Internet Appendix to "Pricing Intertemporal Risk when Investment Opportunities are Unobservable"

Scott Cederburg

University of Arizona

This Internet Appendix contains derivations, details on estimation, and additional empirical results. Section A derives equations in the paper. Section B provides information about the Markov chain Monte Carlo procedure to estimate the predictive system and ICAPM parameters. Section C reports results for a Bayesian predictive regression approach to estimate the intertemporal risk factor and additional results that accompany the primary specifications in the paper.

A Derivations

A.1 Derivation of Equation (7)

Equation (1) implies that the covariance from equation (3),

$$V_{ih} = Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{m,t+1+j}\right),$$

will be positively rewarded when $\gamma > 1$. Given the predictive system in equation (6), the intertemporal hedging factor implied by the model can be expressed as

$$
\begin{aligned}
f_{h,t+1} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} && \text{(A1)}\\
&\approx (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j (r_{f,t+j+1} + r^e_{m,t+1+j}) && \text{(A2)}\\
&= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{f,t+j+1} + (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r^e_{m,t+1+j} && \text{(A3)}\\
&= \sum_{j=1}^{\infty} \rho^j (E_{t+1} - E_t) r_{f,t+j+1} + \sum_{j=1}^{\infty} \rho^j (E_{t+1} - E_t) \overline{r}^e_{m,t+j} && \text{(A4)}\\
&= \sum_{j=1}^{\infty} \rho^j \phi_r^{j-1} (r_{f,t+2} - (1-\phi_r)E_r - \phi_r r_{f,t+1}) + \sum_{j=1}^{\infty} \rho^j \phi_m^{j-1} (\overline{r}^e_{m,t+1} - (1-\phi_m)E_m - \phi_m \overline{r}^e_{m,t}) && \text{(A5)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{r_{f,t+2} - (1-\phi_r)E_r - \phi_r r_{f,t+1}}{\phi_r(1-\phi_r\rho)} + \frac{\overline{r}^e_{m,t+1} - (1-\phi_m)E_m - \phi_m \overline{r}^e_{m,t}}{\phi_m(1-\phi_m\rho)} && \text{(A6)}\\
&= \frac{\eta_{r,t+1}}{\phi_r(1-\phi_r\rho)} + \frac{\eta_{\overline{m},t+1}}{\phi_m(1-\phi_m\rho)}. && \text{(A7)}
\end{aligned}
$$

where the equality in equation (A4) reflects the expectation of the market excess return in equation (6.1) and the law of iterated expectations, the equality in equation (A5) comes from the AR(1) structure of the market risk premium and real risk-free rate in equations (6.5) and (6.6) of the predictive system, and equation (A7) uses the definitions of $\eta_{\overline{m},t+1}$ and $\eta_{r,t+1}$ from equations (6.5) and (6.6). Note that since risk-free rates are determined prior to the period in which they are paid, $E_t r_{f,t+1} = r_{f,t+1}$ and $E_{t+1} r_{f,t+2} = r_{f,t+2}$ in the derivation.

A.2 Derivation of Equation (10)

Equation (10) provides the k-horizon variance ratio as a function of the R^2 from a predictive regression of excess market returns on expected returns and the correlation between shocks to market returns and expected future returns, $\rho_{m\overline{m}}$. Given that I am developing priors based on variance ratios from historical information, the proper variance ratio to consider in the prior predictive analysis is based on the true variance of returns conditional on the predictive system parameter set Ω. The true variance of multiperiod returns is

$$\begin{aligned}
Var(r^e_{m,t\rightarrow t+k}|\Omega) &= Var\left(\sum_{i=1}^{k} r^e_{m,t+i}\right) && \text{(A8)}\\
&= Var\left(\sum_{i=1}^{k} \overline{r}^e_{m,t-1+i} + \sum_{i=1}^{k} \eta_{m,t+i}\right) && \text{(A9)}\\
&= Var\left(\left(\sum_{i=1}^{k} \phi_m^{i-1}\right)\overline{r}^e_{m,t} + \sum_{i=1}^{k-1}\left(\sum_{j=1}^{i} \phi_m^{j-1}\right)\eta_{\overline{m},t+i} + \sum_{i=1}^{k}\eta_{m,t+i}\right), && \text{(A10)}
\end{aligned}$$

where $r^e_{m,t\rightarrow t+k}$ represents the cumulative k-period log excess return on the stock market portfolio. Based on this equation,

$$\begin{aligned}
Var(r^e_{m,t\rightarrow t+k}|\Omega) &= \sum_{i=1}^{k} Var(\eta_{m,t+i}) + \sum_{i=1}^{k-1}\left(\sum_{j=1}^{i}\phi_m^{j-1}\right)^2 Var(\eta_{\overline{m},t+i}) && \text{(A11)}\\
&+ 2\sum_{i=1}^{k}\left(\sum_{j=1}^{i}\phi_m^{j-1}\right) Cov(\eta_{m,t+i},\eta_{\overline{m},t+i}) + \left(\sum_{i=1}^{k}\phi_m^{i-1}\right)^2 Var(\overline{r}^e_{m,t}).
\end{aligned}$$

Thus,

$$\begin{aligned}
Var(r^e_{m,t\rightarrow t+k}|\Omega) &= k\sigma_m^2 + \frac{1}{(1-\phi_m)^2}\left(k-1-2\phi_m\frac{1-\phi_m^{k-1}}{1-\phi_m} + \phi_m^2\frac{1-\phi_m^{2(k-1)}}{1-\phi_m^2}\right)\sigma^2_{\overline{m}} && \text{(A12)}\\
&+ \frac{2}{1-\phi_m}\left(k-1-\phi_m\frac{1-\phi_m^{k-1}}{1-\phi_m}\right)\rho_{m\overline{m}}\sigma_m\sigma_{\overline{m}} + \frac{(1-\phi_m^k)^2}{(1-\phi_m)^2}\frac{\sigma^2_{\overline{m}}}{1-\phi_m^2}
\end{aligned}$$

The variance of one-period returns is

$$Var(r^e_{m,t+1}|\Omega) = \sigma_m^2 + \frac{\sigma^2_{\overline{m}}}{1-\phi_m^2} = \frac{\sigma_m^2}{1-R^2}, \qquad \text{(A13)}$$

where R^2 is the fraction of the unconditional variance of the equity premium to the total unconditional variance of stock returns,

$$R^2 = \frac{\frac{\sigma^2_{\overline{m}}}{1-\phi^2_m}}{\sigma^2_m + \frac{\sigma^2_{\overline{m}}}{1-\phi^2_m}}. \tag{A14}$$

Equation (A12) can be rewritten as

$$Var(r^e_{m,t\to t+k}|\Omega) = k\frac{\sigma^2_m}{1-R^2}\left[1 + A(k)(R^2)^{\frac{1}{2}}(1-R^2)^{\frac{1}{2}}\rho_{m\overline{m}} + B(k)R^2\right], \tag{A15}$$

where $A(k)$ and $B(k)$ are constants conditional on ϕ_m and k,

$$A(k) = 2\left(\frac{1+\phi_m}{1-\phi_m}\right)^{\frac{1}{2}}\left(1 + \frac{1}{k}\left(-1 - \phi_m\frac{1-\phi_m^{k-1}}{1-\phi_m}\right)\right) \tag{A16}$$

and

$$B(k) = -1 + \frac{1+\phi_m}{1-\phi_m}\left(1 + \frac{1}{k}\left(-1 - 2\phi_m\frac{1-\phi_m^{k-1}}{1-\phi_m} + \phi^2_m\frac{1-\phi_m^{2(k-1)}}{1-\phi^2_m} + \frac{(1-\phi^k_m)^2}{1-\phi^2_m}\right)\right). \tag{A17}$$

Finally, using the definition of the k-period variance ratio in equation (9) along with equations (A13) and (A15),

$$\mathrm{VR}(k) = 1 + A(k)(R^2)^{\frac{1}{2}}(1-R^2)^{\frac{1}{2}}\rho_{m\overline{m}} + B(k)R^2. \tag{A18}$$

The $A(k)$ and $B(k)$ constants are positive for $\phi_m > 0$ and increasing in k, and the function $(R^2)^{\frac{1}{2}}(1-R^2)^{\frac{1}{2}}$ is positive and increasing in R^2 over the plausible range of $R^2 \in [0.0, 0.5]$. If $\rho_{m\overline{m}} < 0$ and $R^2 > 0$, the second term in equation (A18) is negative and the third term is positive. The predictive regression R^2 thus has both positive and negative effects on the variance ratio when $\rho_{m\overline{m}} < 0$, and the two parameters interact to determine whether the variance ratio is less than or greater than one.

B Estimation

Draws from the posterior distribution of the parameters can be obtained using the MCMC technique outlined in this appendix. Section B.1 contains the steps to estimate the intertemporal risk factor using the predictive system approach of Pástor and Stambaugh (2009). Section B.2 details the hierarchical Bayes approach to estimate the prices of risk and other parameters of interest for the ICAPM.

B.1 Estimation of the Intertemporal Risk Factor

The following four steps draw a sequence of the market risk premium, real risk-free rate, and related parameters from the system of equations (6). These steps are closely related to those of Pástor and Stambaugh (2009) with the risk-free rate and inflation added into the system. I take 500,000 draws from the posterior distribution and discard the first 100,000 as a burn-in period. From the remaining draws, I keep every fourth draw to reduce serial dependence, which results in a sample of 100,000 draws from the posterior distribution.

1. Draw $\{\overline{r}^e_{m,t}\}_{t=1}^T$, $\{r_{f,t+1}\}_{t=1}^T$, $\{\overline{\pi}_t\}_{t=1}^T | E_x$, E_m, E_r, E_π, ϕ_x, ϕ_m, ϕ_r, ϕ_π, Σ using a forward-filtering, backward-sampling (FFBS) step. All parameter distributions in this step are conditioned on the previous draws of E_x, E_m, E_r, E_π, ϕ_x, ϕ_m, ϕ_r, ϕ_π, and Σ, but the conditioning is suppressed in the notation for simplicity. The set of historical and current returns and state variables observable at time t is denoted by D_t. Defining $r_t \equiv [r^e_{m,t} \quad r_{n,t} \quad \pi_t]'$, $\overline{r}_t \equiv [\overline{r}^e_{m,t} \quad r_{f,t+1} \quad \overline{\pi}_t]'$, $E_{\overline{r}} \equiv [E_m \quad E_r \quad E_\pi]'$, and $\phi_{\overline{r}} = \begin{bmatrix} \phi_m & 0 & 0 \\ 0 & \phi_r & 0 \\ 0 & 0 & \phi_\pi \end{bmatrix}$, the VAR has the form

$$\begin{bmatrix} r_t - E_{\overline{r}} \\ x_t - E_x \\ \overline{r}_t - E_{\overline{r}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & A \\ 0 & \phi_x & 0 \\ 0 & 0 & \phi_{\overline{r}} \end{bmatrix} \begin{bmatrix} r_{t-1} - E_{\overline{r}} \\ x_{t-1} - E_x \\ \overline{r}_{t-1} - E_{\overline{r}} \end{bmatrix} + \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix}, \quad \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} \sim N(0, \Sigma), \qquad \text{(B1)}$$

where $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Define $\tilde{\phi} = \begin{bmatrix} 0 & 0 & A \\ 0 & \phi_x & 0 \\ 0 & 0 & \phi_{\overline{r}} \end{bmatrix}$ and $\Sigma \equiv \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} & \Sigma_{uw} \\ \Sigma_{vu} & \Sigma_{vv} & \Sigma_{vw} \\ \Sigma_{wu} & \Sigma_{wv} & \Sigma_{ww} \end{bmatrix}$.

(a) *Forward Filtering*: The filtering step is used to find sequences for time $t = 1, ..., T$ of the parameters of the distributions

$$\overline{r}_t | D_{t-1} \sim N(a_t, P_t) \tag{B2}$$

and

$$\overline{r}_t | D_t \sim N(b_t, Q_t). \tag{B3}$$

The prior distribution of $\overline{r}_t$ before observing data from time t is given by equation (B2). Once the data from time t is observed, realizations of the market return, nominal risk-free rate, inflation, and state variables give information about the market risk premium, real risk-free rate, and expected inflation. After observing this information, the distribution of $\overline{r}_t$ is updated to equation (B3). Define $z_t = [r_t' \quad x_t']'$, $E_z = [(AE_{\overline{r}})' \quad E_x']'$, and

$$\begin{aligned} &a_t = E(\overline{r}_t | D_{t-1}), && b_t = E(\overline{r}_t | D_t), && P_t = Var(\overline{r}_t | D_{t-1}), \\ &Q_t = Var(\overline{r}_t | D_t), && f_t = E(z_t | D_{t-1}), && R_t = Var(z_t | \overline{r}_t, D_{t-1}), \\ &S_t = Var(z_t | D_{t-1}), && G_t = Cov(z_t, \overline{r}_t | D_{t-1}). \end{aligned} \tag{B4}$$

Also define V as the unconditional variance of $[z' \quad \overline{r}']'$. Then

$$V = \begin{bmatrix} V_{rr} & V_{rx} & V_{r\overline{r}} \\ V_{xr} & V_{xx} & V_{x\overline{r}} \\ V_{\overline{r}r} & V_{\overline{r}x} & V_{\overline{r}\overline{r}} \end{bmatrix} \tag{B5}$$

can be calculated as $\text{vec}(V) = [I - (\tilde{\phi} \otimes \tilde{\phi})]^{-1}\text{vec}(\Sigma)$. Let $V_{zz} = \begin{bmatrix} V_{rr} & V_{rx} \\ V_{xr} & V_{xx} \end{bmatrix}$.

To begin drawing the time series of $\overline{r}_t$, first note that $b_0 = E_{\overline{r}}$ and $Q_0 = V_{\overline{r}\overline{r}}$. Also,

$$\overline{r}_1 | D_0 \sim N(a_1, P_1), \tag{B6}$$

where $a_1 = E_{\overline{r}}$ and $P_1 = V_{\overline{r}\overline{r}}$. Then note that

$$z_1|D_0 \sim N(f_1, S_1), \tag{B7}$$

where $f_1 = E_z$ and $S_1 = V_{zz}$. Further, $G_1 = V_{z\overline{r}}$. Then information about $\overline{r}_1$ is derived from the observation of z_1 since

$$z_1|\overline{r}_1, D_0 \sim N(e_1, R_1), \tag{B8}$$

where $e_1 = f_1 + G_1 P_1^{-1}(\overline{r}_1 - a_1)$ and $R_1 = S_1 - G_1 P_1^{-1} G_1'$. Combining this information using Bayes' rule,

$$\overline{r}_1|D_1 \sim N(b_1, Q_1), \tag{B9}$$

where

$$b_1 = Q_1 \left(P_1^{-1} a_1 + P_1^{-1} G_1' R_1^{-1} (G_1 P_1^{-1} a_1 + (z_1 - f_1)) \right), \tag{B10}$$

and

$$Q_1 = \left(P_1^{-1} + P_1^{-1} G_1' R_1^{-1} G_1 P_1^{-1} \right)^{-1}. \tag{B11}$$

Continuing for $t = 2, ..., T$, we have

$$a_t = (I - \phi_{\overline{r}})E_{\overline{r}} + \phi_{\overline{r}} b_{t-1}, \tag{B12}$$

$$P_t = \phi_{\overline{r}} Q_{t-1} \phi'_{\overline{r}} + \Sigma_{ww}, \tag{B13}$$

$$f_t = \begin{bmatrix} A b_{t-1} \\ (I - \phi_x)E_x + \phi_x x_{t-1} \end{bmatrix}, \tag{B14}$$

$$S_t = \begin{bmatrix} A Q_{t-1} A' + \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix}, \tag{B15}$$

$$G_t = \begin{bmatrix} A Q_{t-1} \phi'_{\overline{r}} \\ 0 \end{bmatrix} + \begin{bmatrix} \Sigma_{uw} \\ \Sigma_{vw} \end{bmatrix}, \tag{B16}$$

$$R_t = S_t - G_t P_t^{-1} G'_t, \tag{B17}$$

$$b_t = Q_t \left(P_t^{-1} a_t + P_t^{-1} G'_t R_t^{-1} (G_t P_t^{-1} a_t + (z_t - f_t)) \right), \tag{B18}$$

$$Q_t = \left(P_t^{-1} + P_t^{-1} G'_t R_t^{-1} G_t P_t^{-1} \right)^{-1}. \tag{B19}$$

The sequences of a_t, b_t, f_t, G_t, P_t, Q_t, and S_t are retained for the backward sampling step.

(b) *Backward Sampling*: Draw $\overline{r}_t | a_t, b_t, f_t, G_t, P_t, Q_t, S_t$ for $t = 0, ..., T$. First, draw

$$\overline{r}_T \sim N(b_T, Q_T). \tag{B20}$$

Then draw $\overline{r}_t$ for $t = T - 1, ..., 0$, where $\overline{r}_t$ is the last three elements of the vector

$$\zeta_t | \zeta_{t+1}, D_t \sim N(h_t, H_t), \tag{B21}$$

where

$$h_t = \begin{bmatrix} r_t \\ x_t \\ b_t \end{bmatrix} + \tilde{\phi} \begin{bmatrix} S_{t+1} & G_{t+1} \\ G'_{t+1} & P_{t+1} \end{bmatrix}^{-1} \begin{bmatrix} z_{t+1} - f_{t+1} \\ \overline{r}_{t+1} - a_{t+1} \end{bmatrix}, \tag{B22}$$

$$H_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Q_t \end{bmatrix} - \tilde{\phi} \begin{bmatrix} S_{t+1} & G_{t+1} \\ G'_{t+1} & P_{t+1} \end{bmatrix}^{-1} \tilde{\phi}', \tag{B23}$$

where

$$\tilde{\phi} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ Q_t A' & 0 & 0 & Q_t \phi'_{\overline{r}} \end{bmatrix}. \tag{B24}$$

The draw of $\overline{r}_{t-1}$ is the $k+4$ through $k+6$ elements of ζ_t, and this draw defines $\overline{r}^e_{m,t-1}$, $r_{f,t}$, and $\overline{\pi}_{t-1}$. Based on the draws of $\overline{r}^e_{m,t}$ and $r_{f,t}$ along with the ϕ_m, ϕ_r, E_m, and E_r parameters, we can calculate $\eta_{\overline{m},t}$ and $\eta_{r,t}$ from equations (6.5) and (6.6) for use in estimating the intertemporal risk factor.

2. Draw $[E'_x \quad E'_{\overline{r}}]' | \{\overline{r}_t\}_{t=1}^T, \phi_x, \phi_{\overline{r}}, \Sigma \sim N(\tilde{E}_{x\overline{r}}, \tilde{V}_{x\overline{r}})$. Let $E_{x\overline{r}} = [E'_x \quad E'_{\overline{r}}]'$. The prior for $E_{x\overline{r}}$ is

$$E_{x\overline{r}} \sim N(E_{x\overline{r}_0}, V_{x\overline{r}_0}). \tag{B25}$$

The prior for $E_{x\overline{r}}$ is centered at zero and diffuse, so the elements of $E_{x\overline{r}_0}$ are set to zero and $V_{x\overline{r}_0}$ is set to $100I$. Information about $E_{x\overline{r}}$ is obtained from the dynamics of the time series of $[x'_t \quad \overline{r}'_t]'$. Defining

$$L_1 = \begin{bmatrix} \phi_x & 0 \\ 0 & \phi_{\overline{r}} \end{bmatrix} \tag{B26}$$

and $L_2 = I - L_1$, the posterior distribution of $E_{x\overline{r}}$ is (conditioning is suppressed for ease of notation),

$$E_{x\bar{r}}|\cdot \sim N(\tilde{E}_{x\bar{r}}, \tilde{V}_{x\bar{r}}), \tag{B27}$$

where

$$\tilde{E}_{x\bar{r}} = \tilde{V}_{x\bar{r}} \left(V_{x\bar{r}_0}^{-1} E_{x\bar{r}_0} + L_2' \Sigma_{vw}^{-1} \sum_{t=1}^{T} \left(\left[\begin{array}{c} x_t \\ \\ \bar{r}_t \end{array} \right] - L_1 \left[\begin{array}{c} x_{t-1} \\ \\ \bar{r}_{t-1} \end{array} \right] \right) \right).$$

and

$$\tilde{V}_{x\bar{r}} = \left(V_{x\bar{r}_0}^{-1} + T L_2' \Sigma_{vw}^{-1} L_2 \right)^{-1}. \tag{B28}$$

3. Draw $\phi_x, \phi_{\bar{r}} | \{\bar{r}_t\}_{t=1}^{T}, E_x, E_{\bar{r}}, \Sigma$. Let $x^k \equiv (x_2^k, ..., x_T^k)'$ be the $(T-1) \times 1$ vector of predictor k in periods $t = 2, ..., T$. Let $x_{(l)}$ be the $(T-1) \times K$ matrix of the K vectors of realizations of the predictors in periods $t = 1, ..., T-1$. Also let $\bar{r}^j \equiv (\bar{r}_2^j, ..., \bar{r}_T^j)'$ and $\bar{r}_{(l)}^j \equiv (\bar{r}_1^j, ..., \bar{r}_{T-1}^j)'$ where $\bar{r}_t^j$ is the market risk premium, real risk-free rate, or expected inflation for j of 1, 2, and 3, respectively, let E_{x^k} be the k-th element of E_x, and let $E_{\bar{r}^j}$ be the j-th element of $E_{\bar{r}}$. Define

$$z = \left[\begin{array}{c} x^1 - \iota_{T-1} E_{x^1} \\ \vdots \\ x^K - \iota_{T-1} E_{x^K} \\ \bar{r}^1 - \iota_{T-1} E_{\bar{r}^1} \\ \bar{r}^2 - \iota_{T-1} E_{\bar{r}^2} \\ \bar{r}^3 - \iota_{T-1} E_{\bar{r}^3} \end{array} \right] \tag{B29}$$

and

$$Z = \begin{bmatrix} x_{(l)} - \iota_{T-1}E'_x & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & x_{(l)} - \iota_{T-1}E'_x & 0 & 0 & 0 \\ 0 & 0 & 0 & \overline{r}^1_{(l)} - \iota_{T-1}E_{\overline{r}^1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \overline{r}^2_{(l)} - \iota_{T-1}E_{\overline{r}^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \overline{r}^3_{(l)} - \iota_{T-1}E_{\overline{r}^3} \end{bmatrix}. \tag{B30}$$

Then z is a $[(T-1)(K+3)] \times 1]$ vector and Z is a $[(T-1)(K+3)] \times (K^2+3)$ matrix. Further define $b \equiv (\text{vec}(\phi'_x)' \quad \text{diag}(\phi_{\overline{r}})')'$. The prior distribution of b is

$$b \sim N(b_0, V_{b_0}) \times 1_{b \in S}, \tag{B31}$$

where $1_{b \in S}$ is equal to one if b is in the space S, which is the space such that the eigenvalues of ϕ_x lie inside the unit circle and each of the three diagonal elements of $\phi_{\overline{r}}$ is in $(-1, 1)$. When b is in S, x_t and each element of $\overline{r}_t$ are stationary. For the uninformative prior results in Section III, the prior parameters are set to produce uninformative priors for b by setting the elements of b_0 to zero and $V_{b_0} = 100I$. In Section IV, the first element of b_0 is 0.97, the first element of V_{b_0} is 0.25, and draws of ϕ_m are restricted to be in $(0, 1)$, such that the priors are somewhat informative about the persistence of the equity premium. This prior information is introduced to facilitate the prior predictive analysis with variance ratios in Section IV.A and has relatively little effect on ICAPM test inferences as shown in Internet Appendix C.2. Information about b arises from the dynamics of the state variables and the market risk premium, real risk-free rate, and expected inflation. The posterior distribution of b is (conditioning is suppressed for notational convenience),

$$b|\cdot \sim N(\tilde{b}, \tilde{V}_b) \times 1_{b \in S}, \tag{B32}$$

where

$$\tilde{b} = \tilde{V}_b \left(V_{b_0}^{-1} b_0 + Z'(\Sigma_{vw}^{-1} \otimes I_{T-1}) z \right),$$

and

$$\tilde{V}_b = \left(V_{b_0}^{-1} + Z'(\Sigma_{vw}^{-1} \otimes I_{T-1}) Z \right)^{-1}. \tag{B33}$$

Draws of b can be obtained by drawing from $N(\tilde{b}, \tilde{V}_b)$ and accepting the draw if $b \in S$. Computational efficiency if often gained by drawing one of the last three elements of b from its marginal truncated normal distribution before drawing the remaining elements from the implied conditional normal distribution. The parameters ϕ_m and $\phi_{\overline{r}}$ are found using the definition of b.

4. Draw $\Sigma | \{\overline{r}_t\}_{t=1}^T, E_x, E_{\overline{r}}, \phi_x, \phi_{\overline{r}}, \Sigma^{(p)}, M_{11}, M_{22}$, where $\Sigma^{(p)}$ denotes the previous draw of Σ.

   The predictive system allows for non-zero correlation between the error terms of realized market excess returns and the market risk premium. Pástor and Stambaugh (2009) note that this correlation is likely negative and develop a method for placing priors on this correlation. Whereas the correlation between unexpected market excess returns and the market risk premium is likely negative, there is no similar expectation for the correlations between the error terms corresponding to the risk-free rate and inflation terms. I therefore follow Pástor and Stambaugh (2009) to place informative priors on the correlation between the error terms from equations (6.1) and (6.5) in some specifications, but put an uninformative prior on the remaining correlations.

   Denote the first element of u corresponding to the error term for the market excess return as u_1, the remaining two elements of u as u_2, the first element of w corresponding to the error term for the market risk premium as w_1, and the remaining two elements as w_2. Then

$$\Sigma \equiv \begin{bmatrix} \Sigma_{u_1u_1} & \Sigma_{u_1u_2} & \Sigma_{u_1v} & \Sigma_{u_1w_1} & \Sigma_{u_1w_2} \\ \Sigma_{u_2u_1} & \Sigma_{u_2u_2} & \Sigma_{u_2v} & \Sigma_{u_2w_1} & \Sigma_{u_2w_2} \\ \Sigma_{vu_1} & \Sigma_{vu_2} & \Sigma_{vv} & \Sigma_{vw_1} & \Sigma_{vw_2} \\ \Sigma_{w_1u_1} & \Sigma_{w_1u_2} & \Sigma_{w_1v} & \Sigma_{w_1w_1} & \Sigma_{w_1w_2} \\ \Sigma_{w_2u_1} & \Sigma_{w_2u_2} & \Sigma_{w_2v} & \Sigma_{w_2w_1} & \Sigma_{w_2w_2} \end{bmatrix}. \tag{B34}$$

Define

$$\Sigma_{11} \equiv \begin{bmatrix} \Sigma_{u_1u_1} & \Sigma_{u_1w_1} \\ \Sigma_{w_1u_1} & \Sigma_{w_1w_1} \end{bmatrix}, \tag{B35}$$

and

$$\hat{\Sigma}_{11,0} \equiv \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix}, \tag{B36}$$

where M_{11}, M_{12}, and M_{22} are prior parameters and M_{12} is a hyperparameter for Σ. A posterior draw of Σ can be obtained using the following two-step process.

(a) Draw $M_{12}|\Sigma^{(p)}, M_{11}, M_{12} \sim p(M_{12}|\Sigma_{11})$, where

$$p(M_{12}|\Sigma_{11}) = |\hat{\Sigma}_{11,0}|^{\frac{T_0-K-4}{2}} \exp\left(-\frac{T_0}{2}\mathrm{tr}(\Sigma_{11}^{-1}\hat{\Sigma}_{11,0})\right),\ M_{12} \in (\underline{c}\sqrt{M_{11}M_{22}}, \overline{c}\sqrt{M_{11}M_{12}}). \tag{B37}$$

The prior parameter value T_0 determines the prior precision for the prior on Σ_{11}. Pástor and Stambaugh (2009) note that T_0 represents the number of pseudo-observations from a hypothetical prior sample. As discussed in the paper, I consider cases with uninformative or informative priors. For the uninformative case, $T_0 = \mathrm{rank}(\Sigma) + 3 = 13$ such that the prior has a minimal impact on the posterior. Informative priors in the paper use $T_0 = T/2$, such that the prior is assumed to contain half as many pseudo-observations as the number of observations in the sample. The M_{11} and M_{22} parameters vary across the prior specifications in the grid over R^2 prior parameter values. The R^2 prior takes 13 evenly spaced values between 0% and 3%, and a value of $x\%$ for this prior parameter indicates that M_{11} is set such that (100-x)% of market variance is attributable to unexpected shocks and M_{22} produces x% of the variance from variation in expected returns given the AR(1) process for returns with an assumed prior persistence parameter of 0.975. Finally, the prior parameters $\underline{c}$ and $\overline{c}$ affect the prior correlation between shocks to current and future expected returns, $\rho_{m\overline{m}}$. In the $\rho_{m\overline{m}}$ dimension, the grid has 13 points for $\overline{c}$ that are evenly spaced between -0.9 and 0.9. These gridpoints produce a grid over the prior mean of $\rho_{m\overline{m}}$ that ranges from -0.9 (when $\overline{c} = -0.9$) to 0.0 (when $\overline{c} = 0.9$). See the Technical Appendix of Pástor and Stambaugh (2009) for details on drawing M_{12} from this distribution. The draw of M_{12} defines a new draw of $\hat{\Sigma}_{11,0}$.

(b) Draw $\Sigma | \{\overline{r}_t\}_{t=1}^T, E_x, E_{\overline{r}}, \phi_x, \phi_{\overline{r}}, \hat{\Sigma}_{11,0}$. First compute the time series of residuals (u_t, v_t, w_t) for $t = 1, ..., T$ as

$$\begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} = \begin{bmatrix} r_t \\ x_t \\ \overline{r}_t \end{bmatrix} - \begin{bmatrix} 0 & 0 & A \\ 0 & \phi_x & 0 \\ 0 & 0 & \phi_{\overline{r}} \end{bmatrix} \begin{bmatrix} r_{t-1} \\ x_{t-1} \\ \overline{r}_{t-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & I - \phi_x & 0 \\ 0 & 0 & I - \phi_{\overline{r}} \end{bmatrix} \begin{bmatrix} 0 \\ E_x \\ E_{\overline{r}} \end{bmatrix}. \quad \text{(B38)}$$

Let X denote the $T \times 2$ matrix of $[u_{1,t} \quad w_{1,t}]$ and $Y_{2,T}$ be the $T \times (K+4)$ matrix of $[u_{2,t} \quad v_t \quad w_{2,t}]$. The posterior distribution of Σ_{11} is inverse Wishart,

$$\Sigma_{11}|\cdot \sim \text{Inverse Wishart}(T_0 \hat{\Sigma}_{11,0} + T\hat{\Sigma}_{11}, T + T_0 - K - 4), \quad \text{(B39)}$$

where $\hat{\Sigma}_{11} = X'X/T$. Further, let $\hat{C} = (X'X)^{-1}X'Y_{2,T}$, $\hat{\Omega} = (Y_{2,T} - X\hat{C})'(Y_{2,T} - X\hat{C})/T$, $V_C = (X_0'X_0 + X'X)^{-1}$, $\tilde{C} = V_C \left[(X_0'X_0)\hat{C}_0 + (X'X)\hat{C}\right]$, and $D = \hat{C}_0'X_0'X_0\hat{C}_0 + \hat{C}'X'X\hat{C} - \tilde{C}'V_C^{-1}\tilde{C}$. The prior parameters $\hat{C}_0$ and $X_0'X_0$ are prior parameters for the covariances between state variables and the market-related shocks, and the prior mean $\hat{C}_0$ is set to be a $2 \times (K+4)$ matrix of zeros and $X_0'X_0 = 10^{-6}I$, such that the prior is uninformative. Then the posterior distribution of Ω is

$$\Omega|\cdot \sim \text{Inverse Wishart}(S_0\hat{\Omega}_0 + T\hat{\Omega} + D, T + S_0), \quad \text{(B40)}$$

where S_0 and $\hat{\Omega}_0$ are prior parameter values. I set S_0 to be $K+7$, which produces an uninformative prior with $\hat{\Omega}_0 = 0.01I$. Then letting $\tilde{c} = \text{vec}(\tilde{C})$, the posterior distribution of $c = \text{vec}(C)$ is

$$c|\Omega, \cdot \sim N(\tilde{c}, \Omega \otimes V_C). \tag{B41}$$

Finally, given (Σ_{11}, C, Ω), construct a posterior draw of Σ using $\Sigma_{11} \equiv \begin{bmatrix} \Sigma_{u_1u_1} & \Sigma_{u_1w_1} \\ \Sigma_{w_1u_1} & \Sigma_{w_1w_1} \end{bmatrix}$, $\begin{bmatrix} \Sigma_{u_2u_1} & \Sigma_{u_2w_1} \\ \Sigma_{vu_1} & \Sigma_{vw_1} \\ \Sigma_{w_2u_1} & \Sigma_{w_2w_1} \end{bmatrix} = C\Sigma_{11}$, and $\begin{bmatrix} \Sigma_{u_2u_2} & \Sigma_{u_2v} & \Sigma_{u_2w_2} \\ \Sigma_{vu_2} & \Sigma_{vv} & \Sigma_{vw_2} \\ \Sigma_{w_2u_2} & \Sigma_{w_2v} & \Sigma_{w_2w_2} \end{bmatrix} = \Omega + C\Sigma_{11}C'$. The correlation between u_1 and w_1 implied by the draw of Σ_{11} is denoted by $\rho_{m\overline{m}}$ in the paper.

Based on the draws of $\eta_{\overline{m},t}$, $\eta_{r,t}$, ϕ_m, and ϕ_r, the time series of the intertemporal risk factor is calculated for each draw using equation (7).

B.2 Testing the ICAPM

The following steps draw from the posterior distribution of the prices of risk for the ICAPM factors. The hiearchical Bayes approach is similar to the one developed by Davies (2010). I take 100,000 draws from the posterior distribution conditional on the draws from the predictive system above, and I discard the first 20,000 draws as a burn-in period. Inferences are made using the remaining 80,000 draws.

1. Draw $\beta_{i,y}|\sigma^2_{i,y}, \{f_{h,t,y}\}_{t=1}^T, \lambda_y, \sigma^2_y \sim N(\overline{\beta}_{i,y}, \overline{V}_{\beta,i,y})$ for $i = 1, ..., N_y$ and $y = 1, ..., Y$, where N_y is the number of assets in period y, $\beta_{i,y} = [\alpha_{i,y} \quad \beta^m_{i,y} \quad \beta^h_{i,y}]'$,

$$\overline{\beta}_{i,y} = \overline{V}_{\beta,i,y}\left(\sigma^{-2}_{i,y}X'_{1,y}R^e_{i,y} + \sigma^{-2}_y X'_{2,y}(\overline{R}^e_{i,y} - \lambda_{0,y})\right), \tag{B42}$$

$$\overline{V}_{\beta,i,y} = \left(\sigma^{-2}_{i,y}X'_{1,y}X_{1,y} + \sigma^{-2}_y X'_{2,y}X_{2,y}\right)^{-1}, \tag{B43}$$

$X_{1,y} = \left[\iota_T \quad \{r^e_{m,t,y}\}_{t=1}^T \quad \{f_{h,t,y}\}_{t=1}^T\right]$, $R^e_{i,y} = \{r^e_{i,t,y}\}_{t=1}^T$, $X_{2,y} = [0 \quad \lambda_{m,y} \quad \lambda_{h,y}]$, $\overline{R}^e_{i,y} = \overline{r}^e_{i,y} + \frac{s^2_{i,y}}{2}$, and $s^2_{i,y}$ is the sample return variance for firm i in period y. For firms with missing returns during period y, the $X_{1,y}$ matrix is altered to match the months of data in $R^e_{i,y}$.

2. Draw $\sigma^2_{i,y}|\beta_{i,y}, \{f_{h,t,y}\}_{t=1}^T, \underline{\nu}, \underline{s}^2 \sim$ Inverse Gamma$(\frac{\overline{s}^2}{2}, \frac{\overline{\nu}}{2})$ for $i = 1, ..., N_y$ and $y = 1, ..., Y$, where $\overline{\nu} = \underline{\nu} + T$, $\overline{s}^2 = \underline{\nu}\underline{s}^2 + s^2$, and

$$s^2 = \sum_{t=1}^{T} \left(r^e_{i,t,y} - \alpha_{i,y} - \beta^m_{i,y} r^e_{m,t,y} - \beta^h_{i,y} f_{h,t,y} \right)^2. \tag{B44}$$

The prior parameters $\underline{\nu}$ and $\underline{s}^2$ can be viewed as the number of pseudo-observations and the sum of squared errors from those observations under the prior. I set $\underline{\nu} = 4$ and $\underline{s}^2$ equal to the average sample variance of the errors from time-series OLS regressions matching equation (8.1).

3. Draw $\lambda_y|\{\beta_{i,y}\}_{i=1}^{N_y}, \sigma^2_y, \overline{\lambda}, V_\lambda \sim N(\hat{\lambda}_y, \hat{V}_{\lambda y})$ for $y = 1, ..., Y$, where

$$\hat{\lambda}_y = \hat{V}_{\lambda y}(V_\lambda^{-1}\overline{\lambda} + \sigma_y^{-2} X'_{1,y} \overline{R}^e_y), \tag{B45}$$

$$\hat{V}_{\lambda y} = (V_\lambda^{-1} + \sigma_y^{-2} X'_{1,y} X_{1,y})^{-1}, \tag{B46}$$

$\lambda_y = [\lambda_{y,0} \quad \lambda_{y,m} \quad \lambda_{y,h}]'$, $X_{1,y} = \begin{bmatrix} \iota_{N_y} & \{\beta^m_{i,y}\}_{i=1}^{N_y} & \{\beta^h_{i,y}\}_{i=1}^{N_y} \end{bmatrix}$, and $\overline{R}^e_y = \{\overline{R}^e_{i,y}\}_{i=1}^{N_y}$.

4. Draw $\sigma^2_{\lambda y}|\lambda_y, \{\beta_{i,y}\}_{i=1}^{N_y}, \underline{\nu}_\lambda, \underline{s}^2_\lambda \sim$ Inverse Gamma$(\frac{\overline{s}^2_\lambda}{2}, \frac{\overline{\nu}_\lambda}{2})$ for $y = 1, ..., Y$, where $\overline{\nu}_\lambda = \underline{\nu}_\lambda + N_y$, $\overline{s}^2_\lambda = \underline{\nu}_\lambda \underline{s}^2_\lambda + s^2_\lambda$, and $s^2_\lambda = \sum_{i=1}^{N_y} \left(\overline{R}^e_{i,y} - \lambda_{0,y} - \lambda_{m,y}\beta^m_{i,y} - \lambda_{h,y}\beta^h_{i,y} \right)^2$. I set $\underline{\nu}_\lambda = 4$ and $\underline{s}^2_\lambda$ equal to the average sample variance of the errors from cross-sectional OLS regressions matching equation (8.2) where the betas are those estimated using OLS to set the prior in step 2.

5. Draw $V_\lambda|\overline{\lambda}, \{\lambda_y\}_{y=1}^Y, g, G \sim$ Inverse Wishart$(Y + g, \sum_{y=1}^Y (\lambda_y - \overline{\lambda})(\lambda_y - \overline{\lambda})' + G)$, where $g = 6$ and $G = gI$.

6. Draw $\overline{\lambda}|\{\lambda_y\}_{y=1}^{Y}, V_\lambda, \underline{V}, \underline{\lambda} \sim N(\tilde{\lambda}, \tilde{V}_\lambda)$, where

$$\tilde{\lambda} = \tilde{V}_\lambda \left((V_\lambda / Y)^{-1} \sum_{y=1}^{Y} \frac{\lambda_y}{Y} + \underline{V}^{-1} \underline{\lambda} \right), \tag{B47}$$

$$\tilde{V}_\lambda = \left((V_\lambda / Y)^{-1} + \underline{V}^{-1} \right)^{-1}, \tag{B48}$$

$\underline{\lambda} = [0 \quad 0 \quad 0]'$, and $\underline{V} = 100I$, which creates a diffuse prior for $\overline{\lambda}$.

C Additional Empirical Results

C.1 Estimation of the Intertemporal Risk Factor Based on a Predictive Regression Approach

The primary specification of the intertemporal risk factor in the paper is based on the predictive system of Pástor and Stambaugh (2009). In this appendix, I use an alternative predictive regression approach to estimate the intertemporal risk factor. In particular, the posterior distribution of the risk factor is produced using a Bayesian predictive regression. The predictive regression is estimated with diffuse normal-inverse-Wishart priors such that the posterior means of the Bayesian predictive regression coefficients closely match their ordinary least squares (OLS) counterparts.

The first step to produce a posterior distribution of the intertemporal risk factor is to estimate the following equations using a Bayesian multivariate regression,

$$r^e_{m,t} = a_m + B_m x_{t-1} + \nu_{m,t}, \tag{C1.1}$$

$$r_{n,t+1} - \pi_t = a_r + B_r x_{t-1} + \nu_{r,t}, \tag{C1.2}$$

$$\nu_t \sim N(0, \Sigma^*). \tag{C1.3}$$

The predictive regression MCMC chain first produces 25,000 independent draws from the posterior distribution. The predictive regression approach assumes that the expected market return and real risk-free rate are perfect linear functions of the predictor variables, such that $\overline{r}^e_{m,t} \equiv a_m + B_m x_t$ and

$r_{f,t+1} \equiv a_r + B_r x_t$. Each draw from the posterior distribution of the predictive regression coefficients from equation (C1) produces a time series of the expected market return and real risk-free rate.

Figure C.1 shows the posterior means and 90% credible intervals for these time series. Comparing the predictive regression estimates in Figure C.1 to the predictive system estimates in Figure 1 reveals similar time-series patterns, though the predictive regression approach tends to produce more extreme movements in the market risk premium. Notably, the 90% credible interval for the market risk premium from the predictive regression is much tighter than the corresponding credible interval using the predictive system. Pástor and Stambaugh (2012) note that allowing for imperfect prediction of the market risk premium using a predictive system increases parameter uncertainty relative to the predictive regression model, which assumes that the risk premium is an exact linear function of the observed state variables. Pástor and Stambaugh (2009) suggest a diagnostic to test whether a predictive regression is correctly specified, in the sense that expected market return is an exact linear function of the state variables. Specifically, the autocorrelation of the residuals should be zero if the predictive regression is properly specified. Figure C.2 shows the posterior distribution of the autocorrelation of predictive regression residuals. The posterior distribution is tightly centered around 0.09 and all 25,000 draws are positive. Thus, the diagnostic test provides strong evidence that a predictive system should be used in place of a predictive regression to account for predictor imperfection, in line with the primary specifications in the paper.

Next, for each of the 25,000 posterior draws of the time series $\{\overline{r}_{m,t}\}_{t=0}^{T}$ and $\{r_{f,t+1}\}_{t=0}^{T}$, I estimate the autoregressions

$$\overline{r}^{e}_{m,t} = (1-\phi_m)E_m + \phi_m \overline{r}^{e}_{m,t-1} + \eta_{\overline{m},t} \tag{C2}$$

and

$$r_{f,t+1} = (1-\phi_r)E_r + \phi_r r_{f,t} + \eta_{r,t} \tag{C3}$$

using Bayesian regressions with $\eta_{\overline{m},t} \sim N(0, \sigma^2_{\overline{m}})$ and $\eta_{r,t} \sim N(0, \sigma^2_r)$. I draw the ϕ_m and ϕ_r parameters and the time series of $\eta_{\overline{m},t}$ and $\eta_{r,t}$ from the autoregressions from the posteriors. Equation (7) is then used to calculate the posterior draw of the intertemporal risk factor time series corresponding to each of the 25,000 posterior draws.

Table C.I shows the posterior means of the parameters from equations (C1) to (C3). Figure C.3 plots the posterior distributions of the predictive regression coefficients for the market risk premium from the Bayesian predictive regression. These posteriors can be compared with the posteriors of the implied predictive regression slopes from the predictive system with uninformative priors that are shown in Figure 3. Unlike the predictive regression approach, the predictive system allows for these predictor variables to be imperfect predictors for the market risk premium. As such, the impact of each state variable on the predicted market return can differ across the two specifications. Both the market risk premium and real risk free rate are estimated to be highly persistent processes with autoregression coefficient posterior means of 0.96 and 0.98, respectively.

Summary statistics for the intertemporal risk factor implied by the predictive regression approach are given in Table C.II. Consistent with the tighter credible interval on the market risk premium observed in Figure C.1 relative to the corresponding predictive system estimates in Figure 1, uncertainty about the hedging factor is much smaller with the predictive regression approach. The average posterior standard deviation of monthly factor draws is 1.98% for the predictive regression compared with 3.84% for the predictive system (shown in Table II). The predictive regression enforces a constraint that the market risk premium is a linear function of given state variables, and Pástor and Stambaugh (2012) note that it is easier to learn about return predictability with this assumption. The predictive system with imperfect prediction does not make this assumption, such that it is likely to more accurately reflect uncertainty about the market risk premium if the perfect prediction assumption of the predictive regression is violated.

Table C.III reports results from testing the ICAPM using the intertemporal risk factor implied by the predictive regression approach. The ICAPM test results using the predictive regression factor are similar to the results from the predictive system approach with uninformative priors in the paper. The estimated price of risk for the intertemporal risk factor is 0.25% per month with a 90% credible interval of -0.30% to 0.80%. The posterior mean of the ICAPM RMSE using the predictive regression is somewhat larger than the predictive system approach with uninformative priors in the paper, but the two models generally produce similar inferences.

Overall, the results using a predictive regression approach largely match those from using the predictive system approach in the paper. I concentrate on the predictive system approach in the paper because I am able to place informative priors on economically meaningful quantities using that method. Further, the predictive system fully accounts for uncertainty in the market risk premium by allowing for imperfect predictors, which is important for measuring the impact of uncertainty about the intertemporal hedging factor for ICAPM inferences.

C.2 Supplementary Results for Primary Specification

Table C.IV reports posterior means of the predictive system parameters under uninformative priors. This predictive system is used in the ICAPM tests in Section III of the paper.

The discussion in Section IV.A.3 refers to variance decomposition results based on the predictive systems with informative priors. Following Campbell and Ammer (1993), total stock market return variance can be decomposed into components attributable to variation in cash flows, the market risk premium, and the risk-free rate:

$$\begin{aligned} Var(r^e_{m,t+1}) = {} & Var\left(\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j}\right) + Var\left(\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}\right) \\ & + Var\left(\sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right) - 2Cov\left(\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j}, \sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}\right) \\ & - 2Cov\left(\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j}, \sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right) + 2Cov\left(\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}, \sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right). \end{aligned} \tag{C4}$$

Further, the variance attributable to discount rate variation is given by $Var\left(\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}\right) + Var\left(\sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right) + 2Cov\left(\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}, \sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right)$. Figure C.4 shows quantiles of posteriors for the proportion of total stock market return variance attributable to cash flows (Panel A), discount rates (Panel B), the market risk premium (Panel C), and the risk-free rate (Panel D) across prior specifications.

Figure 7 in the paper shows quantiles of posteriors for the time-series standard deviation and the correlation between the market and intertemporal risk factors across prior specifications. Figures C.5 and C.6 report information about the remaining intertemporal risk factor statistics that are reported in Table II for the case with uninformative priors. Specifically, Figure C.5 displays posteriors of the time-series mean of the intertemporal risk factor, and Figure C.6 shows the time-series average of the posterior standard deviation of the intertemporal risk factor for each prior specification. In line with the results in Figure 7, the posterior uncertainty about the intertemporal risk factor tends to be lower in the cases in which intertemporal risk is significantly priced.

Figures C.7 and C.8 provide versions of Figures 4 and 9 that do not use interpolation across the gridpoints. For each gridpoint in the 13 by 13 grid, I calculate the percentage of draws in which historical variance ratios for horizons of two to eight years lie outside of the 90% credible intervals from a prior predictive analysis. I also calculate the percentage of posterior draws in which the price of risk for the intertemporal risk factor, $\overline{\lambda}_h$, is negative. To create Figures 4 and 9 in the paper, I interpolate across the gridpoints and show the regions in which the interpolated values satisfy the conditions. Figures C.7 and C.8 report results using information only from the gridpoints without interpolation. This analysis ignores some information from the percentages at each gridpoint but avoids the assumptions of the interpolation.

Figure C.9 provides a version of Figure 9 with intertemporal risk factors that are estimated using a predictive system with an uninformative prior for the persistence parameter of the market risk premium, ϕ_m. The predictive systems in Section IV of the paper are estimated with a somewhat informative prior, $\phi_m \sim N(0.97, 0.25)$ and $\phi_m \in [0, 1]$. I specify this prior in the paper to facilitate the prior predictive analysis for the variance ratios, which is used to determine what region of the prior parameter space is reasonable given historical data. A comparison of Figures 9 and C.9 shows that ICAPM inferences are similar under the uninformative and informative priors on ϕ_m. Figure C.9 does not show the region associated with the prior predictive analysis. Given an uninformative prior for ϕ_m that is centered at zero, $\phi_m < 0$ in half of draws from the prior. Variance ratios that are calculated with a negative value for the persistence of the market risk premium are highly variable across draws, and this scenario does not provide a reasonable reflection of reality. I therefore omit the prior predictive analysis from Figure C.9 and use a base specification with a somewhat informative prior for ϕ_m in the paper.

Table C.I: Predictive Regression Parameter Estimates

This table presents parameter estimates for the predictive regression model given by equations (C1) to (C3) with uninformative priors. Mean draws from the posterior distribution of each parameter are shown. The $TERM$, DEF, DY, and RF variables are the term spread, default spread, dividend yield, and short rate state variables, respectively. The reported figures for the Σ^* parameter are multiplied by 100 for ease of presentation and interpretation. The sample period is January 1952 – December 2014.

$$r^e_{m,t} = a_m + B_m x_{t-1} + \nu_{m,t}$$
$$r_{n,t} - \pi_t = a_r + B_r x_{t-1} + \nu_{r,t}$$
$$\nu_t \sim N(0, \Sigma^*)$$

$$\overline{r}^e_{m,t} = (1-\phi_m)E_m + \phi_m \overline{r}^e_{m,t-1} + \eta_{\overline{m},t}$$
$$r_{f,t+1} = (1-\phi_r)E_r + \phi_r r_{f,t} + \eta_{r,t}$$

a_m	-0.86			
a_r	0.07			
	$TERM$	DEF	DY	RF
B_m	0.30	0.26	0.49	-0.16
B_r	-0.06	-0.00	-0.03	0.04
$\Sigma^* \times 100$	$r^e_{m,t}$	$r_{n,t} - \pi_t$		
$r^e_{m,t}$	1850.56	2.20		
$r_{n,t} - \pi_t$	2.20	2.53		
ϕ_m	0.96			
ϕ_r	0.98			

Table C.II: Intertemporal Risk Factor Time-Series Statistics Using a Predictive Regression Approach

This table presents statistics for the intertemporal risk factor, which is estimated based on the predictive regression in equations (C1) to (C3). The mean, standard deviation, and correlation with the market factor reported for the intertemporal risk factor are calculated as averages of these time-series statistics across posterior draws, and the numbers in brackets show the 90% credible interval for the posterior distribution of each statistic. The average factor uncertainty is the time-series mean of the posterior standard deviation of the monthly intertemporal risk factor. The mean and standard deviation are reported in percent per month. The sample period is January 1952 – December 2014.

Statistic	Intertemporal Risk Factor
Mean	0.00
	[−0.32, 0.32]
Standard Deviation	4.23
	[2.11, 7.28]
Correlation with Market Factor	−0.41
	[−0.57, −0.18]
Average Factor Uncertainty	1.98

Table C.III: Estimated Prices of Risk for ICAPM Factors Using a Predictive Regression Approach

This table reports the prices of risk for the market and intertemporal risk factors. The intertemporal risk factor is estimated based on the predictive regression in equations (C1) to (C3). The table shows estimates for the prices of risk from the system of equations (8) along with the root mean squared error (RMSE). The numbers in brackets show the 90% credible interval for the posterior distribution of each parameter. The sample period is January 1952 – December 2014.

	$E_t r_i^e + \frac{V_{ii}}{2} = \overline{\lambda}_0 + \overline{\lambda}_m \beta_i^m + \overline{\lambda}_h \beta_i^h$			
Model	$\overline{\lambda}_0$	$\overline{\lambda}_m$	$\overline{\lambda}_h$	RMSE
ICAPM – Predictive Regression	0.237 [−0.152, 0.628]	0.576 [0.223, 0.926]	0.247 [−0.303, 0.803]	1.946 [1.788, 2.117]

Table C.IV: Predictive System Parameter Estimates

This table presents parameter estimates for the predictive system in equation (6) under the base specification with uninformative priors. Information on the prior distributions for this model is given in Internet Appendix B. Mean draws from the posterior distribution of each parameter are shown. The $TERM$, DEF, DY, and RF variables are the term spread, default spread, dividend yield, and short rate state variables, respectively. The reported figures for the Σ parameter are multiplied by 100 for ease of presentation and interpretation. The sample period is January 1952 – December 2014.

$$r^e_{m,t} = \overline{r}^e_{m,t-1} + \eta_{m,t}$$
$$r_{n,t} = r_{f,t} + \overline{\pi}_{t-1} + \eta_{n,t}$$
$$\pi_t = \overline{\pi}_{t-1} + \eta_{\pi,t}$$
$$x_t = (I - \phi_x)E_x + \phi_x x_{t-1} + \eta_{x,t}$$
$$\overline{r}^e_{m,t} = (1 - \phi_m)E_m + \phi_m \overline{r}^e_{m,t-1} + \eta_{\overline{m},t}$$
$$r_{f,t+1} = (1 - \phi_r)E_r + \phi_r r_{f,t} + \eta_{r,t}$$
$$\overline{\pi}_t = (1 - \phi_\pi)E_\pi + \phi_\pi \overline{\pi}_{t-1} + \eta_{\overline{\pi},t}$$
$$\eta_t \sim N(0, \Sigma)$$

E_m	0.47	ϕ_m	0.97
E_r	0.05	ϕ_r	0.98
E_π	0.24	ϕ_π	0.98

	$TERM$	DEF	DY	RF
E_x	0.90	1.01	2.86	3.83

ϕ_x	$TERM$	DEF	DY	RF
$TERM$	0.95	0.00	0.00	0.01
DEF	0.07	0.96	-0.01	-0.00
DY	-0.04	-0.01	0.96	0.01
RF	-0.04	0.00	0.00	0.99

$\Sigma \times 100$	$r^e_{m,t}$	$r_{n,t}$	π_t	$TERM$	DEF	DY	RF	$\overline{r}^e_{m,t}$	$r_{f,t+1}$	$\overline{\pi}_t$
$r^e_{m,t}$	1870.80	-1.51	0.29	-1.85	12.82	-74.48	-24.44	-10.77	-0.29	-1.49
$r_{n,t}$	-1.51	0.14	-0.01	-0.02	-0.33	0.20	0.54	-0.07	0.05	0.02
π_t	0.29	-0.01	1.15	-0.09	-0.00	0.12	0.08	0.01	0.01	-0.01
$TERM$	-1.85	-0.02	-0.09	1.15	0.38	0.01	-1.18	0.28	-0.05	-0.07
DEF	12.82	-0.33	-0.00	0.38	6.69	-1.16	-8.87	1.72	-0.68	-0.19
DY	-74.48	0.20	0.12	0.01	-1.16	7.31	2.26	0.99	0.17	0.03
RF	-24.44	0.54	0.08	-1.18	-8.87	2.26	17.57	-3.07	1.31	0.30
$\overline{r}^e_{m,t}$	-10.77	-0.07	0.01	0.28	1.72	0.99	-3.07	1.04	-0.23	-0.06
$r_{f,t+1}$	-0.29	0.05	0.01	-0.05	-0.68	0.17	1.31	-0.23	0.28	-0.09
$\overline{\pi}_t$	-1.49	0.02	-0.01	-0.07	-0.19	0.03	0.30	-0.06	-0.09	0.18

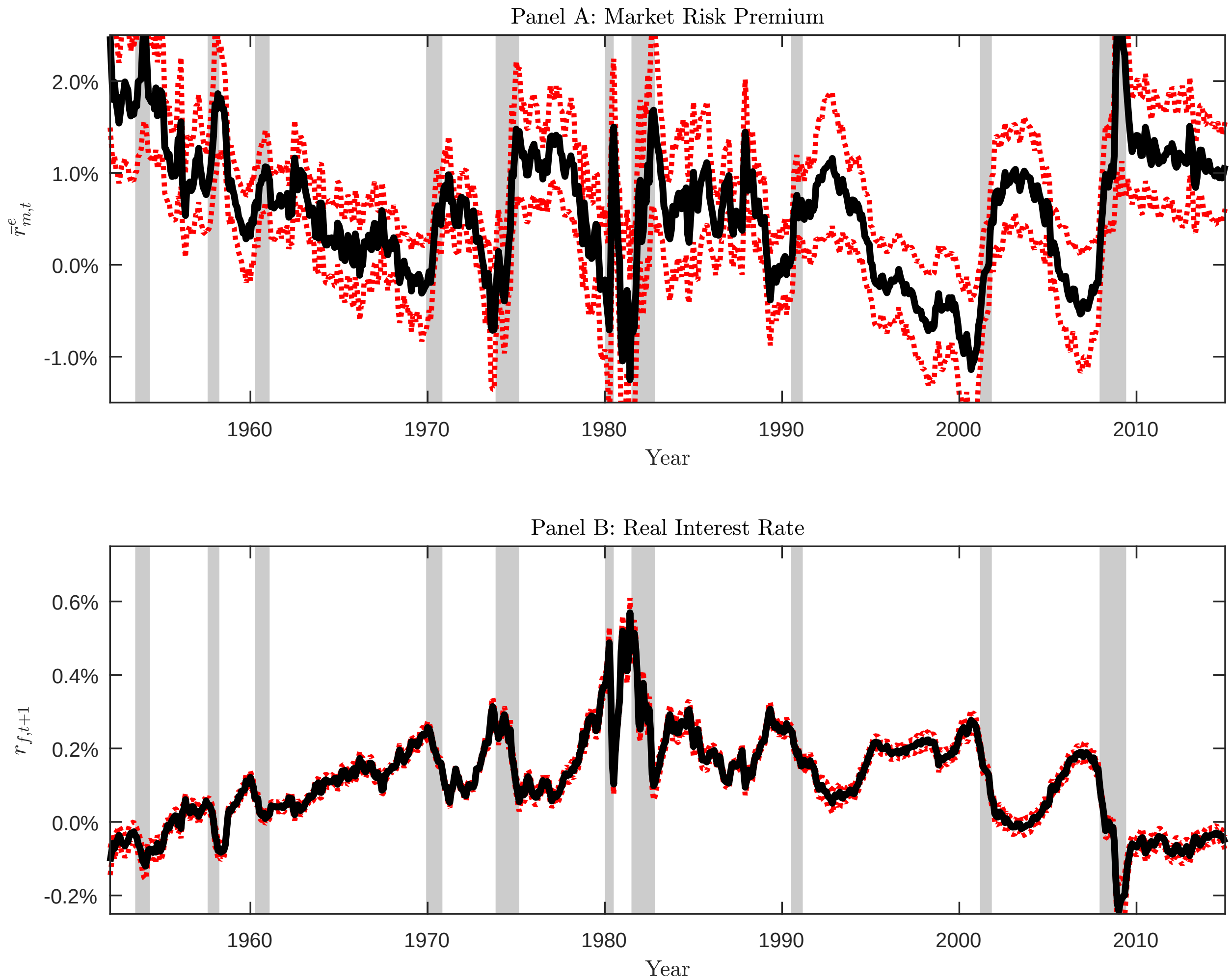


Figure C.1: Predictive Regression Estimates of Latent Processes

This figure shows estimates of the time series of the market risk premium (Panel A) and real interest rate (Panel B) from the predictive regression in equation (C1). The black solid lines represent the posterior means and the red dotted lines show 90% credible intervals for the latent processes. Each variable is expressed in percent per month. The sample period is January 1952 – December 2014, and NBER recessions are shaded.

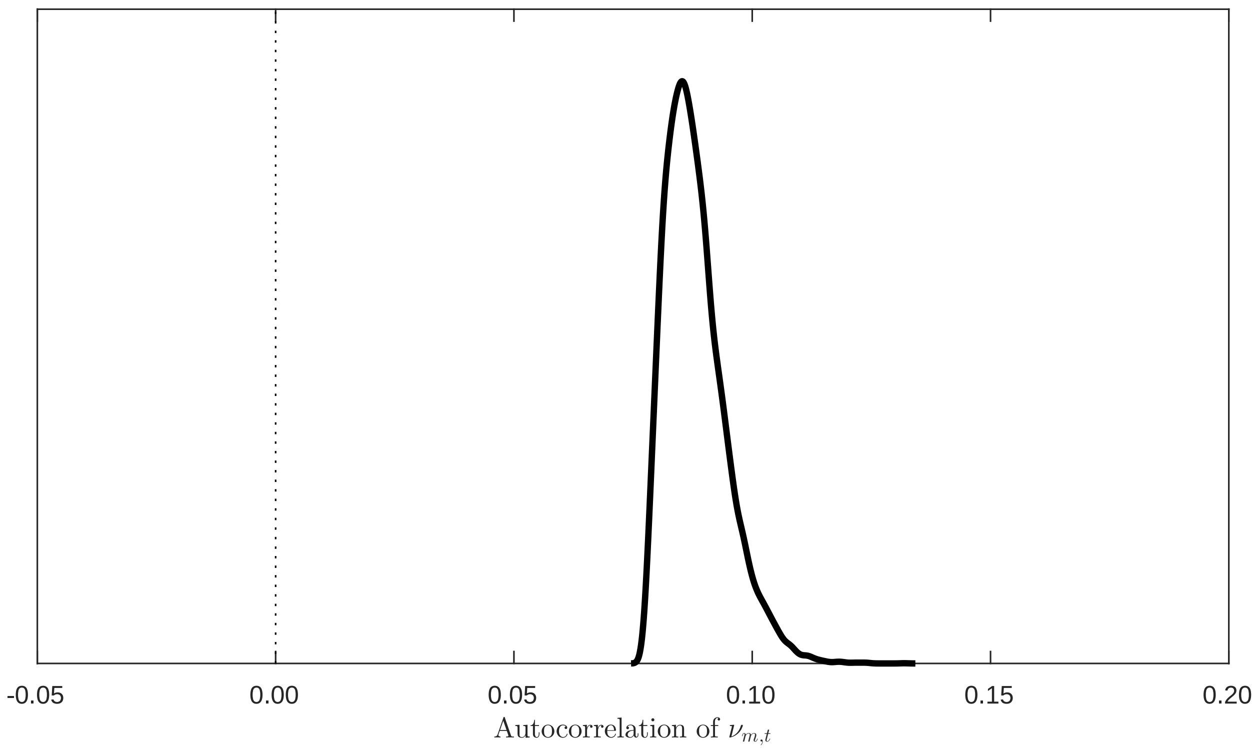


Figure C.2: Posterior of Autocorrelation of Predictive Regression Errors

This figure shows the posterior distribution of the autocorrelation of the error term from the Bayesian predictive regression in equation (C1). The autocorrelation should be zero if the predictive regression is properly specified, and zero is denoted by the dotted line. The sample period is January 1952 – December 2014.

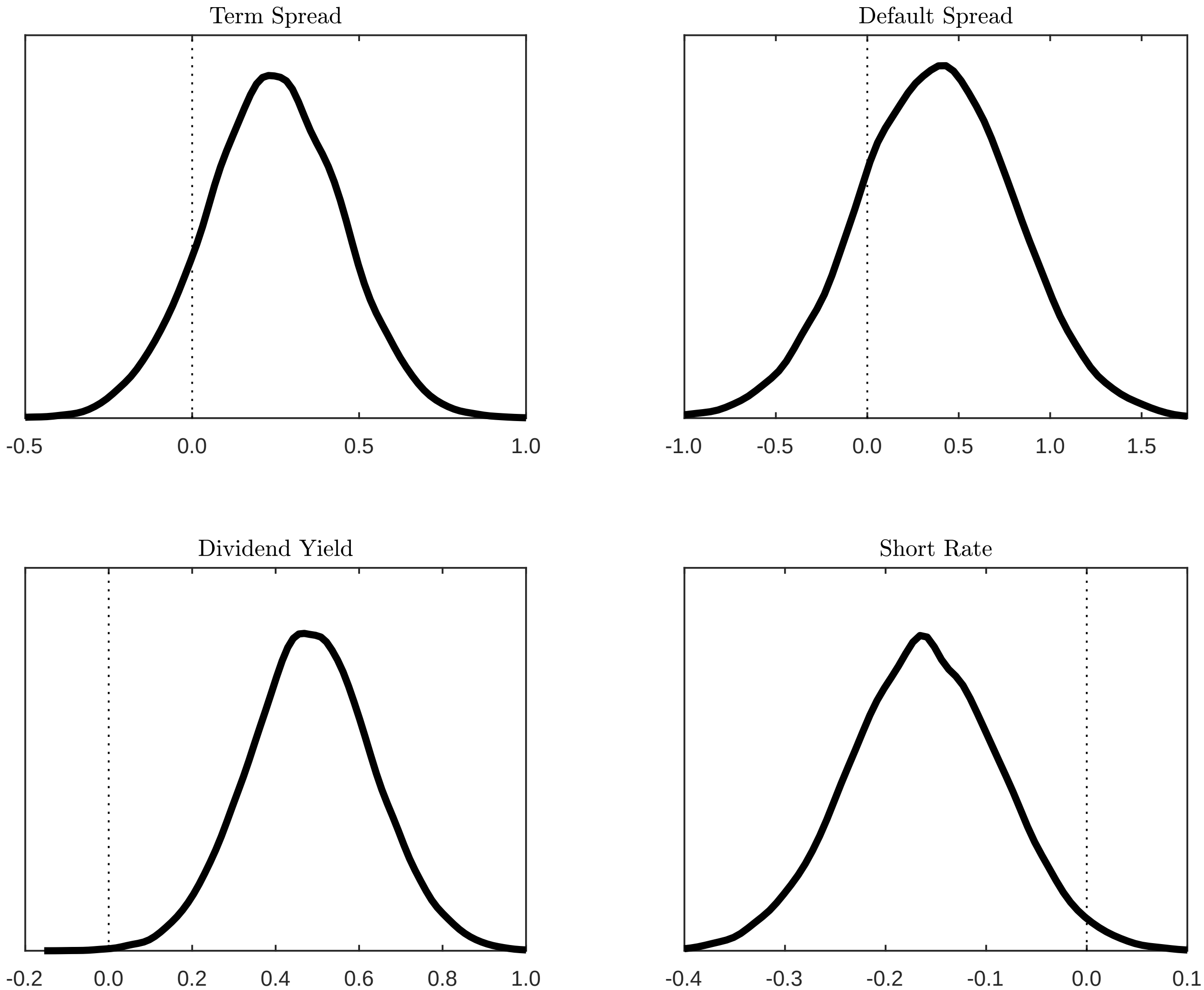


Figure C.3: Posteriors of Predictive Regression Coefficients

This figure shows posterior distributions of the predictive regression slopes from the Bayesian predictive regression in equation (C1). The slopes measure the relation between the market risk premium and the term spread, default spread, dividend yield, and short rate. The sample period is January 1952 – December 2014.

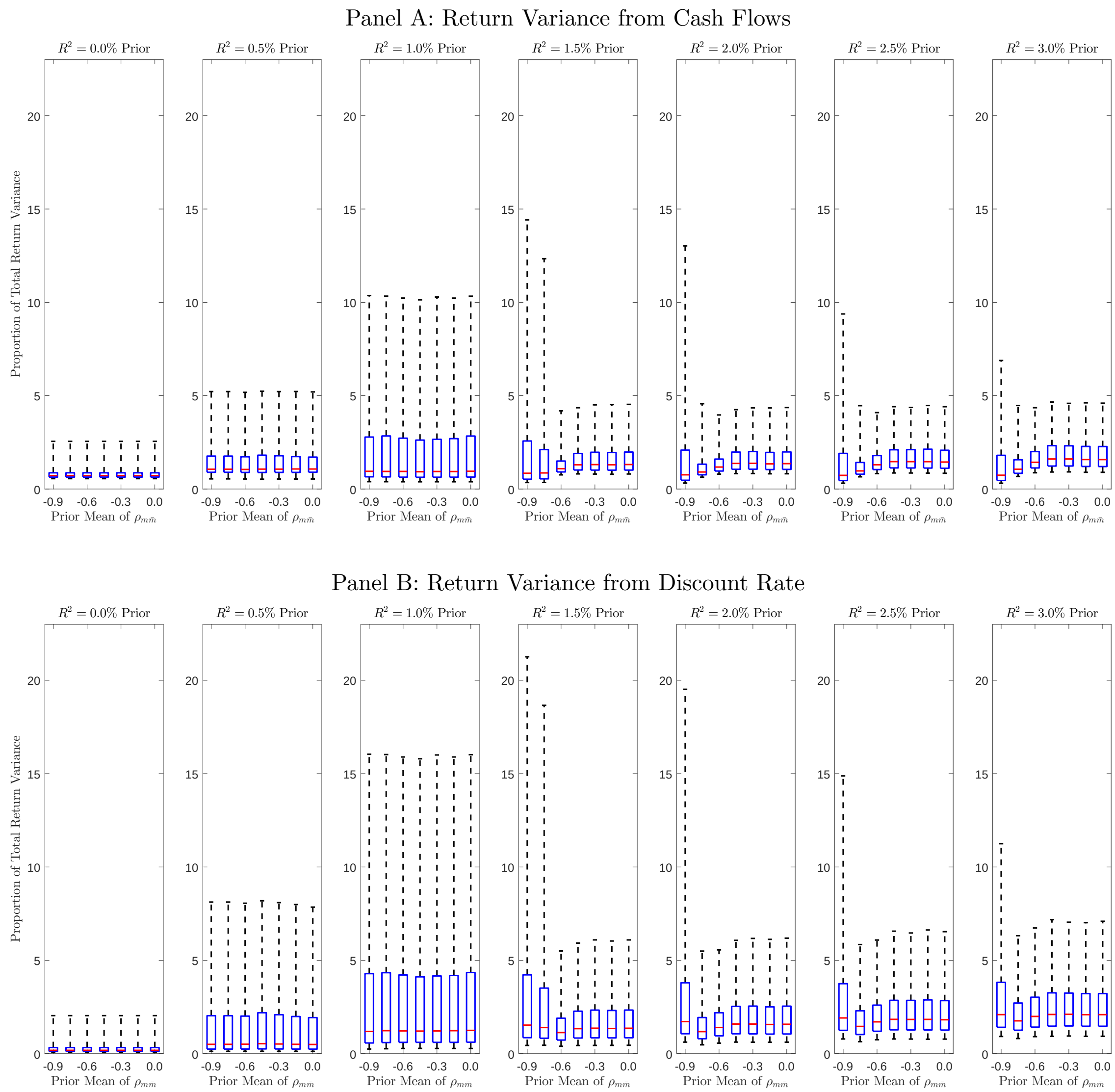
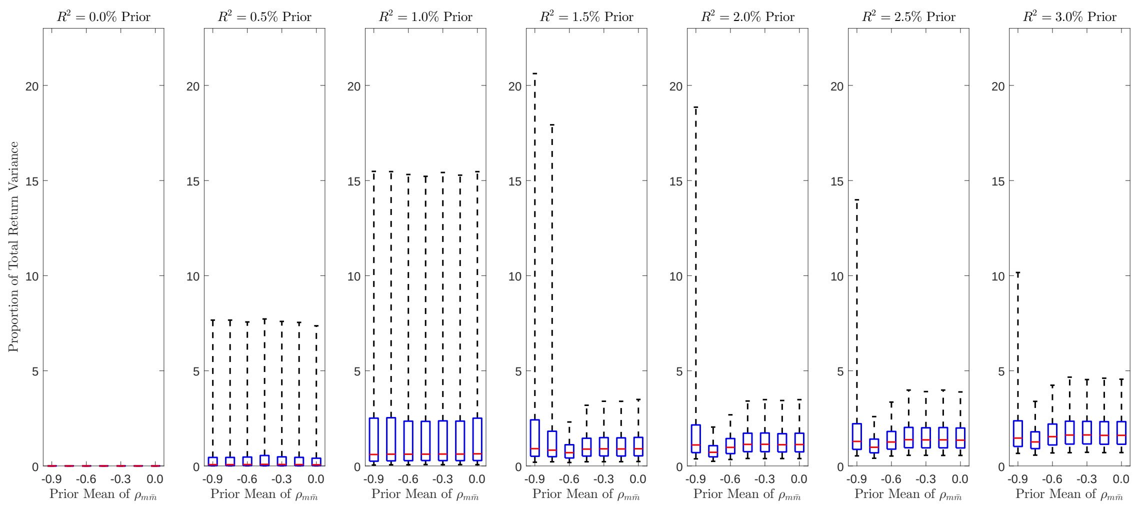


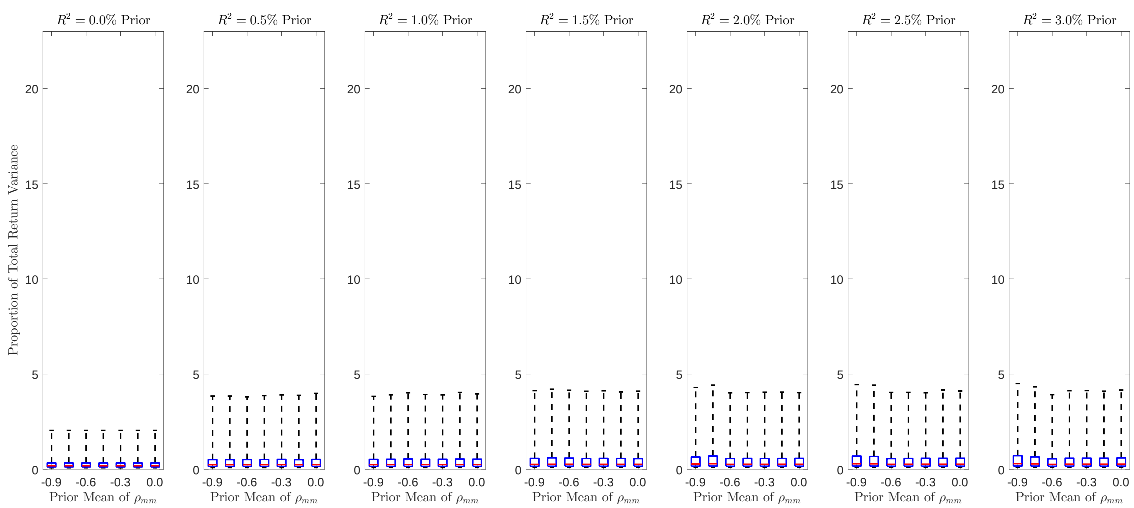
Figure C.4: Variance Decomposition Posteriors

This figure shows quantiles of the posterior distributions of components of stock return variance across specifications of the prior parameters for the predictive regression R^2 and correlation between shocks to current returns and the market risk premium. The figure reports posteriors of the proportions of total stock market return variance that are attributable to cash flows (Panel A), discount rates (Panel B), the market risk premium (Panel C), and the risk-free rate (Panel D). The box shows the median and 25th and 75th percentiles, and the whiskers encompass the 90% credible interval. The variance decomposition is estimated with the predictive system in equation (6) with informative priors as described in Section IV.A. The sample period is January 1952 – December 2014.

Panel C: Return Variance from Market Risk Premium



Panel D: Return Variance from Risk-Free Rate



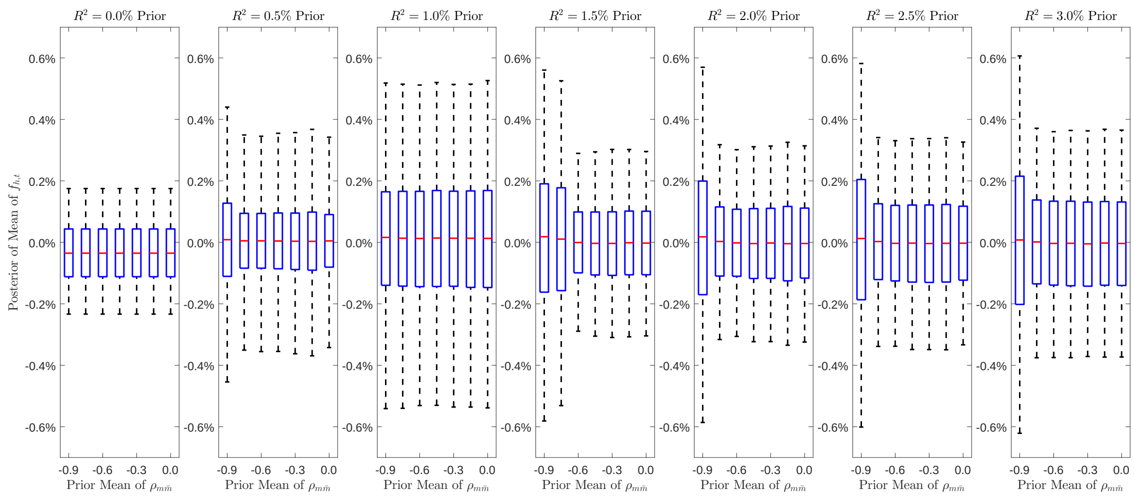


Figure C.5: Posteriors of the Intertemporal Risk Factor Mean across Prior Specifications

This figure shows quantiles of the posterior distributions of the time-series mean of the intertemporal risk factor across specifications of the predictive system prior parameters for the predictive regression R^2 and correlation between shocks to current returns and the market risk premium. The box shows the median and 25th and 75th percentiles, and the whiskers encompass the 90% credible interval. The intertemporal risk factor is estimated with the predictive system in equation (6) with informative priors as described in Section IV.A. The mean of the intertemporal risk factor is expressed in percent per month. The sample period is January 1952 – December 2014.

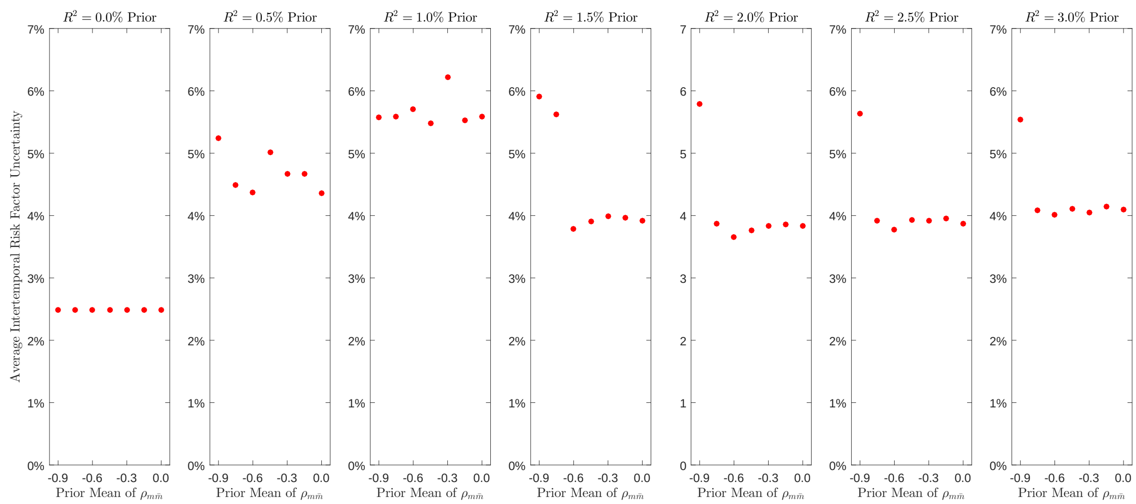


Figure C.6: Average Uncertainty about the Intertemporal Risk Factor across Prior Specifications

This figure shows the time-series average of the posterior standard deviation of the intertemporal risk factor across specifications of the predictive system prior parameters for the predictive regression R^2 and correlation between shocks to current returns and the market risk premium. For each month in the sample, I calculate the posterior standard deviation of $f_{h,t}$ and report the time-series average of these standard deviations. The intertemporal risk factor is estimated with the predictive system in equation (6) with informative priors as described in Section IV.A. The average standard deviation is expressed in percent per month. The sample period is January 1952 – December 2014.

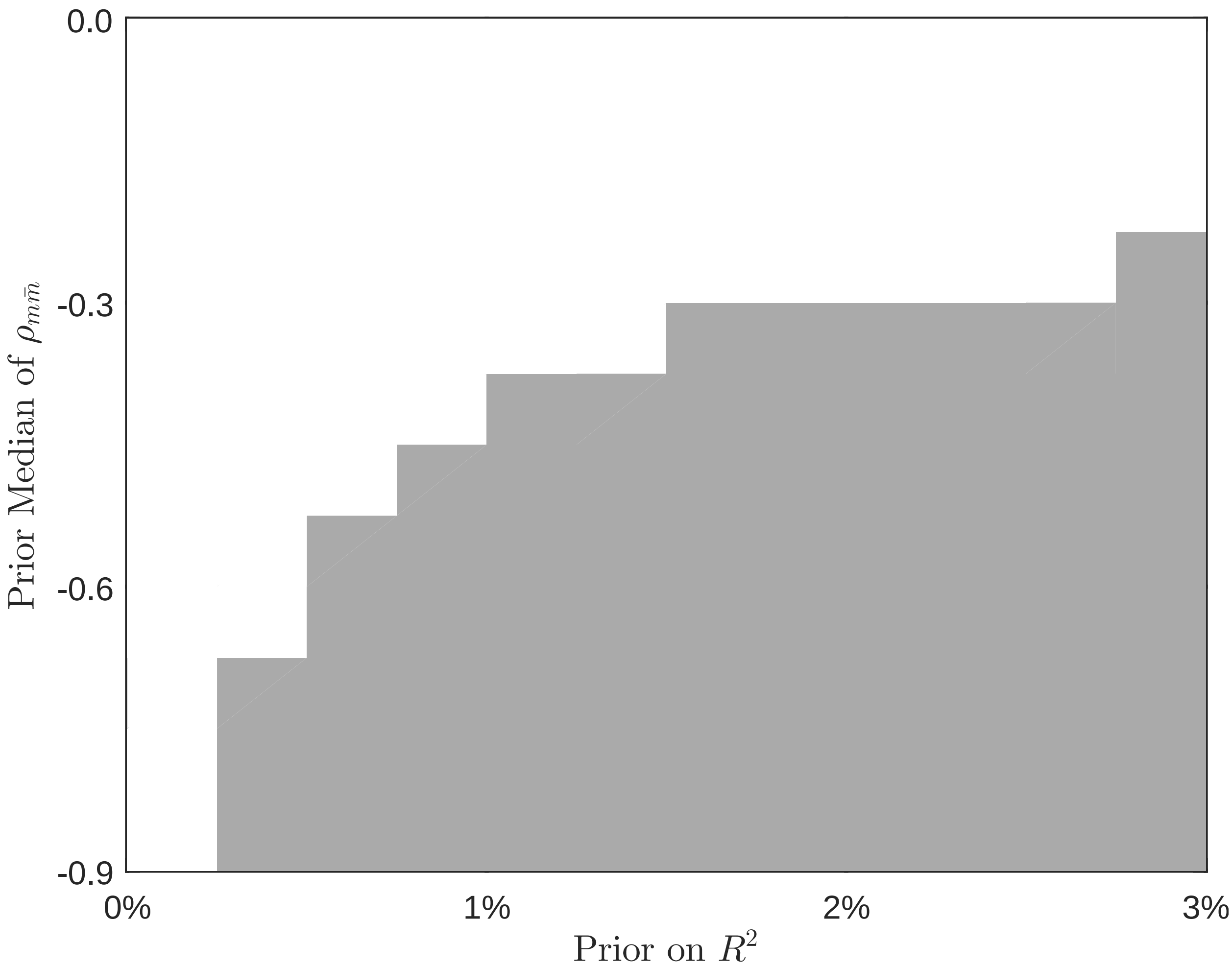


Figure C.7: Region of the Prior Parameter Space that is Consistent with Historical Variance Ratios

This figure shows the region of the prior parameter space in which the historical variance ratios at horizons of two to eight years lie within the 90% credible interval of the variance ratios from a prior predictive analysis. The prior parameter space is a two-dimensional grid over prior parameters corresponding to the predictive regression R^2 and the correlation between shocks to current returns and the market risk premium, $\rho_{m\overline{m}}$. This figure differs from Figure 4 in the paper because this figure does not use interpolation over the gridpoints for the proportion of posterior draws in which the historical variance ratios lie outside of the 90% credible interval. Historical variance ratios are calculated using annual real stock market return data from 1802–1951.

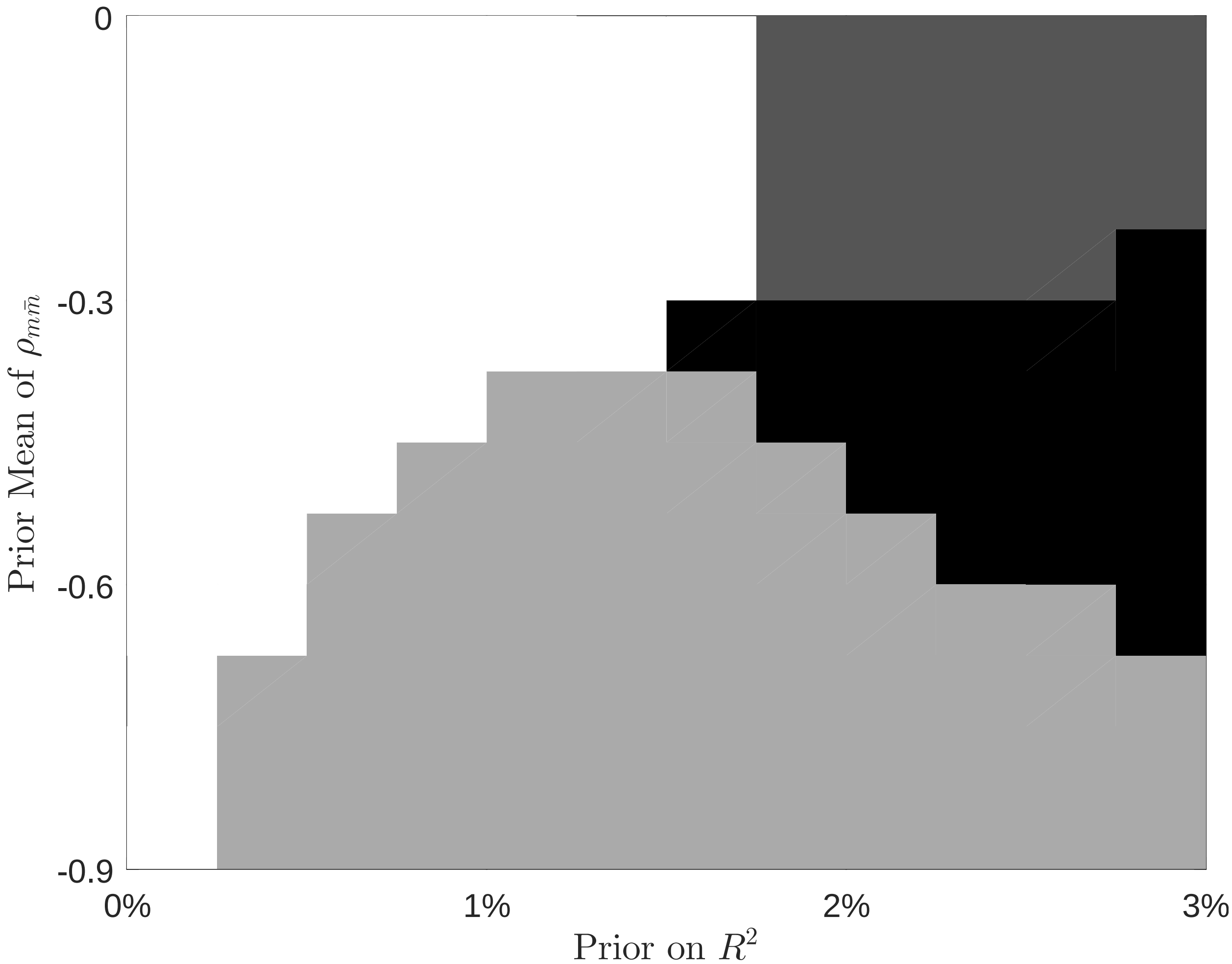


Figure C.8: Region of the Prior Parameter Space in which Intertemporal Risk is Priced

This figure shows the region of the prior parameter space in which zero is not in the 90% credible interval of the price of risk for the intertemporal risk factor in dark gray, the region from Figure C.7 in which the historical variance ratios at horizons of two to eight years lie within the 90% credible interval of the variance ratios from a prior predictive analysis in light gray, and the overlapping region in black. The prior parameter space is a two-dimensional grid over prior parameters corresponding to the predictive regression R^2 and the correlation between shocks to current returns and the market risk premium, $\rho_{m\overline{m}}$. This figure differs from Figure 9 in the paper because this figure does not use interpolation over the gridpoints for the proportions of posterior draws in which the historical variance ratios and $\overline{\lambda}_h$ lie outside of the 90% credible interval. Historical variance ratios are calculated using annual real stock market return data from 1802–1951, and the sample period for ICAPM tests is January 1952 – December 2014.

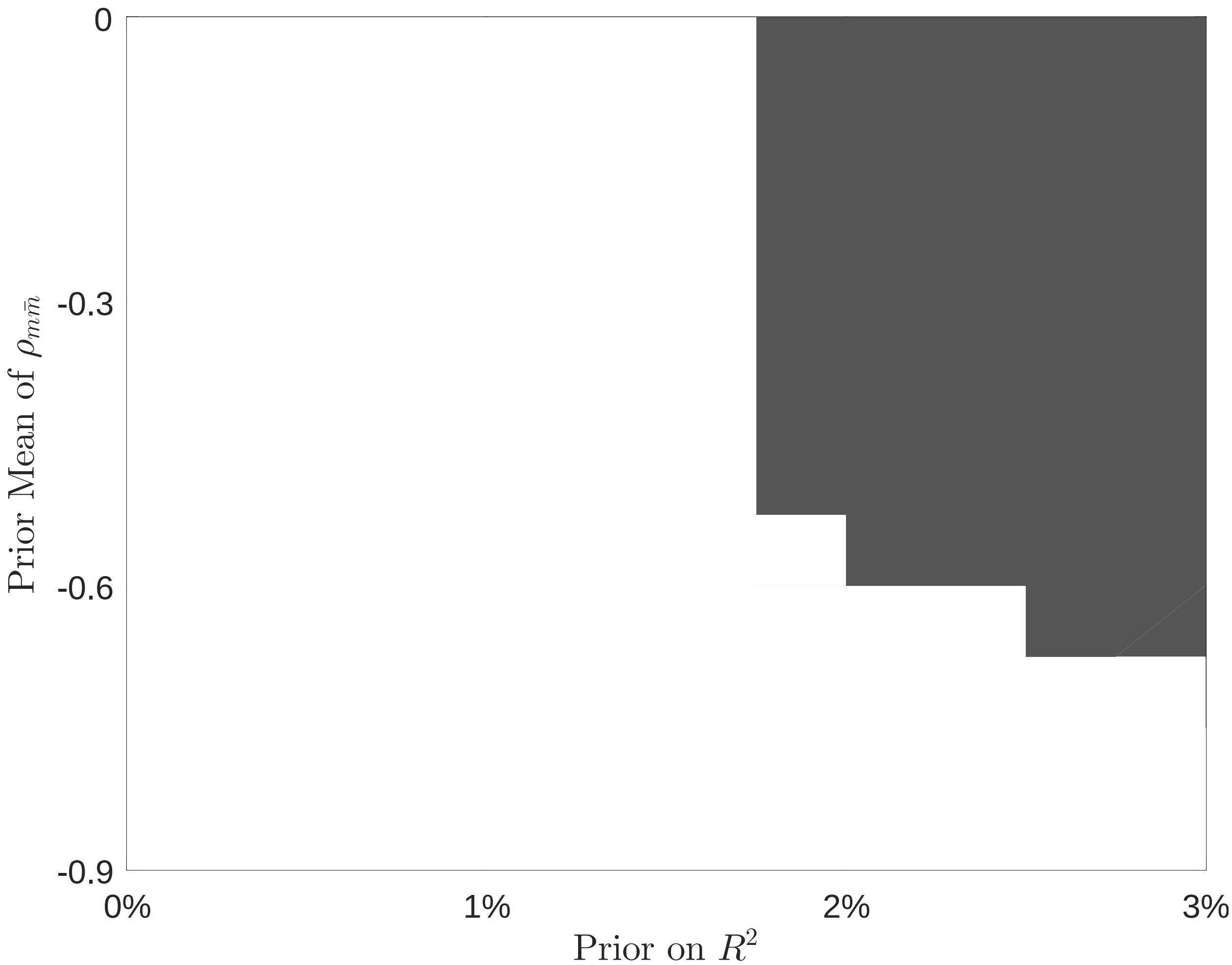


Figure C.9: Region of the Prior Parameter Space in which Intertemporal Risk is Priced with Uninformative Prior on Market Risk Premium Persistence

This figure shows the region of the prior parameter space in which zero is not in the 90% credible interval of the price of risk for the intertemporal risk factor in dark gray. The prior parameter space is a two-dimensional grid over prior parameters corresponding to the predictive regression R^2 and the correlation between shocks to current returns and the market risk premium, $\rho_{m\overline{m}}$. This figure differs from Figure C.8 because the intertemporal risk factor is estimated using a predictive system with an uninformative prior on the persistence parameter for the market risk premium, ϕ_m. The sample period is January 1952 – December 2014.