020	$\sigma_x(1)$	0.0033	ψ	1.5
σ_c	0.0063	$\sigma_x(2)$	0.0142	γ	10
ϕ	4.0	p_1	0.9965		
σ_d	0.0224	p_2	0.915		

Table D-2. Root mean squared errors for the dividend growth rate forecasts

For the in-sample forecasts, we use data from 1979:M12 to 2004:M12 (U.S.) and from 1994:M12 to 2004:M12 (Europe and Japan) to elicit prior information. The estimation sample is from 2005:M1 to 2017:M2. Our prediction sample is from 2005:M1 to 2013:M2. For the out-of-sample forecasts, we use data from 1979:M12 to 2000:M12 (U.S.) and from 1994:M12 to 2000:M12 (Europe and Japan) to elicit prior information. The initial estimation sample is from 2001:M1 to 2004:M12. We increase the estimation sample recursively as we move the forecast origin. The root mean squared errors based on the out-of-sample forecasts are indicated with o.o.s.

	Panel A: U.S.			Panel B: Europe			Panel C: Japan		
	SSP	LSP	LSP	SSP	LSP	LSP	SSP	LSP	LSP
n			o.o.s.			o.o.s.			o.o.s.
1y	8.41	9.82	10.10	8.46	7.59	20.14	11.62	12.79	18.32
2y	9.51	8.60	9.43	10.24	7.10	15.73	9.91	8.70	14.98
3y	9.19	7.03	8.10	9.11	6.38	12.01	7.96	6.78	11.80
4y	8.20	5.87	6.89	6.32	4.26	10.05	5.68	5.22	9.00
5y	6.73	4.68	5.75	4.98	3.07	9.33	4.44	3.93	8.26

Table D-3. Out-of-sample forecasts of dividend growth, prices, and returns: U.S.

We provide the annualized average expected dividend growth rates $E_t[g_{t+n}]$ ("Exp. growth"); the expected discount rate $E_t[r_{t+n}]$ ("Exp. return"), computed as in (12): $E_t[r_{t+n}] = e_{n,t} + E_t[g_{d,t+n}]$; the expected excess return $E_t rx_{t+n}$ ("Premium"), computed as in (13): $E_t[rx_{t+n}] = E_t[r_{t+n}] - y^r_{n,t}$; the Sharpe ratio $SR_{n,t}$ ("Sharpe ratio"), computed as in (15): $SR_{n,t} = \frac{E_t[rx_{t+n}]}{\sqrt{V_t[g_{d,t+n}]}}$; the forward equity yields $e^f_{n,t}$ ("STRIPS"); and the nominal bond yields $y_{n,t}$ ("YLD"). Results are based on the LSP approach, a 3-variable VAR approach that includes 5y-1y nominal bond yield spread, asset dividend to earnings ratio, and dividend growth. The initial estimation sample for the LSP approach is from 2001:M1 to 2004:M12. We increase the estimation sample recursively as we move the forecast origin, i.e., 2005:M1 to 2013:M2. The sample average of inflation rates is around 2%.

	Exp. growth	Exp. return	Premium	Sharpe ratio	STRIPS	YLD
			Entire period			
1y	2.93	0.39	0.35	0.08	-4.58	2.04
2y	2.61	0.82	0.65	0.23	-3.96	2.17
3y	2.59	1.30	0.93	0.38	-3.66	2.37
4y	2.62	1.75	1.16	0.50	-3.46	2.59
5y	2.65	2.22	1.39	0.65	-3.26	2.83
5y-1y	-0.28	1.83	1.04	0.57	1.32	0.79
			Expansion period			
1y	5.03	-2.99	-3.28	-0.96	-10.31	2.29
2y	3.87	-1.38	-1.77	-0.61	-7.64	2.39
3y	3.40	-0.20	-0.73	-0.29	-6.14	2.53
4y	3.19	0.53	-0.18	-0.09	-5.37	2.72
5y	3.07	1.16	0.25	0.10	-4.83	2.92
5y-1y	-1.96	4.14	3.52	1.06	5.48	0.62
			Recession period			
1y	-3.45	10.65	11.39	3.24	12.84	1.26
2y	-1.22	7.51	8.00	2.78	7.22	1.52
3y	0.10	5.86	5.99	2.38	3.89	1.87
4y	0.88	5.45	5.25	2.32	2.36	2.20
5y	1.36	5.43	4.86	2.34	1.49	2.58
5y-1y	4.81	-5.22	-6.54	-0.90	-11.34	1.32

Table D-4. The spread between 5-year and 1-year forecasts

We provide the results based on the in-sample and out-of-sample forecasts for $E_t[g_{d,t+5} - g_{d,t+1}]$ and $E_t[r_{t+5} - r_{t+1}]$ computed as in (12): $E_t[r_{t+n}] = e_{n,t} + E_t[g_{d,t+n}]$. For the in-sample forecasts, we use data from 1979:M12 to 2004:M12 (U.S.) and from 1994:M12 to 2004:M12 (Europe and Japan) to elicit prior information. The estimation sample is from 2005:M1 to 2017:M2. Our prediction sample is from 2005:M1 to 2013:M2. For the out-of-sample forecasts, we use data from 1979:M12 to 2000:M12 (U.S.) and from 1994:M12 to 2000:M12 (Europe and Japan) to elicit prior information. The initial estimation sample is from 2001:M1 to 2004:M12. We increase the estimation sample recursively as we move the forecast origin. We use $*$ to indicate the statistical significance at the 90% confidence level.

	In-sample						Out-of-sample					
	Exp. return			Exp. growth			Exp. return			Exp. growth		
	50%	[5%	95%]	50%	[5%	95%]	50%	[5%	95%]	50%	[5%	95%]
						Entire period						
U.S.	2.52*	[1.13,	3.85]	0.40	[-0.98,	1.73]	1.83*	[1.40,	2.25]	-0.28	[-0.72,	0.14]
Europe	-0.28	[-2.68,	2.25]	-0.45	[-2.85,	2.07]	2.21	[-0.32,	4.96]	2.21	[-0.30,	4.99]
Japan	-3.07*	[-5.27,	-0.90]	-1.06	[-3.26,	1.11]	-2.76*	[-5.54,	-0.04]	-0.62	[-3.41,	2.18]
						Expansion period						
U.S.	4.18*	[3.01,	5.32]	-1.91*	[-3.09,	-0.78]	4.14*	[3.83,	4.47]	-1.96*	[-2.28,	-1.63]
Europe	0.04	[-2.17,	2.39]	-7.23*	[-9.44,	-4.88]	9.85*	[7.23,	12.75]	2.85	[-0.47,	5.75]
Japan	1.19	[-0.85,	3.13]	-2.34*	[-4.38,	-0.44]	1.68	[-0.90,	4.23]	-1.92	[-4.49,	0.63]
						Recession period						
U.S.	-2.55*	[-4.58,	-0.64]	7.48*	[5.45,	9.38]	-5.22*	[-6.00,	-4.51]	4.81*	[4.03,	5.51]
Europe	-0.68	[-3.34,	2.06]	8.44*	[5.78,	11.19]	-7.79*	[-10.21,	-5.25]	1.37	[-1.05,	3.91]
Japan	-7.79*	[-10.18,	-5.37]	0.36	[-2.03,	2.78]	-7.48*	[-10.49,	-4.41]	0.75	[-2.25,	3.83]

Table D-5. The expected dividend growth rates and expected excess returns: US

Tables are generated from models that are estimated from 2005:M1 to 2017:M2 sample. We show two approaches: (1) SSP and (2) LSP, i.e., a 3-variable VAR approach that includes the 5y-1y nominal bond spread, asset dividend to earnings ratio, and dividend growth.

	RMSE		Premium		Exp. return		Sharpe Ratio		Exp. growth		STRIPS	YLD	INF
horizon	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP			
Entire period													
1y	8.41	9.82	1.43	1.22	1.46	1.26	0.19	0.14	4.01	3.80	-4.58	2.04	2.00
2y	9.51	8.60	2.22	1.95	2.39	2.13	0.37	0.24	4.18	3.91	-3.96	2.17	2.00
3y	9.19	7.03	2.65	2.39	3.02	2.76	0.48	0.34	4.31	4.05	-3.66	2.37	2.00
4y	8.20	5.87	2.94	2.69	3.53	3.28	0.56	0.44	4.40	4.14	-3.46	2.59	2.00
5y	6.73	4.68	3.21	2.95	4.04	3.78	0.63	0.54	4.47	4.21	-3.26	2.83	2.00
5y-1y	-	-	1.78	1.72	2.58	2.52	0.44	0.40	0.46	0.41	1.32	0.79	-
Positive strips spread													
1y	5.90	5.24	-0.60	-1.86	-0.30	-1.57	-0.08	-0.21	7.71	6.44	-10.31	2.29	2.00
2y	8.41	7.01	1.58	-0.58	1.97	-0.20	0.27	-0.07	7.21	5.05	-7.64	2.39	2.00
3y	8.65	6.72	2.69	0.53	3.23	1.06	0.49	0.08	6.83	4.66	-6.14	2.53	2.00
4y	7.55	5.90	3.16	1.19	3.88	1.91	0.60	0.20	6.53	4.56	-5.37	2.72	2.00
5y	5.82	4.66	3.47	1.71	4.39	2.62	0.68	0.31	6.30	4.53	-4.83	2.92	2.00
5y-1y	-	-	4.07	3.57	4.69	4.18	0.76	0.52	-1.41	-1.91	5.48	0.62	-
Negative strips spread													
1y	13.33	17.41	7.59	10.61	6.84	9.87	1.01	1.21	-7.25	-4.23	12.84	1.26	2.00
2y	12.12	12.16	4.17	9.67	3.69	9.19	0.70	1.19	-5.05	0.45	7.22	1.52	2.00
3y	10.51	7.84	2.52	8.06	2.39	7.93	0.46	1.15	-3.37	2.17	3.89	1.87	2.00
4y	9.77	5.71	2.28	7.24	2.48	7.44	0.43	1.18	-2.08	2.88	2.36	2.20	2.00
5y	8.85	4.66	2.41	6.73	2.99	7.31	0.47	1.20	-1.08	3.24	1.49	2.58	2.00
5y-1y	-	-	-5.17	-3.87	-3.85	-2.55	-0.54	-0.01	6.17	7.48	-11.34	1.32	-

Table D-6. The expected dividend growth rates and the expected excess returns: Europe

Tables are generated from models that are estimated from 2005:M1 to 2017:M2 sample. We show two approaches: (1) SSP and (2) LSP, i.e., a 3-variable VAR approach that includes the 5y-1y nominal bond spread, asset dividend to earnings ratio, and dividend growth.

	RMSE		Premium		Exp. return		Sharpe Ratio		Exp. growth		STRIPS	YLD	INF
horizon	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP			
Entire period													
1y	8.46	7.59	3.70	2.76	3.56	2.62	0.47	0.38	-0.53	-1.47	2.12	1.96	2.10
2y	10.24	7.10	5.08	4.30	5.07	4.30	0.68	0.63	-0.80	-1.58	3.78	2.09	2.10
3y	9.11	6.38	4.16	3.33	4.33	3.50	0.55	0.51	-0.89	-1.72	2.95	2.27	2.10
4y	6.32	4.26	3.37	2.46	3.76	2.84	0.46	0.41	-0.90	-1.81	2.17	2.49	2.10
5y	4.98	3.07	2.83	1.83	3.37	2.37	0.40	0.33	-0.88	-1.88	1.61	2.64	2.10
5y-1y	-	-	-0.87	-0.93	-0.18	-0.24	-0.08	-0.05	-0.36	-0.42	-0.51	0.68	-
Positive strips spread													
1y	6.39	7.61	2.01	0.85	2.46	1.30	0.26	0.12	8.55	7.39	-8.64	2.55	2.10
2y	10.34	8.14	3.05	1.06	3.58	1.59	0.41	0.16	5.90	3.91	-4.94	2.63	2.10
3y	9.99	7.43	3.19	0.87	3.85	1.53	0.42	0.13	4.23	1.91	-3.14	2.77	2.10
4y	7.00	4.60	2.95	0.60	3.78	1.44	0.40	0.10	3.14	0.80	-2.29	2.93	2.10
5y	5.24	3.08	2.63	0.38	3.60	1.35	0.37	0.07	2.42	0.17	-1.89	3.06	2.10
5y-1y	-	-	0.62	-0.47	1.14	0.05	0.11	-0.05	-6.13	-7.22	6.75	0.52	-
Negative strips spread													
1y	10.49	7.48	5.90	5.25	4.99	4.34	0.76	0.72	-12.42	-13.07	16.22	1.19	2.10
2y	9.99	5.40	7.73	8.54	7.02	7.83	1.03	1.26	-9.58	-8.77	15.21	1.39	2.10
3y	7.68	4.61	5.43	6.56	4.95	6.08	0.72	1.01	-7.59	-6.46	10.92	1.62	2.10
4y	5.23	3.76	3.92	4.88	3.73	4.69	0.53	0.81	-6.20	-5.24	8.02	1.91	2.10
5y	4.56	3.05	3.08	3.72	3.08	3.72	0.43	0.67	-5.21	-4.57	6.19	2.10	2.10
5y-1y	-	-	-2.81	-1.53	-1.91	-0.62	-0.32	-0.05	7.21	8.50	-10.03	0.90	-

Table D-7. The expected dividend growth rates and the expected excess returns: Japan

Tables are generated from models that are estimated from 2005:M1 to 2017:M2 sample. We show two approaches: (1) SSP and (2) LSP, i.e., a 3-variable VAR approach that includes the 5y-1y nominal bond spread, asset dividend to earnings ratio, and dividend growth.

	RMSE		Premium		Exp. return		Sharpe Ratio		Exp. growth		STRIPS	YLD	INF
horizon	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP	SSP	LSP			
Entire period													
1y	11.62	12.79	8.86	11.99	8.91	12.04	0.73	0.87	6.45	9.58	2.18	0.28	0.23
2y	9.91	8.70	8.88	11.25	9.02	11.39	0.94	1.27	7.07	9.44	1.57	0.38	0.23
3y	7.96	6.78	8.28	9.80	8.53	10.05	0.97	1.34	7.55	9.08	0.49	0.49	0.23
4y	5.68	5.22	8.11	8.97	8.50	9.35	1.01	1.38	7.93	8.78	-0.04	0.62	0.23
5y	4.44	3.93	8.17	8.46	8.68	8.97	1.06	1.42	8.22	8.51	-0.28	0.74	0.23
5y-1y	-	-	-0.69	-3.53	-0.24	-3.08	0.33	0.55	1.77	-1.07	-2.46	0.46	-
Positive strips spread													
1y	12.60	13.39	5.43	5.07	5.49	5.12	0.44	0.37	11.51	11.14	-6.31	0.29	0.23
2y	8.47	8.33	6.54	5.30	6.71	5.47	0.70	0.60	11.25	10.01	-4.94	0.40	0.23
3y	7.65	7.02	7.08	5.55	7.36	5.83	0.83	0.76	11.05	9.52	-4.20	0.51	0.23
4y	5.81	4.55	7.43	5.63	7.84	6.04	0.92	0.87	10.90	9.10	-3.70	0.64	0.23
5y	4.37	2.61	7.77	5.76	8.30	6.29	1.00	0.97	10.79	8.78	-3.25	0.76	0.23
5y-1y	-	-	2.34	0.70	2.81	1.16	0.56	0.60	-0.72	-2.37	3.06	0.47	-
Negative strips spread													
1y	10.62	12.33	12.67	19.67	12.71	19.71	1.04	1.42	0.83	7.84	11.61	0.27	0.23
2y	11.31	9.14	11.48	17.75	11.59	17.95	1.22	2.01	2.45	8.81	8.80	0.35	0.23
3y	8.21	6.44	9.60	14.45	9.83	14.73	1.13	1.98	3.68	8.58	5.69	0.46	0.23
4y	5.50	5.82	8.87	12.59	9.23	13.03	1.10	1.95	4.63	8.42	4.01	0.59	0.23
5y	4.47	4.96	8.62	11.46	9.09	11.93	1.11	1.92	5.37	8.21	3.01	0.71	0.23
5y-1y	-	-	-4.06	-8.22	-3.62	-7.78	0.08	0.50	4.54	0.38	-8.59	0.44	-

Table D-8. The spread between 5-year and 1-year forecasts: the SSP approach

We provide the SSP results based on the in-sample forecasts. The estimation sample is from 2005:M1 to 2017:M2. Our prediction sample is from 2005:M1 to 2013:M2. We use $*$ to indicate the statistical significance at the 90% confidence level.

	Exp. return			Exp. growth		
	50%	[5%	95%]	50%	[5%	95%]
			Entire period			
U.S.	2.58*	[0.80,	3.60]	0.19	[-1.23,	1.57]
Europe	-0.18	[-2.86,	1.73]	-0.36	[-2.96,	1.63]
Japan	-0.24	[-3.44,	1.78]	1.77	[-1.41,	3.81]
			Expansion period			
U.S.	4.69*	[2.16,	5.83]	-1.41*	[-2.88,	-0.22]
Europe	1.14	[-0.41,	2.80]	-6.13*	[-8.67,	-4.46]
Japan	2.81	[-0.26,	3.49]	-0.72*	[-2.79,	-0.08]
			Recession period			
U.S.	-3.85*	[-4.40,	-1.07]	6.17*	[4.72,	7.95]
Europe	-1.91	[-4.45,	0.65]	7.21*	[5.68,	10.75]
Japan	-3.62*	[-5.92,	-0.16]	4.54*	[2.42 ,	8.18]

References

Bansal, R., and H. Zhou, 2002, Term Structure of Interest Rates with Regime Shifts, *Journal of Finance* 57, 1997–2043.

Binsbergen, Jules van, Wouter Hueskes, Ralph Koijen, and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri, 2015, Prior Selection For Vector Autoregressions, *Review of Economics and Statistics* 97, 436–451.

Kim, Chang-Jin, and Charles R. Nelson, 1999, *State Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* (The MIT Press).