Internet appendix of

Ambiguity in the Cross-Section of Expected Returns:

An Empirical Assessment

Julian Thimme Clemens Völkert

This internet appendix provides information on summary statistics, the estimation technique, as well as on the simulation study that are not contained in the paper. Section 1 comprises a table with summary statistics of the data used in the paper. Section 2 describes how we estimate the model parameters and conduct tests with GMM. In Section 3, we briefly review the long-run risks model and its solution. We then discuss the conditional expected value, a key feature of the smooth ambiguity model. Afterwards, we extensively discuss the finite sample properties of our estimation technique based on simulated data.

1 Summary statistics

The following table contains summary statistics of the consumption and return data used in the estimation of the three decision models. It only provides information about the 30 test portfolios that are used for our main results presented in Table 1 of the paper. Summary statistics of the alternative portfolios used in the robustness section (Table 2 of the paper) are available upon request.

	Mean	Min	Max	Std. Dev.	Skew	Kurt	AC1
Log Consumption Growth	0.46	-1.25	2.10	0.47	-0.37	4.25	0.45
Risk-Free Rate	0.29	-1.74	03.61	0.67	0.49	5.65	0.47
Log Price-Dividend Ratio	3.50	2.78	4.50	0.40	0.51	2.59	0.97
Market Return (<i>CRSP</i>)	1.90	-26.93	22.35	8.47	-0.51	3.65	0.10
Market Return (<i>cay</i>)	1.49	-1.99	04.98	0.82	-0.04	5.44	0.14
Size 1	2.80	-33.50	41.11	12.96	0.13	3.43	0.07
Size 2	2.62	-32.74	42.30	12.15	-0.03	3.55	0.01
Size 3	2.73	-31.47	38.17	11.54	-0.15	3.60	-0.01
Size 4	2.58	-33.61	34.87	11.13	-0.24	3.65	-0.01
Size 5	2.62	-28.08	34.48	10.73	-0.26	3.50	0.02
Size 6	2.47	-28.94	30.67	10.00	-0.34	3.48	0.04
Size 7	2.45	-26.72	32.97	9.97	-0.30	3.59	0.06
Size 8	2.31	-23.58	27.72	9.55	-0.39	3.36	0.04
Size 9	2.14	-24.90	21.65	8.72	-0.59	3.56	0.08
Size 10	1.73	-29.11	22.55	7.95	-0.54	3.88	0.14
BM 1	1.78	-35.71	25.86	9.85	-0.46	3.98	0.11
BM 2	1.83	-26.73	25.92	8.85	-0.39	3.89	0.02
BM 3	2.03	-25.32	22.58	8.66	-0.35	3.34	0.05
BM 4	1.97	-24.72	24.60	8.69	-0.46	3.70	0.10
BM 5	2.16	-24.85	20.79	8.02	-0.53	3.40	0.10
BM 6	2.17	-25.54	22.89	8.30	-0.44	3.53	0.13
BM 7	2.17	-20.78	27.97	8.45	-0.23	3.30	0.15
BM 8	2.63	-34.29	35.58	8.98	-0.27	4.69	0.12
BM 9	2.70	-22.32	31.59	9.27	-0.19	3.73	0.06
BM 10	3.00	-31.11	48.05	11.26	0.25	4.80	0.12
Industry 1 (NoDur)	2.42	-31.39	30.08	8.45	-0.46	4.42	0.05
Industry 2 (Durbl)	1.97	-37.50	42.51	11.79	-0.11	3.91	0.12
Industry 3 (Manuf)	2.05	-29.49	27.87	9.29	-0.46	3.77	0.09
Industry 4 (Enrgy)	2.29	-27.03	26.34	8.89	-0.39	3.38	0.09
Industry 5 (HiTec)	2.42	-35.00	39.07	12.09	-0.24	4.17	0.08
Industry 6 (Telcm)	1.77	-23.23	25.91	8.41	-0.18	3.28	0.15
Industry 7 (Shops)	2.25	-35.38	43.46	10.00	-0.11	5.17	0.04
Industry 8 (Hlth)	2.49	-35.44	24.55	9.40	-0.53	4.22	0.05
Industry 9 (Utils)	1.82	-23.57	26.11	7.28	-0.38	3.93	0.10
Industry 10 (Other)	1.92	-27.14	24.14	10.15	-0.50	3.35	0.09

Table 1: Summary Statistics

The table shows descriptive statistics of log consumption growth, the risk-free rate, the log price-dividend ratio, two proxies for the market return, and the 30 equity test assets. The log consumption growth rate and returns are expressed in percentage terms. The sample period is from the first quarter of 1952 to the third quarter of 2011. The data is described in Appendix A in the paper.

2 GMM Estimation and Model Evaluation

Let $\mathbf{R}_{t,t+1} = (R_{t,t+1}^1, \dots, R_{t,t+1}^n)'$ denote the vector of gross returns. The population moment conditions are

$$\mathbb{E} [\mathbf{f}(\Theta, \mathbf{R}_{t,t+1})] = 0.$$

For gross returns, we have $\mathbf{f}(\Theta, \mathbf{R}_{t,t+1}) = \xi_{t,t+1}(\Theta)\mathbf{R}_{t,t+1} - \mathbf{1}$ and for excess returns $\mathbf{f}(\Theta, \mathbf{R}_{t,t+1}) = \xi_{t,t+1}(\Theta) (\mathbf{R}_{t,t+1} - R_{t,t+1}^f \mathbf{1})$, where we make the dependence of the pricing kernel on the parameter vector Θ explicit. The sample equivalent of the population orthogonality conditions is

$$\mathbf{g}_T(\Theta, \mathbf{R}) = \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{f}(\Theta, \mathbf{R}_{t,t+1}).$$

The GMM estimator of Θ , denoted by $\hat{\Theta}$, minimizes the criterion function

$$\hat{\Theta} = \operatorname{argmin}_{\Theta} \mathbf{g}_T(\Theta, \mathbf{R})' \mathbf{W}_T \mathbf{g}_T(\Theta, \mathbf{R}),$$

where $\mathbf{W}_T$ is a positive semi-definite weighting matrix which converges in probability to a positive definite matrix of constants $\mathbf{W}$.

Let $\mathbf{S}$ denote the variance covariance matrix of the moment conditions. Hansen (1982) shows that the weighting matrix that minimizes the asymptotic variance of

the parameter estimates is $\mathbf{W} = \mathbf{S}^{-1}$. Even if $\mathbf{S}$ is not used as weighting matrix, we still need to estimate $\mathbf{S}$ to construct standard errors and to test hypotheses. We estimate the long-run covariance matrix $\mathbf{S}$ according to Newey and West (1987b). As in Yogo (2006), the lag length is set to one to account for time aggregation in consumption data.

Hansen (1982) shows that the variance of the parameter estimates and the estimated moment conditions are

$$\begin{aligned}\text{Var}(\hat{\Theta}) &= \frac{1}{T} (\mathbf{D}'\mathbf{W}\mathbf{D})^{-1} \mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D} (\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}, \\ \text{Var}\left(\mathbf{g}_T(\hat{\Theta}, \mathbf{R})\right) &= \frac{1}{T} (\mathbf{I}_n - \mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}')' \mathbf{S} (\mathbf{I}_n - \mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'),\end{aligned}$$

where $\mathbf{D} = \mathbb{E} [\partial \mathbf{f}(\Theta, \mathbf{R}_{t,t+1}) / \partial \Theta']$ and $\mathbf{I}_n$ is an n -dimensional identity matrix.

To perform inference on the parameters, we use Wald tests. Let $\mathbf{C}$ denote a $k \times k$ matrix and $\mathbf{c}$ a $k \times 1$ vector, where k denotes the number of estimated parameters. Newey and West (1987a) show that the restrictions

$$\mathbf{C}\Theta - \mathbf{c} = 0$$

can be tested using the following test statistic

$$(\mathbf{C}\hat{\Theta} - \mathbf{c})' \left(\mathbf{C} \text{Var}(\hat{\Theta}) \mathbf{C}' \right)^{-1} (\mathbf{C}\hat{\Theta} - \mathbf{c}),$$

which is χ^2 distributed with degrees of freedom equal to the rank of $\mathbf{C}$. We use this test to check for ambiguity neutrality ($\gamma = \eta$). We also test the hypothesis $\gamma = \rho$, which corresponds to time-separability in the EZ model and to pure ambiguity preferences in the SA model.

We perform Hansen's J-test for overidentifying restrictions

$$\mathbf{g}_T(\hat{\Theta}, \mathbf{R})' \text{Var} \left(\mathbf{g}_T(\hat{\Theta}, \mathbf{R}) \right)^{-1} \mathbf{g}_T(\hat{\Theta}, \mathbf{R}),$$

which has a χ^2 distribution with n (number of moments) minus k (number of estimated parameters) degrees of freedom under the null of correct model specification. Since the variance covariance matrix of the moment conditions is singular, we use the Moore-Penrose pseudoinverse.

We compare root mean squared errors (RMSE) and Hansen and Jagannathan (1997) distances (HJD) of different preference specifications and parameter vectors. The RMSE is the square-root of the objective function that is minimized if the identity matrix is used for weighting

$$\text{RMSE}(\Theta, \mathbf{R}) = \sqrt{\mathbf{g}_T(\Theta, \mathbf{R})' \mathbf{I}_n \mathbf{g}_T(\Theta, \mathbf{R})}.$$

The Hansen and Jagannathan (1997) distance is given by

$$\text{HJD}(\Theta, \mathbf{R}) = \sqrt{\mathbf{g}_T(\Theta, \mathbf{R})' G(\mathbf{R})^{-1} \mathbf{g}_T(\Theta, \mathbf{R})},$$

i.e. the square-root of the objective function if the second moment matrix $G(\mathbf{R})$, with $G(\mathbf{R})_{i,j} := \frac{1}{T} \sum_{t=0}^{T-1} R_{t,t+1}^i R_{t,t+1}^j$, is used for weighting.

We use the methodology of Jagannathan and Wang (1996) and Parker and Julliard (2005) to test if the RMSE or the HJD is zero. Let

$$\mathbf{A} = \mathbf{S}^{1/2} (\mathbf{W} - \mathbf{W} \mathbf{D} (\mathbf{D}' \mathbf{W} \mathbf{D})^{-1} \mathbf{D}' \mathbf{W}) (\mathbf{S}^{1/2})',$$

where $\mathbf{S}^{1/2}$ is the upper-triangular matrix from the Cholesky decomposition of $\mathbf{S}$. The matrix $\mathbf{A}$ has $n-k$ positive eigenvalues, denoted by $\lambda_1, \dots, \lambda_{n-k}$. The asymptotic sampling distribution of the distance is

$$T \mathbf{g}_T(\hat{\Theta}, \mathbf{R})' \mathbf{W}_T \mathbf{g}_T(\hat{\Theta}, \mathbf{R}) \xrightarrow{d} u = \sum_{j=1}^{n-k} \lambda_j v_j \text{ as } T \rightarrow \infty,$$

where $v_1, \dots, v_{n-k}$ are independent χ^2 random variables with one degree of freedom. The empirical p -value of the statistic can be computed by drawing $T(n-k)$ independent random numbers from a $\chi^2(1)$ distribution and counting the number of cases where u exceeds the test statistic. If the optimal weighting matrix is used, i.e. $\mathbf{W} = \mathbf{S}^{-1}$, all eigenvalues are unity and this test coincides with the J-test.

The cross-sectional R^2 is computed as

$$R^2 = 1 - \frac{\text{Var}(\mathbb{E}[\mathbf{R}_{t,t+1}] - \mathbb{E}_p[\mathbf{R}_{t,t+1}])}{\text{Var}(\mathbb{E}[\mathbf{R}_{t,t+1}])},$$

where the predicted returns are calculated using the covariance decomposition

$$\mathbb{E}_p [\mathbf{R}_{t,t+1}] = \frac{1 - \text{Cov} [\xi_{t,t+1}, \mathbf{R}_{t,t+1}]}{\mathbb{E} [\xi_{t,t+1}]}.$$

3 Long-Run Risks Model

In this section, we estimate the SA, EZ, and PA models based on simulated consumption and return data to investigate the performance of our estimation technique in finite samples. As in Bansal and Yaron (2004), the laws of motion of consumption and dividend growth are driven by two long-run risk factors.

3.1 Data generating process

In Bansal and Yaron (2004), the distribution of consumption and dividend growth, hence the time $t + 1$ economic model, depends on the realizations of two state variables that characterize the economic circumstances at time $t + 1$. We use a similar endowment process and assume that the representative investor is ambiguous about the state variables that drive consumption and dividend growth in the long-run. Short-run consumption and dividend uncertainty are treated as risk and evaluated with the risk aversion coefficient γ . The long-run growth factor and the conditional volatility of consumption growth are perceived as ambiguous. The parameter η determines the investor's attitude towards these sources of uncertainty.

The laws of motion of log consumption growth $\Delta c_{t+1} = \log C_{t+1} - \log C_t$, log dividend growth $\Delta d_{t+1} = \log D_{t+1} - \log D_t$, and of the state variables are

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_{t+1} + \sigma_{t+1} w_{t+1}^c, \\
\Delta d_{t+1} &= \mu_d + \lambda x_{t+1} + \sigma_{t+1} \phi_{d,\sigma} \left(\rho_{cd} w_{t+1}^c + \sqrt{1 - \rho_{cd}^2} w_{t+1}^d \right), \\
x_{t+1} &= \varphi_x x_t + \phi_x \sigma_{t+1} w_{t+1}^x, \\
\sigma_{t+1}^2 &= \sigma^2 + \varphi_\sigma (\sigma_t^2 - \sigma^2) + \phi_\sigma w_{t+1}^\sigma,
\end{aligned} \tag{1}$$

where $w_{t+1}^c, w_{t+1}^d, w_{t+1}^x, w_{t+1}^\sigma \sim i.i.d. \mathcal{N}(0, 1)$. Consumption and dividend growth contain a persistent long-run growth component x_{t+1} and the conditional volatilities are driven by a time-varying uncertainty factor σ_{t+1}^2 . The solution of the model is sketched in Section 3.2. Note that in Bansal and Yaron (2004), consumption growth depends on x_t and σ_t^2 and therefore, at time t , the investor knows that $\Delta c_{t+1} \sim \mathcal{N}(\mu_c + x_t, \sigma_t^2)$. Consequently, she cannot be ambiguous about the distribution. In contrast, we assume that consumption growth depends on x_{t+1} and σ_{t+1}^2 which are ambiguous. More precisely, we assume that a candidate model π_{t+1} is characterized by a pair $(x_{t+1}^\pi, (\sigma_{t+1}^2)^\pi)$ that pins down the distribution of consumption and dividend growth at time t .

3.2 Model Solution

We solve the model in the same manner as Bansal and Yaron (2004), Bansal et al. (2012a), and Beeler and Campbell (2012) using analytical approximations. We assume that the log wealth-consumption ratio z is affine in the state variables

$$z_t = A + B_x x_t + B_\sigma (\sigma_t^2 - \sigma^2).$$

For the log return on the consumption claim $r_{t,t+1}^w = \log R_{t,t+1}^w$, we use the log-linear return approximation of Campbell and Shiller (1988)

$$r_{t,t+1}^w = k_0 + k_1 z_{t+1} - z_t + \Delta c_{t+1},$$

where k_0 and k_1 are linearizing constants. It holds that $k_0 = \log(1 + e^{\bar{z}}) - k_1 \bar{z}$, where $k_1 = \frac{e^{\bar{z}}}{1 + e^{\bar{z}}}$ and $\bar{z}$ is the long-run mean of the log wealth-consumption ratio. Using the Euler equation $\mathbb{E}_t [e^{\log \xi_{t,t+1} + r_{t,t+1}^w}] = 1$ yields the following coefficients of the wealth-consumption ratio

$$\begin{aligned} A &= \frac{1}{1 - k_1} \left(-\delta + k_0 + (1 - \rho)\mu_c + (1 - k_1\varphi_\sigma) \frac{B_\sigma}{\varphi_\sigma} \sigma^2 + \frac{1 - \eta}{2(1 - \rho)} \left(\frac{B_\sigma}{\varphi_\sigma} \phi_\sigma \right)^2 \right), \\ B_x &= \frac{1 - \rho}{1 - k_1\varphi_x} \varphi_x, \\ B_\sigma &= \frac{(1 - \gamma)(1 - \rho)}{2(1 - k_1\varphi_\sigma)} \left(1 + \frac{1 - \eta}{1 - \gamma} \left(\frac{\phi_x}{1 - k_1\varphi_x} \right)^2 \right) \varphi_\sigma. \end{aligned} \tag{2}$$

By substituting the return on the consumption claim into Equation (3) in the paper, we obtain an expression of the log pricing kernel in terms of the state variables

$$\log \xi_{t,t+1} = s_0 + s_x x_t + s_\sigma (\sigma_t^2 - \sigma^2) - \Lambda_c \sigma_{t+1} w_{t+1}^c - \Lambda_x \phi_x \sigma_{t+1} w_{t+1}^x - \Lambda_\sigma \phi_\sigma w_{t+1}^\sigma,$$

with the drift characterized by the coefficients

$$\begin{aligned} s_0 &= -\delta - \rho \mu_c - \frac{(1-\eta)(\rho-\eta)}{2(1-\rho)^2} \left(\frac{B_\sigma}{\varphi_\sigma} \phi_\sigma \right)^2 + \frac{s_\sigma}{\varphi_\sigma} \sigma^2, \\ s_x &= -\rho \varphi_x, \\ s_\sigma &= \frac{1}{2}(1-\gamma)(\gamma-\rho)\varphi_\sigma + \frac{1}{2}(1-\eta)(\eta-\rho) \left(\frac{\phi_x}{1-k_1\varphi_x} \right)^2 \varphi_\sigma. \end{aligned}$$

The market prices of uncertainty in consumption, expected consumption growth, and volatility are determined by

$$\begin{aligned} \Lambda_c &= \gamma, \\ \Lambda_x &= \frac{\eta-\rho}{1-\rho} k_1 B_x + \eta, \\ \Lambda_\sigma &= \frac{\eta-\rho}{1-\rho} k_1 B_\sigma + \frac{1}{2}(\eta-\gamma)(1-\gamma). \end{aligned} \tag{3}$$

Given the log pricing kernel, the continuously compounded risk-free rate is

$$\begin{aligned} r_{t,t+1}^f &= -\log \mathbb{E}_t \left(e^{\log \xi_{t,t+1}} \right) \\ &= r_0^f + r_x^f x_t + r_\sigma^f (\sigma_t^2 - \sigma^2), \end{aligned}$$

with

$$\begin{aligned}
r_0^f &= -s_0 - \frac{1}{2} \left(\Lambda_\sigma - \frac{1}{2} (\Lambda_c^2 + \Lambda_x^2 \phi_x^2) \right)^2 \phi_\sigma^2 - \frac{1}{2} (\Lambda_c^2 + \Lambda_x^2 \phi_x^2) \sigma^2, \\
r_x^f &= -s_x, \\
r_\sigma^f &= -s_\sigma - \frac{1}{2} (\Lambda_c^2 + \Lambda_x^2 \phi_x^2) \varphi_\sigma.
\end{aligned}$$

To solve for the price-dividend ratio z^d , we rely on the log-linear approximation of the log return on the dividend claim $r_{t,t+1}^d = \log R_{t,t+1}^d$

$$r_{t,t+1}^d = k_0^d + k_1^d z_{t+1}^d - z_t^d + \Delta d_{t+1},$$

where $k_0^d = \log(1 + e^{\bar{z}^d}) - k_1^d \bar{z}^d$, $k_1^d = \frac{e^{\bar{z}^d}}{1 + e^{\bar{z}^d}}$, and $\bar{z}^d$ denotes the long-run mean of the log price-dividend ratio. We conjecture that the log price-dividend ratio z^d is affine in the state variables

$$z_t^d = A^d + B_x^d x_t + B_\sigma^d (\sigma_t^2 - \sigma^2).$$

The coefficients of the log price-dividend ratio follow by applying the Euler equation

to the log return on the dividend claim

$$\begin{aligned}
A^d &= \frac{s_0 + k_0^d + \mu_d + \frac{1}{2} \left(\frac{B_\sigma^d}{\varphi_\sigma} - \frac{s_\sigma}{\varphi_\sigma} - \Lambda_\sigma \right)^2 \phi_\sigma^2 + \left((1 - k_1^d \varphi_\sigma) \frac{B_\sigma^d}{\varphi_\sigma} - \frac{s_\sigma}{\varphi_\sigma} \right) \sigma^2}{(1 - k_1^d)}, \\
B_x^d &= \frac{\lambda - \rho}{1 - k_1^d \varphi_x} \varphi_x, \\
B_\sigma^d &= \frac{s_\sigma + \frac{1}{2} \left((k_1^d B_x^d + \lambda - \Lambda_x)^2 \phi_x^2 + \gamma^2 + \phi_{d,\sigma}^2 - 2\gamma\rho_{cd}\phi_{d,\sigma} \right) \varphi_\sigma}{1 - k_1^d \varphi_\sigma}.
\end{aligned}$$

The conditional expected return on the dividend claim is given by

$$\begin{aligned}
\mathbb{E}_t [r_{t,t+1}^d] &= k_0^d + (k_1^d - 1)A^d + \mu_d \\
&\quad + ((k_1^d \varphi_x - 1)B_x^d + \lambda \varphi_x)x_t + (k_1^d \varphi_\sigma - 1)B_\sigma^d(\sigma_t^2 - \sigma^2).
\end{aligned}$$

The equity premium follows by subtracting the risk-free rate from the expected return on the dividend claim

$$\begin{aligned}
\mathbb{E}_t [r_{t,t+1}^d] - r_{t,t+1}^f &= \frac{1}{2} \left(\left(\Lambda_\sigma - \frac{1}{2}(\Lambda_c^2 + \Lambda_x^2 \phi_x^2) \right)^2 - \left(\Lambda_\sigma - \frac{B_\sigma^d - s_\sigma}{\varphi_\sigma} \right)^2 \right) \phi_\sigma^2 \\
&\quad + \left(\Lambda_x (k_1^d B_x^d + \lambda) \phi_x^2 - \frac{1}{2} (k_1^d B_x^d + \lambda)^2 \phi_x^2 + \Lambda_c \phi_{d,\sigma} \rho_{cd} - \frac{1}{2} \phi_{d,\sigma}^2 \right) \sigma^2 \\
&\quad + \left(\Lambda_x (k_1^d B_x^d + \lambda) \phi_x^2 - \frac{1}{2} (k_1^d B_x^d + \lambda)^2 \phi_x^2 + \Lambda_c \phi_{d,\sigma} \rho_{cd} - \frac{1}{2} \phi_{d,\sigma}^2 \right) \varphi_\sigma (\sigma_t^2 - \sigma^2).
\end{aligned}$$

3.3 Conditional expectation

Using the dynamics in Equation (1) and the coefficients of the wealth-consumption ratio in Equation (2), the conditional expectation of Y_{t+1} equals

$$\begin{aligned}\mathbb{E}_{\pi_{t+1}}[Y_{t+1}] &= \mathbb{E}_{\pi_{t+1}} \left[e^{\theta_1(q - B_x x_t + (k_1 B_x + (1-\rho))x_{t+1} - B_\sigma \sigma_t^2 + k_1 B_\sigma \sigma_{t+1}^2 + (1-\rho)\sigma_{t+1} w_{t+1}^c)} \right] \\ &= e^{\theta_1(q - B_x x_t + (k_1 B_x + (1-\rho))x_{t+1}^\pi - B_\sigma \sigma_t^2 + k_1 B_\sigma (\sigma_{t+1}^2)^\pi)} e^{\frac{1}{2}(1-\gamma)^2(\sigma_{t+1}^2)^\pi},\end{aligned}\tag{4}$$

with $q = -\delta + (1-\rho)\mu_c + k_0 + (k_1 - 1)(A - B_\sigma \sigma^2)$. In this equation, the measure π_{t+1} , i.e. the economic model, is fully characterized by the realizations x_{t+1}^π and $(\sigma_{t+1}^2)^\pi$. Hence, taking expectations conditional on the measure π means conditional on the state variables. In general, imposing affine dynamics on the endowment process in Equation (1) implies that the conditional expectation is an exponential affine function of the underlying state variables.

In order to use the methodology outlined in Section 3.3 of the paper, we need to identify a set of predictor variables whose realizations map one-to-one and onto the set Π_t of economic models. For the endowment process in Equation (1), the vector of state variables and their first lags has this property

$$\log(Y_{t+1}) = \alpha + \beta_1 x_{t+1} + \beta_2 x_t + \beta_3(\sigma_{t+1}^2 - \sigma^2) + \beta_4(\sigma_t^2 - \sigma^2) + (1-\gamma)\sigma_{t+1}\varepsilon_{t+1}.$$

In reality, we are not able to observe the underlying state vector. Thus, it is necessary to identify which observable variables describe the dynamics of the economy.

The risk-free rate and the log price-dividend ratio are observable at the market, show a clear business cycle pattern, and have a long tradition as predictors of stock and bond returns. Cochrane (2005, Ch. 20) provides a detailed review of the literature. Furthermore, in standard affine asset pricing models these variables are approximately affine in the state vector (see Drechsler and Yaron (2011), Lustig et al. (2012), and the references therein). Among others, Constantinides and Ghosh (2011) and Bansal et al. (2012b) exploit this relation to estimate LRR models. In the models they consider, these two variables span the state space. This also holds if the representative investor is ambiguous about the distribution of consumption and dividend growth. They invert the expressions for the log price-dividend ratio and the risk-free rate to express the state variables in terms of observables. Because of their economic relevance and motivated by the relation in affine asset pricing models, we use these two quantities and their first lags as predictor variables.

3.4 Finite sample evidence

Bansal and Yaron (2004) and Bansal et al. (2012a) calibrate their models assuming a monthly decision interval. However, we use quarterly data over the post-war sub-period to estimate the preference parameters in Section 4 of the paper. In order to be consistent with the empirical sampling frequency, we use a quarterly calibration and convert the model parameters from a monthly into a quarterly frequency. Table 2 contains the parameters used in our simulation study.

Preferences			
ρ	δ	γ	η
0.6667	0.0033	0.6667 – 10	10 – 20

Consumption and Dividends				
μ_c	μ_d	λ	$\phi_{d,\sigma}$	ρ_{cd}
0.0045	0.0045	2.5	6.5024	0.40

Long-Run Growth Rate	
φ_x	ϕ_x
0.9269	0.1140

Volatility		
φ_σ	ϕ_σ	σ
0.9970	$4.84 \cdot 10^{-6}$	0.0125

Table 2: Parameters of the LRR Model

The table shows preference and model parameters in quarterly terms. The values of the structural parameters are obtained by converting the monthly values from Bansal et al. (2012a) to a quarterly frequency.

The models are estimated based on simulated consumption and return data with the EIS restricted to 1.5. To simulate data given SA preferences, we choose a low value of the risk aversion coefficient ($\gamma = 5$) and set the ambiguity parameter η to 20. In line with Bansal and Yaron (2004), the risk aversion coefficient γ is assumed to be 10 for EZ preferences. In case of PA preferences, we set $\eta = 20$ and fix $\gamma = \rho = 1/1.5$. The subjective time discount rate is set to 0.0033 for all models. We simulate the models 1,000 times at a quarterly frequency with a sample size approximately equivalent to the actual data (60 years). To investigate the finite

sample behavior, we also estimate the parameters based on longer samples. We start the simulations at the unconditional means of the state variables and discard the first 10 years of each simulated path. In addition to the return on the dividend claim, we generate 30 additional equity test assets to represent the cross-section of returns. The log dividend growth of asset i follows from

$$\Delta d_{t+1}^i = \mu_d^i + \lambda^i x_{t+1} + \sigma_{t+1} \phi_{d,\sigma}^i \left(\rho_{cd}^i w_{t+1}^c + \sqrt{1 - (\rho_{cd}^i)^2} w_{t+1}^i \right),$$

i.e. each of the 30 assets is characterized by a vector $(\mu_d^i, \lambda^i, \phi_{d,\sigma}^i, \rho_{cd}^i)'$. The returns on these assets differ with respect to their means, volatilities, and covariances with the return on the consumption claim, aggregate consumption growth, and returns on the other assets. Bansal et al. (2012b) show that by varying the parameter vectors and hence the distributions of dividend growths of the different claims, the model is able to replicate the cross-sectional differences in return characteristics. We choose a diverse calibration that ensures that the cross-sectional differences within the asset menu used in the empirical exercise in the paper is well-represented.

The results of the estimations are summarized in Table 3. Panels 1 and 2 present the results for the SA model. While Panel 1 shows the median estimated preference parameters and hypotheses tests with the analytic formula for the conditional expectation in Equation (4), Panel 2 displays results for the approximation technique outlined in Section 3.3 of the paper. We report median rather than mean

N	δ	γ	η	R^2	RMSE	HJD	$\gamma = \rho$	$\eta = \gamma$	J-test
Panel 1: SA with R^w and true conditional expectation ($\rho = 0.6667$)									
60	0.003 (0.000)	6.107 (6.528)	19.294 (3.809)	0.603 (-)	0.031 (0.411)	0.406 (0.183)	0.700 (0.403)	2.647 (0.104)	39.751 (0.088)
Panel 2: SA with R^w and approximated conditional expectation ($\rho = 0.6667$)									
60	0.003 (0.000)	6.016 (6.622)	19.245 (3.800)	0.603 (-)	0.031 (0.407)	0.406 (0.185)	0.627 (0.428)	2.626 (0.105)	39.770 (0.088)
Panel 3: EZ with R^w ($\rho = 0.6667$)									
60	0.003 (0.000)	10.080 (2.947)		0.498 (-)	0.030 (0.459)	0.379 (0.220)	10.110 (0.001)		36.929 (0.179)
Panel 4: EZ with R^d ($\rho = 0.6667$)									
60	0.000 (0.010)	1.704 (0.484)		0.165 (-)	0.054 (0.002)	0.538 (0.000)	4.676 (0.031)		45.869 (0.032)
Panel 5: EZ with R^w , data based on SA model ($\rho = 0.6667$)									
60	0.004 (0.000)	14.293 (3.037)		0.435 (-)	0.037 (0.126)	0.441 (0.048)	20.091 (0.000)		46.546 (0.028)
Panel 6: PA with R^w ($\rho = 0.6667$)									
60	0.003 (0.000)	0.667 (-)	19.954 (4.714)	0.497 (-)	0.035 (0.440)	0.418 (0.179)		16.843 (0.000)	41.192 (0.084)
Panel 7: PA with R^w , data based on SA model ($\rho = 0.6667$)									
60	0.003 (0.000)	0.667 (-)	22.224 (4.908)	0.489 (-)	0.036 (0.4044)	0.423 (0.222)		19.475 (0.000)	40.780 (0.091)

Table 3: Parameter Estimates (Simulated Data)

The table shows GMM estimates of the preference parameters. SA refers to the smooth ambiguity model, EZ to Epstein and Zin (1989) preferences, and PA to pure ambiguity preferences. HAC standard errors are in parentheses. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The table also reports the cross-sectional R^2 , the Wald tests of the hypotheses $\gamma = \rho$ and $\eta = \gamma$, and the J-test for overidentifying restrictions (p -values in parentheses). Details on the tests are provided in Section 2. All estimates are based on 1,000 simulation runs of approximately equivalent length to the data ($N = 60$ years of quarterly data, i.e. the sample length is 240). The true parameters are $\Theta^{\text{SA}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 5, \eta = 20)$ for the smooth ambiguity model, $\Theta^{\text{EZ}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 10)$ for Epstein and Zin (1989) preferences, and $\Theta^{\text{PA}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 0.6667, \eta = 20)$ for pure ambiguity preferences.

estimates as means are biased in some cases due to numerical instabilities during the minimization of the GMM objective function.

The median estimated parameters are close to the true values. Furthermore, the approximated conditional expectation delivers almost identical results compared to the true conditional expectation. The standard errors of the estimated parameters are quite large, indicating that it is rather difficult to estimate the risk aversion and ambiguity parameters jointly in small samples. For the hypotheses tests, we report the median test statistics and p -values (in parentheses). The specification tests do not reject the model. Due to the low value of γ , as well as its large standard error, the Wald test does not reject pure ambiguity preferences ($\rho = \gamma$). The large standard error of the risk aversion coefficient also implies that ambiguity neutrality ($\gamma = \eta$) is difficult to reject. Even though the model is true, the median p -value of the Wald test is above 10%.

Panels 3 and 4 of Table 3 report results assuming that the investor has EZ preferences. Panel 3 shows that the median estimated parameter vector $\hat{\Theta}^{\text{EZ}} = (0.667, 0.003, 10.080)$ is very close to the assumed preference parameters. The subjective time discount rate is estimated very precisely, while the standard error is larger for the risk aversion coefficient. The model is not rejected by any of the three specification tests and the Wald test rejects the null hypothesis of time-separability.

We substitute the return on the consumption claim by the return on the dividend claim. In the LRR model, the return on the consumption claim has a mean

of 0.73% and a standard deviation of 1.52% per quarter. In contrast, the mean and volatility of the return on the dividend claim are 1.15% and 8.94%. The standard deviation is about 6 times larger. The correlation between the two return series is 0.53. Panel 4 shows that using the return on the dividend claim as a proxy for the return on wealth biases the parameter estimates dramatically. The median estimated parameter vector is $\hat{\Theta}^{EZ} = (0.667, 0.000, 1.752)$. The subjective time discount rate is close to zero and the estimated risk aversion coefficient is far too low, which is mainly due to the high volatility of the return on the dividend claim. All specification tests reject the model. This indicates that in our model economy, substituting the return on a stock market index for the return on wealth is inappropriate. This finding emphasizes the importance of carefully specifying a proxy for the return on the wealth portfolio. It indicates that relying on a broad stock market index may lead to biased results, if returns on this portfolio are more volatile than the return on the wealth portfolio.

In addition, we investigate the bias if the parameters are estimated on the basis of EZ preferences but the data was generated by the SA model. Our main conclusion from Panel 5 is that risk aversion is estimated with an upward bias if ambiguity aversion is present. $\hat{\gamma}$ may be interpreted as the investor's effective risk aversion which is upward biased in the presence of ambiguity if the investor is ambiguity averse. See the discussion in Section 3.4 of the paper. The p -values of the specification test based on the HJD and the J-test are slightly below 5%. The RMSE

is close to the one in Panel 1. These results show that even if ambiguity aversion is present by definition, it is rather difficult to discriminate between the models solely based on their pricing errors in small samples. Considering the plausibility of the estimated preference parameters is also important.

Panels 6 and 7 report median estimation results under the assumption of PA preferences, first if the PA model is used for simulating data (Panel 6), and if the SA model is used (Panel 7). Based on these estimates, it is not possible to distinguish PA from SA preferences, given that γ is low. This could have been expected as $\gamma = \rho$ could not be rejected in Panel 2. The results in Panel 7 indicate that η is estimated with a slight upward bias, if $\gamma = \rho$ is counterfactually assumed.

Table 4 reports results for the SA model given larger sample sizes N of 100, 200, 500, and 1000 years (results for the other models are available upon request). As expected, the estimated parameters are close to the true values and their standard errors decrease in N . Concerning the hypotheses tests for the SA model, the Wald test for ambiguity neutrality is rejected for a sample size of 100 years and above, while the Wald test rejects pure ambiguity preferences for 500 and 1000 years. The SA model is not rejected by the specification tests. Concerning the finite sample properties of these tests, we are able to draw several conclusions.

Table 5 shows the rejection rates of the specification tests, i.e. the proportion of estimations in which the empirical p -value is smaller than the asymptotic size. It is well-known that the J-test has poor finite sample properties. We observe that in

N	δ	γ	η	R^2	RMSE	HJD	$\gamma = \rho$	$\eta = \gamma$	J-test
100	0.003 (0.000)	5.554 (5.071)	19.845 (2.965)	0.719 (-)	0.025 (0.469)	0.306 (0.265)	0.996 (0.318)	5.216 (0.022)	35.192 (0.198)
200	0.003 (0.000)	5.545 (3.577)	19.967 (2.121)	0.842 (-)	0.017 (0.500)	0.209 (0.363)	1.885 (0.170)	10.540 (0.001)	30.919 (0.369)
500	0.003 (0.000)	5.229 (2.250)	20.152 (1.343)	0.930 (-)	0.011 (0.499)	0.130 (0.422)	4.011 (0.045)	28.341 (0.000)	29.232 (0.453)
1000	0.003 (0.000)	5.126 (1.587)	20.166 (0.951)	0.963 (-)	0.008 (0.488)	0.092 (0.426)	7.781 (0.005)	57.533 (0.000)	28.805 (0.475)

Table 4: Parameter Estimates (Simulated Data and Larger Sample Size)

The table shows GMM estimates of the preference parameters. Data are simulated and parameters are estimated based on the smooth ambiguity model. ρ is fixed at 0.667. We use the technique explained in Section 3.3 in the paper to approximate the conditional expectation. HAC standard errors are in parentheses. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The table also reports the cross-sectional R^2 , the Wald tests for pure ambiguity preferences ($\gamma = \rho$) and ambiguity neutrality ($\eta = \gamma$), and the J-test for overidentifying restrictions (p -values in parentheses). Details on the tests are provided in Section 2. The estimates are based on 1,000 simulation runs with sample size equal to N years (quarterly data, i.e. sample size = $4 \cdot N$). The true parameters are $\Theta^{\text{SA}} = (\rho = 0.6667, \delta = 0.0033, \gamma = 5, \eta = 20)$.

our simulation study the rejection rates are far too large for samples smaller than 500 years. This confirms that the test rejects too often for sample sizes typically used in empirical tests of consumption-based asset pricing models. Similar to Ahn and Gadarowski (2004), we find that the specification test based on the Hansen and Jagannathan (1997) distance also performs poorly in small samples. In contrast to this, the test based on the RMSE behaves superior in finite samples. The rejection rates for a sample size of 60 years differ only slightly from those with 1000 years and are close to the asymptotic values.

Test	N	1%	5%	10%
RMSE	60	1.40	7.40	13.90
HJD	60	10.20	25.40	36.40
J-test	60	23.40	42.50	53.10
RMSE	100	0.70	6.90	12.20
HJD	100	4.80	17.20	26.20
J-test	100	10.10	24.70	33.10
RMSE	200	1.30	5.90	11.00
HJD	200	3.60	11.80	20.50
J-test	200	4.60	14.40	22.00
RMSE	500	2.30	8.20	13.80
HJD	500	3.40	10.20	17.50
J-test	500	3.10	9.60	14.80
RMSE	1000	0.90	5.80	10.90
HJD	1000	2.60	8.60	16.50
J-test	1000	1.70	6.50	12.40

Table 5: Finite Sample Behavior of Specification Tests

The table shows the proportion of the empirical p -values of the specification tests being smaller than the asymptotic size. The smooth ambiguity model is used for simulation and estimation. The RMSE is the square-root of the mean squared Euler equation error. HJD denotes the Hansen and Jagannathan (1997) distance. The J-test refers to Hansen’s test for overidentifying restrictions. Details on the tests are provided in Section 2. All estimates are based on 1,000 simulation runs. The results are reported for three significance levels (1%, 5%, and 10%) and for different values of the sample size N in years (quarterly data, i.e. sample size = $4 \cdot N$).

References

- Ahn, S. and Gadarowski, C. (2004), “Small Sample Properties of the GMM Specification Test Based on the Hansen-Jagannathan Distance.” *Journal of Empirical Finance*, 11, 109–132.
- Bansal, R., Kiku, D., and Yaron, A. (2012a), “An Empirical Evaluation of the Long-Run Risks Model for Asset Prices.” *Critical Finance Review*, 1, 183–221.
- (2012b), “Risks for the Long Run: Estimation with Time Aggregation.” NBER Working Paper 18305.
- Bansal, R. and Yaron, A. (2004), “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.” *Journal of Finance*, 59, 1481–1509.
- Beeler, J. and Campbell, J. (2012), “The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment.” *Critical Finance Review*, 1, 141–182.
- Campbell, J. and Shiller, R. (1988), “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors.” *Review of Financial Studies*, 1, 195–228.
- Cochrane, J. (2005), *Asset Pricing*. New Jersey: Princeton University Press.
- Constantinides, G. and Ghosh, A. (2011), “Asset Pricing Tests with Long-Run Risks in Consumption Growth.” *Review of Asset Pricing Studies*, 1, 96–136.
- Drechsler, I. and Yaron, A. (2011), “What’s Vol Got to Do with It.” *Review of Financial Studies*, 24, 1–45.
- Epstein, L. and Zin, S. (1989), “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework.” *Econometrica*, 57, 937–969.
- Hansen, L. (1982), “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica*, 50, 1029–1084.
- Hansen, L. and Jagannathan, R. (1997), “Assessing Specification Errors in Stochastic Discount Factor Models.” *Journal of Finance*, 52, 557–590.
- Jagannathan, R. and Wang, Z. (1996), “The Conditional CAPM and the Cross-Section of Expected Returns.” *Journal of Finance*, 51, 3–53.
- Lustig, H., Van Nieuwerburgh, S., and Verdelhan, A. (2012), “The Wealth-Consumption Ratio.” *Review of Asset Pricing Studies*, 3, 38–94.
- Newey, W. and West, K. (1987a), “Hypothesis Testing with Efficient Method of Moments Estimation.” *International Economic Review*, 28, 777–787.

- (1987b), “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, 55, 703–708.
- Parker, J. and Julliard, C. (2005), “Consumption Risk and the Cross Section of Expected Returns.” *Journal of Political Economy*, 113, 185–222.
- Yogo, M. (2006), “A Consumption-Based Explanation of Expected Stock Returns.” *Journal of Finance*, 61, 539–580.