Does Smooth Ambiguity Matter for Asset Pricing?

A. Ronald Gallant — Penn State University*

Mohammad R. Jahan-Parvar — Federal Reserve Board†

Hening Liu — University of Manchester‡§

December 2017

Abstract

We use the Bayesian method introduced by Gallant and McCulloch (2009) to estimate consumption-based asset pricing models featuring smooth ambiguity preferences. We rely on semi-nonparametric estimation of a flexible auxiliary model in our structural estimation. Based on the market and aggregate consumption data, our estimation provides statistical support for asset pricing models with smooth ambiguity. Statistical model comparison shows that models with ambiguity, learning and time-varying volatility are preferred to the long-run risk model. We also analyze asset pricing implications of the estimated models.

JEL Classification: C61; D81; G11; G12.

Keywords: Ambiguity, Bayesian estimation, Equity premium, Markov-switching, Long-run risk.

*Department of Economics, Pennsylvania State University, 511 Kern Graduate Building, University Park, PA 16802 U.S.A. e-mail: aronaldg@gmail.com.

†Corresponding Author, Board of Governors of the Federal Reserve System, 20th St. NW and Constitution Ave., Washington, DC 20551 U.S.A. e-mail: Mohammad.Jahan-Parvar@frb.gov.

‡Accounting and Finance Group, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK. e-mail: Hening.Liu@manchester.ac.uk.

§We thank Jaroslav Borovicka, Yoosoon Chang, Pierre Collin-Dufresne, Stefanos Delikouras, Marco Del Negro, Luca Guerrieri, Philipp K. Illeditsch, Shaowei Ke, Nour Meddahi, Thomas Maurer, Jianjun Miao, James Nason, Joon Y. Park, Eric Renault, Michael Stutzer, seminar participants at Federal Reserve Board, George Washington University, Georgetown University, Indiana University, North Carolina State University, University of Colorado, University of Maryland, 2016 SFS Cavalcade, and 2018 Zurich Asset Pricing Workshop for helpful comments and discussions. The analysis and the conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. Any remaining errors are ours.

1 Introduction

Several major asset pricing puzzles have posed a remarkable challenge to the standard consumption-based model with a fully rational representative-agent. One prominent puzzle is the "equity premium puzzle" first documented by Mehra and Prescott (1985), which states that the standard model requires an implausibly high level of risk aversion to explain the historical equity premium in the U.S. data. Other important stylized facts of the stock market include excess return volatility, countercyclical equity premium and equity volatility, and return predictability.[1] A recent strand of the literature proposes to embed ambiguity in an otherwise standard model to explain various asset pricing puzzles. An ambiguity-averse agent recognizes uncertainty about an objective law governing the state process and is averse to such uncertainty. There are two popular approaches to model ambiguity in the asset pricing literature, the multiple-priors approach and the smooth ambiguity approach.[2] Existing consumption-based models with ambiguity are largely confined to model calibration. However, calibration does not provide the likelihood of the model given observed macroeconomic and financial variables, and thus statistical support for the importance of ambiguity in asset pricing is still limited in the literature on estimation of structural models.

In this paper, we use the "General Scientific Models" (henceforth, GSM) Bayesian estimation method developed by Gallant and McCulloch (2009) to estimate a set of three consumption-based asset pricing models with smooth ambiguity preferences. Our Bayesian estimation jointly produces estimates of preferences parameters and parameters governing state dynamics in a structural model. We consider models with an ambiguity-averse representative agent who is uncertain about the conditional mean growth rate of aggregate consumption. The agent's preferences are represented by generalized recursive smooth ambiguity utility advanced by Hayashi and Miao (2011) and Ju and Miao (2012). This class of preferences builds on the seminal work of Klibanoff, Marinacci, and Mukerji (2005, 2009) and allows for the separation among risk aversion, ambiguity aversion and the elasticity of intertemporal substitution (EIS). Altug, Collard, Çakmakli, Mukerji, and Özsöylev

[1] See Shiller (1981), Fama and French (1989), Schwert (1989), and Fama and French (1988a) for related empirical evidence.

[2] See Epstein and Wang (1994), Chen and Epstein (2002), Epstein and Schneider (2008), and Drechsler (2013) for applications of the multiple-priors preferences and Ju and Miao (2012), and Collard, Mukerji, Sheppard, and Tallon (2017) for applications of smooth ambiguity preferences. The smooth ambiguity utility model has a connection with risk-sensitive control and robustness, see Klibanoff, Marinacci, and Mukerji (2009), Hansen (2007) and Hansen and Sargent (2010).

(2017) show that in the special case of the EIS and relative risk aversion being inversely related, Ju and Miao's generalized smooth ambiguity utility function is not equivalent to the preferences proposed by Klibanoff et al. (2005, 2009). Our estimation suggests a clear distinction of the EIS from risk aversion for all models and a preference for early resolution of uncertainty.

We examine three models featuring smooth ambiguity. The first is the original model of Ju and Miao (2012). The growth rate of consumption follows a Markov-switching process in which the mean growth rate depends on a hidden state. The hidden state evolves according to a two-state Markov chain. The agent cannot observe the state but can learn about the state in a Bayesian fashion by observing realized growth rates of consumption. Ambiguity arises since the mean growth rate is unobservable. Because the hidden state evolves dynamically over time, learning cannot resolve the agent's ambiguity in the long run. The agent is ambiguity-averse in that he dislikes a mean-preserving spread in the continuation value led by the agent's belief about the hidden state. As a result, compared with a solely risk-averse agent, the ambiguity-averse agent effectively assigns more probability weight to "bad" states that are associated with lower levels of the continuation values.

The second model is an extension of the first and incorporates time-varying conditional volatility. We postulate that conditional volatility of consumption growth follows another two-state Markov chain that is independent of the chain for the mean growth state, as in McConnell and Perez-Quiros (2000) and Lettau, Ludvigson, and Wachter (2008). A number of studies have examined the role of time-varying volatility and found that volatility risk is significantly priced in the stock market, see Bollerslev, Tauchen, and Zhou (2009), Drechsler (2013) and Bansal, Kiku, Shaliastovich, and Yaron (2014), among others. By estimating a model with ambiguity, learning and time-varying volatility, we aim to investigate whether (1) inclusion of time-varying volatility affects the estimated impact of ambiguity on asset prices, and (2) the model with time-varying volatility represents an improvement over the original model.

The third model is built on the long-run risk model of Bansal and Yaron (2004) and the smooth ambiguity model of Collard et al. (2017). The motivation for examining this model is to study the impact of ambiguity in a long-run risk setting. The persistent long-run risk component in the conditional mean of consumption growth is empirically difficult to detect. Thus, it is reasonable to postulate that the agent also faces the same difficulty as an econometrician does.[3] Similar to the model setup of Ju and Miao (2012), the agent cannot observe the long-run risk component governing mean consumption growth but can learn about it in a Bayesian fashion by observing realizations of consumption and dividend growth rates. In addition, we incorporate stochastic volatility as an exogenous process as in Bansal and Yaron (2004).[4] By estimating this model, we want to investigate to what extent the estimated level of ambiguity aversion depends on specifications of state processes and the information structure.

In all three models, the agent's ambiguity aversion endogenously generates pessimistic beliefs about the distribution of consumption growth. In contrast to an ambiguity-neutral investor, the ambiguity-averse agent always slants his belief toward states with low levels of conditional mean growth of consumption. This pessimism is manifested by a sharp increase in the pricing kernel when the economy experiences a negative shock after staying at the "normal" growth rate for several periods. The pessimistic distortion to the pricing kernel raises its volatility and thus implies a high market price of risk and high equity premium.

In addition to models with smooth ambiguity, we also estimate two baseline models with Epstein and Zin (1989)'s recursive utility for model comparison. The first baseline model is the long-run risk model studied by Bansal, Kiku, and Yaron (2012), which is an improved formulation of the original model of Bansal and Yaron (2004). The second baseline model is a special case of Ju and Miao's model by suppressing ambiguity aversion. In this model, the agent without endogenous pessimism makes Bayesian inference to evaluate mean consumption growth. By estimating a series of structural models with and without smooth ambiguity, we address two important questions: (1) does a structural estimation with macro-finance data lend statistical support to the class of smooth ambiguity preferences that have sound decision-theoretic basis? (2) Based on a standard Bayesian model comparison between those featuring smooth ambiguity and Epsetin-Zin's preferences, which estimated model is the preferred model? We find a significant distinction between risk aversion

[3] Bidder and Dew-Becker (2016) study a related framework where the agent estimates the consumption process non-parameterically and prices assets using a pessimistic model. They find that long-run risks arise endogenously as the worst-case outcome. Collard et al. (2017) consider a more elaborate model in which the agent is not only ambiguous about the latent mean growth rate of consumption but also ambiguous about whether the latent variable comes from a highly persistent process or a moderately persistent process.

[4] Incorporating an unobservable stochastic volatility component together with learning is beyond the scope of our study. We leave estimation of the model in which the agent also has ambiguity about the volatility state for future research.

and ambiguity aversion in the estimated models. Moreover, the distinction is robust to different specifications of consumption dynamics.

A model comparison exercise based on posterior likelihoods and the Bayesian information criteria (BIC) shows that the two models featuring smooth ambiguity and time-varying volatility are preferred to the long-run risk model and the Epstein-Zin's recursive utility model with regime-switching consumption growth and learning. Prior to our study, Bansal, Gallant, and Tauchen (2007) and Aldrich and Gallant (2011) concluded that the long-run risk model is a preferred model. In addition, we find that the estimated smooth ambiguity models can match moments of asset returns better than the Epstein-Zin's models do. The estimated Ju and Miao's model, albeit receives less statistical support than the long-run risk model does, can match the equity premium and variance risk premium in the data well.

We use the projection method to solve all models examined in this paper. The log-linear approximation method that has been widely used in the long-run risk literature is not applicable to our smooth ambiguity models. This is because learning induces nonlinearities in the dynamics of the agent's beliefs. Additionally, the smooth ambiguity utility function is highly nonlinear. To keep our quantitative analysis consistent, we also use the projection method to solve the long-run risk model. In a recent work, Pohl, Schmedders, and Wilms (2017) assess numerical accuracy of the log-linear approximation method and find that applying log-linearization to solve long-run risk models can yield biased results due to neglecting higher-order effects. The bias becomes more pronounced when the long-run risk and stochastic volatility components are highly persistent. Using the log-linear approximation and a mixed data frequency approach, Schorfheide, Song, and Yaron (2017) perform Bayesian estimation of long-run risk models with several specifications of stochastic volatility and find that the long-run risk component and stochastic volatilities are highly persistent. While our estimation is based on annual data and Bayesian indirect inference, we also find persistent long-run risk and stochastic volatility components as well as a high EIS.

Similar to other macro-finance applications, we face sparsity of data because we use annual data for estimation. In addition, the likelihood of a structural asset pricing model is not readily available. As has become standard in the macro-finance empirical literature, we use prior information and a Bayesian estimation methodology to overcome data sparsity. Specifically, we use the GSM Bayesian estimation method developed by Gallant and McCulloch (2009). GSM is the Bayesian counterpart to the classical "indirect inference" and "efficient method of moments" (hereafter, EMM) methods introduced by Gouriéroux, Monfort, and Renault (1993) and Gallant and Tauchen (1996, 1998, 2010). These are simulation-based inference methods that rely on an auxiliary model for implementation. The Bayesian estimation of GSM relies on the theoretical results of Gallant and Long (1997) in its construction of a likelihood. In particular, Gallant and McCulloch synthesize a likelihood by means of an auxiliary model and simulations from the structural model.

A comparison of Aldrich and Gallant (2011) with Bansal et al. (2007) displays the advantages of a Bayesian EMM approach relative to a frequentist EMM approach, particularly for the purpose of model comparison. An indirect inference approach is an appropriate estimation methodology in the context of this study since the estimated equilibrium model is highly nonlinear and does not admit analytically tractable solutions, thereby severely inhibiting accurate numerical construction of a likelihood by means other than GSM. GSM uses a sieve (see Section 3) specially tailored to macroeconomic and financial time-series applications as the auxiliary model. When a suitable sieve is used as the auxiliary model, as in this study, the GSM method synthesizes the exact likelihood implied by the model.[5] In this instance, the synthesized likelihood model departs significantly from a normal-errors likelihood, which suggests that alternative econometric methods based on normal approximations will give biased results. In particular, in addition to the generalized autoregressive conditional heteroscedasticity (GARCH) effect, the four-dimensional error distribution implied by the smooth ambiguity model is skewed in all four components and has fat-tails for consumption growth, dividend growth and stock returns, and thin tails for bond returns.

This paper contributes to a growing body of literature on ambiguity, learning and macro-finance. We discuss closely related papers here. Epstein and Schneider (2007) develop a model with learning under ambiguity. They use the multiple-priors approach to model ambiguity and assume a set of priors and a set of likelihoods for signals. Beliefs are updated by Bayes' rule in an appropriate way. Epstein and Schneider (2008) apply this model to study information quality and asset prices. Leippold, Trojani, and Vanini (2008) adopt the continuous-time multiple-priors framework of Chen

[5] Gallant and McCulloch (2009) use the terms "scientific model" and "statistical model" instead of the terms "structural model" and "auxiliary model" used in the indirect inference econometric literature. We will follow the conventions of the econometric literature. The structural models here are equilibrium asset pricing models.

and Epstein (2002) to analyze asset pricing implications of learning under ambiguity. Cogley and Sargent (2008) examine the impacts of pessimistic beliefs on the market price of risk and equity premium. Hansen and Sargent (2010) consider robustness concerns in learning and study time-varying model uncertainty premia. Collard et al. (2017) assume that a representative agent with smooth ambiguity preferences faces both model uncertainty and state uncertainty and analyze the dynamics of risk premia conditioning on the historical data. Johannes, Lochstoer, and Mou (2016) and Collin-Dufresne, Johannes, and Lochstoer (2016) study parameter learning and asset prices in the consumption-based framework with recursive utility. Jeong, Kim, and Park (2015) estimate an asset pricing model in which the agent has multiple-priors utility. Their estimation results suggest that ambiguity on the true probability law governing fundamentals carries a sizable premium.

Jahan-Parvar and Liu (2014), Backus, Ferriere, and Zin (2015) and Altug et al. (2017) examine both business cycle and asset pricing implications in dynamic stochastic general equilibrium (DSGE) models with smooth ambiguity. Ilut and Schneider (2014) estimate a DSGE model with multiple-priors utility. Their estimation suggests that time-varying confidence in future total factor productivity explains a significant fraction of the business cycle fluctuations. Bianchi, Ilut, and Schneider (2016) estimate another DSGE model to explain joint dynamics of asset prices and real economic activity in the postwar data. They show that time-varying ambiguity about corporate profits leads to high equity premium and excess volatility. They further show that the recursive multiple priors utility model provides a tractable way to analyze DSGE models with time varying uncertainty and facilitates estimation by means of likelihood methods.

The rest of the paper proceeds as follows. Section 2 presents consumption-based asset pricing models with smooth ambiguity. Section 3 discusses the estimation method and empirical findings. Section 4 presents asset pricing implications. Section 5 concludes. Numerical solution methods and additional results are included in the Internet Appendix.

2 Asset Pricing Models

The intuitive notions behind any consumption-based asset pricing model are that agents receive income (wage, interest, and dividends) that they use to purchase consumption goods. Agents reallocate their consumption over time by trading stocks that pay random dividends and bonds that pay interest with certainty. This is done for consumption smoothing over time. The budget constraint implies that the purchase of consumption, bonds, and stocks cannot exceed income in any period. Agents are endowed with a utility function that depends on the entire consumption path. The first-order conditions of their utility maximization deliver an intertemporal relation of prices of stocks and bonds. Among all tradable assets, we focus on the risky asset that pays aggregate dividends and the one-period risk-free bond with zero net supply.

2.1 Asset Pricing Models Featuring Smooth Ambiguity

We examine three consumption-based asset pricing models in which a representative agent is endowed with smooth ambiguity preferences. These models include (1) Ju and Miao (2012)'s model in which the mean of consumption growth follows a hidden Markov chain with two states, abbreviated as "AAMS", (2) an extended version of Ju and Miao's model with time-varying conditional volatility, abbreviated as "AAMSSV", and (3) a long-run risk model featuring ambiguity in which the long-run risk component is assumed to be unobservable, abbreviated as "AALRRSV" model. The latter model shares many features with the models introduced by Collard et al. (2017). In all these models, the agent cannot observe the state determining mean consumption growth but learns about the state in a Bayesian fashion. The unobservable mean growth state implies that the agent is ambiguous about the data-generating process of fundamentals. Smooth ambiguity utility captures the agent's aversion toward this ambiguity.

2.1.1 The AAMS model

Aggregate consumption follows the process

$$\Delta c_t \equiv \ln\left(\frac{C_t}{C_{t-1}}\right) = \mu\left(s_t\right) + \sigma_c \epsilon_{c,t}, \quad \epsilon_{c,t} \sim N\left(0,1\right),$$

where $\epsilon_{c,t}$ is an *i.i.d.* standard normal random variable, and s_t indicates the state of mean consumption growth and follows a two-state Markov chain. Suppose that "l" and "h" indicate low and high mean growth states respectively. The transition probabilities are given by

$$\Pr\left(s_t = l | s_{t-1} = l\right) = p_{ll}, \;\; \Pr\left(s_t = h | s_{t-1} = h\right) = p_{hh}$$

Because aggregate dividends are more volatile than aggregate consumption (see Abel, 1999 and Bansal and Yaron, 2004), the dividend growth process is given by

$$\Delta d_t \equiv \ln\left(\frac{D_t}{D_{t-1}}\right) = \lambda \Delta c_t + g_d + \tilde{\sigma}_d \epsilon_{d,t} \tag{1}$$

where $\epsilon_{d,t}$ is an *i.i.d.* standard normal random variable that is independent of all other shocks in the model. The parameter λ represents the leverage ratio; see Abel (1999). We pin down the parameters g_d and $\tilde{\sigma}_d$ by the estimates of unconditional mean and volatility of dividend growth. We set the unconditional mean of dividend growth to that of consumption growth implied by the Markov-switching model. In addition, we denote the unconditional volatility of dividend growth by σ_d.

The agent cannot observe the mean growth state but can learn about it through observing the history of consumption and dividends. The agent knows the parameters in the consumption and dividend processes, namely, $\{\mu_l, \mu_h, p_{ll}, p_{hh}, \sigma_c, \lambda, g_d, \tilde{\sigma}_d\}$. Suppose that the agent's belief is $\pi_t = \Pr(s_{t+1} = h | \mathcal{I}_t)$ where $\mathcal{I}_t$ denotes information available at time t. With respect to learning about the unobservable state, dividends do not contain additional information compared to consumption. As a result, given the prior belief π_0 and full information, the agent updates his beliefs according to Bayes' rule:

$$\pi_{t+1} = \frac{p_{hh} f(\Delta c_{t+1} | s_{t+1} = h)\pi_t + (1 - p_{ll}) f(\Delta c_{t+1} | s_{t+1} = l)(1 - \pi_t)}{f(\Delta c_{t+1} | s_{t+1} = h)\pi_t + f(\Delta c_{t+1} | s_{t+1} = l)(1 - \pi_t)}$$

where $f(\Delta c_t | s_t)$ is conditional density with mean $\mu(s_t)$ and variance σ_c^2:

$$f(\Delta c_t | s_t) \propto \exp\left[-\frac{(\Delta c_t - \mu(s_t))^2}{2\sigma_c^2}\right].$$

The generalized recursive smooth ambiguity utility function proposed by Hayashi and Miao (2011) and Ju and Miao (2012) implies that given consumption plans $C = (C_t)_{t \geq 0}$ the value function $V_t = V(C; \pi_t)$ is given by

$$V_t(C; \pi_t) = \left[(1-\beta) C_t^{1-1/\psi} + \beta \left\{\mathcal{R}_t\left(V_{t+1}(C; \pi_{t+1})\right)\right\}^{1-1/\psi}\right]^{\frac{1}{1-1/\psi}},$$

where $\beta \in (0,1)$ is the subjective discount factor, ψ is the elasticity of intertemporal substitution (EIS) parameter, γ is the coefficient of relative risk aversion, and $\mathcal{R}_t\left(V\left(C_{t+1};\pi_{t+1}\right)\right)$ is the certainty equivalent of the continuation value given by

$$\mathcal{R}_t\left(V_{t+1}\left(C;\pi_{t+1}\right)\right) = \left(\mathbb{E}_{\pi_t}\left[\left(\mathbb{E}_{\{s_{t+1},,t\}}\left[V_{t+1}\left(C;\pi_{t+1}\right)^{1-\gamma}\right]\right)^{\frac{1-\eta}{1-\gamma}}\right]\right)^{\frac{1}{1-\eta}}. \tag{2}$$

Ambiguity aversion is characterized by the parametric restriction $\eta > \gamma$, where η is the ambiguity aversion parameter. By setting $\eta = \gamma$, we obtain Epstein-Zin's recursive utility under ambiguity neutrality.[6] In the certainty equivalent (2), the expectation operator $\mathbb{E}_{s_{t+1},t}[\cdot]$ is taken with respect to the conditional distribution of consumption growth in state s_{t+1} and all other information at time t. The expectation operator $\mathbb{E}_{\pi_t}$ is taken with respect to the posterior belief about the unobservable state.

Following Hayashi and Miao (2011), the stochastic discount factor (SDF) in this model is given by

$$M_{t,t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-1/\psi}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{1/\psi-\gamma}\left(\frac{\left(\mathbb{E}_{\{s_{t+1},t\}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma).}$$

The last multiplicative term in the SDF arises due to ambiguity aversion. This term makes the SDF more countercyclical than in the case of Epstein and Zin's recursive utility and induces large variations in the SDF. The risk-free rate, R_t^f, is the reciprocal of the conditional expectation of the SDF,

$$R_t^f = \frac{1}{\mathbb{E}_t\left[M_{t,t+1}\right]}.$$

Stock returns, defined by $R_{t+1} = \dfrac{P_{t+1}+D_{t+1}}{P_t}$, satisfy the Euler equation

$$\mathbb{E}_t\left[M_{t,t+1}R_{t+1}\right] = 1.$$

[6] We follow Ju and Miao (2012) and do not consider $\eta < \gamma$ in our estimation as this parametric restriction might imply "ambiguity loving", see also Hayashi and Miao (2011).

We rewrite the Euler equation as

$$0=\tilde{\pi}_t\mathbb{E}_{h,t}\left[M^{EZ}_{t,t+1}\left(R_{t+1}-R^f_t\right)\right]+\left(1-\tilde{\pi}_t\right)\mathbb{E}_{l,t}\left[M^{EZ}_{t,t+1}\left(R_{t+1}-R^f_t\right)\right],$$

where $\mathbb{E}_{h,t}\left[\cdot\right]$ denotes $\mathbb{E}_{s_{t+1},t}\left[\cdot\right]$ for $s_{t+1}=h$ and similarly for state l. We interpret the term $M^{EZ}_{t,t+1}$ as the SDF under recursive utility:

$$M^{EZ}_{z_{t+1},t+1}=\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\frac{1}{\psi}-\gamma}.$$

We interpret $\tilde{\pi}_t$ as the ambiguity-distorted belief and represent it by:

$$\tilde{\pi}_t=\frac{\pi_t\left(\mathbb{E}_{h,t}\left[V^{1-\gamma}_{t+1}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\pi_t\left(\mathbb{E}_{h,t}\left[V^{1-\gamma}_{t+1}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}+\left(1-\pi_t\right)\left(\mathbb{E}_{l,t}\left[V^{1-\gamma}_{t+1}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}.$$

As long as $\eta>\gamma$, distorted beliefs are not equivalent to Bayesian beliefs. The distortion driven by ambiguity aversion is an equilibrium outcome and implies pessimistic beliefs; see Section 4.

We rewrite the Euler equation to solve for the price-dividend ratio,

$$\frac{P_t}{D_t}=\mathbb{E}_t\left[M_{t,t+1}\left(1+\frac{P_{t+1}}{D_{t+1}}\right)\frac{D_{t+1}}{D_t}\right].$$

Since $\frac{P_t}{D_t}$ is a functional of the state variable π_t, $\frac{P_t}{D_t}=\Phi\left(\pi_t\right)$, the Euler equation becomes

$$\Phi\left(\pi_t\right)=\mathbb{E}_t\left[M_{t,t+1}\left(1+\Phi\left(\pi_{t+1}\right)\right)\exp\left(\Delta d_{t+1}\right)\right].$$

2.1.2 The AAMSSV model

We follow McConnell and Perez-Quiros (2000) and Lettau et al. (2008) and extend Ju and Miao's model by incorporating time-varying conditional volatility. We assume that the conditional mean and volatility states follow two independent Markov chains. The consumption process takes the form

$$\Delta c_t=\mu\left(s^\mu_t\right)+\sigma\left(s^\sigma_t\right)\epsilon_{c,t},\quad \epsilon_{c,t}\sim N\left(0,1\right)$$

with transition probabilities

$$\begin{aligned}
\Pr\left(s_t^\mu = l | s_{t-1}^\mu = l\right) &= p_{ll}^\mu,\ \Pr\left(s_t^\mu = h | s_{t-1}^\mu = h\right) = p_{hh}^\mu, \\
\Pr\left(s_t^\sigma = l | s_{t-1}^\sigma = l\right) &= p_{ll}^\sigma,\ \Pr\left(s_t^\sigma = h | s_{t-1}^\sigma = h\right) = p_{hh}^\sigma \ .
\end{aligned}$$

To ease the analysis, we assume that the mean state s_t^μ is unobservable while the volatility state s_t^σ is observable. We make this simplifying assumption for two reasons: First, empirical studies such as Bryzgalova and Julliard (2015) have established that estimation and characterization of the mean consumption growth is more difficult than consumption volatility. Second, according to the existing literature, volatility states are very persistent, leading to filtered probabilities of the volatility state close to 1. These results imply that ambiguity has little room with respect to consumption volatility states.[7]

The agent updates beliefs according to Bayes' rule as

$$\pi_{t+1} = \frac{p_{hh}^\mu f\left(\Delta c_{t+1} | s_{t+1}^\mu = h, s_{t+1}^\sigma\right)\pi_t + \left(1 - p_{ll}^\mu\right) f\left(\Delta c_{t+1} | s_{t+1}^\mu = l, s_{t+1}^\sigma\right)(1 - \pi_t)}{f\left(\Delta c_{t+1} | s_{t+1}^\mu = h, s_{t+1}^\sigma\right)\pi_t + f\left(\Delta c_{t+1} | s_{t+1}^\mu = l, s_{t+1}^\sigma\right)(1 - \pi_t)}$$

where $f\left(\Delta c_{t+1} | s_{t+1}^\mu, s_{t+1}^\sigma\right)$ is conditional density

$$f\left(\Delta c_{t+1} | s_{t+1}^\mu, s_{t+1}^\sigma\right) \propto \frac{1}{\sigma\left(s_{t+1}^\sigma\right)} \exp\left[-\frac{\left(\Delta c_{t+1} - \mu\left(s_{t+1}^\mu\right)\right)^2}{2\sigma\left(s_{t+1}^\sigma\right)^2}\right]$$

The value function is given by

$$\begin{aligned}
V_t\left(C; \pi_t, s_t^\sigma\right) &= \left[(1-\beta) C_t^{1-1/\psi} + \beta\left\{\mathcal{R}_t\left(V_{t+1}\left(C; \pi_{t+1}, s_{t+1}^\sigma\right)\right)\right\}^{1-1/\psi}\right]^{\frac{1}{1-1/\psi}}, \\
\mathcal{R}_t\left(V_{t+1}\left(C; \pi_{t+1}, s_{t+1}^\sigma\right)\right) &= \left(\mathbb{E}_{\pi_t}\left[\left(\mathbb{E}_{\left\{s_{t+1}^\mu, s_t^\sigma, t\right\}}\left[V_{t+1}\left(C; \pi_{t+1}, s_{t+1}^\sigma\right)^{1-\gamma}\right]\right)^{\frac{1-\eta}{1-\gamma}}\right]\right)^{\frac{1}{1-\eta}}
\end{aligned}$$

in which $\mathbb{E}_{\left\{s_{t+1}^\mu, s_t^\sigma, t\right\}}[\cdot]$ denotes the expectation conditional on the history up to time t including the volatility state s_t^σ, and a probability distribution of consumption growth given state s_{t+1}^μ. The

[7] We have also examined the model in which both the conditional mean and volatility states are unobservable. But solving the model requires substantial run time to achieve convergence. For some parameter values, the numerical algorithm fails to locate a fixed point for the wealth-consumption ratio. These difficulties make our Bayesian MCMC estimation infeasible. Lettau et al. (2008) also point out the convergence issue for Epstein and Zin's recursive utility.

conditional expectation can be explicitly written as

$$\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t},t\}}\left[V^{1-\gamma}_{t+1}\right]=\begin{cases} p^{\sigma}_{ll}\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t+1},t\}}\left[V^{1-\gamma}_{t+1}|s^{\sigma}_{t+1}=l\right]+\left(1-p^{\sigma}_{ll}\right)\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t+1},t\}}\left[V^{1-\gamma}_{t+1}|s^{\sigma}_{t+1}=h\right], s^{\sigma}_{t}=l \\ \left(1-p^{\sigma}_{hh}\right)\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t+1},t\}}\left[V^{1-\gamma}_{t+1}|s^{\sigma}_{t+1}=l\right]+p^{\sigma}_{hh}\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t+1},t\}}\left[V^{1-\gamma}_{t+1}|s^{\sigma}_{t+1}=h\right], s^{\sigma}_{t}=h \end{cases}$$

where

$$\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t+1},t\}}\left[V^{1-\gamma}_{t+1}\right]\propto\int\frac{1}{\sigma\left(s^{\sigma}_{t+1}\right)}\exp\left(-\frac{\left(\Delta c_{t+1}-\mu\left(s^{\mu}_{t+1}\right)\right)^{2}}{2\sigma\left(s^{\sigma}_{t+1}\right)^{2}}\right)V^{1-\gamma}_{t+1}d\left(\Delta c_{t+1}\right).$$

The SDF in this model is

$$M_{t,t+1}=\beta\left(\frac{C_{t+1}}{C_t}\right)^{-1/\psi}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{1/\psi-\gamma}\left(\frac{\left(\mathbb{E}_{\{s^{\mu}_{t+1},s^{\sigma}_{t},t\}}\left[V^{1-\gamma}_{t+1}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}$$

The dividend growth process is specified in the same form as in the AAMS model, i.e., in equation (1). Stock returns and the risk-free rate are defined in a similar way accordingly. The price-dividend ratio ($\frac{P_t}{D_t}=\Phi\left(\pi_t,s^{\sigma}_{t}\right)$) satisfies the Euler equation

$$\Phi\left(\pi_t,s^{\sigma}_{t}\right)=\mathbb{E}_t\left[M_{t,t+1}\left(1+\Phi\left(\pi_{t+1},s^{\sigma}_{t+1}\right)\right)\exp\left(\Delta d_{t+1}\right)\right]$$

2.1.3 The AALRRSV model

We consider the long-run risk model of Bansal and Yaron (2004), the specification of which is given by

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c+x_{t+1}+\sigma_t\epsilon_{c,t+1} \\
\Delta d_{t+1} &= \mu_d+\lambda x_{t+1}+\varphi_d\sigma_t\epsilon_{d,t+1} \\
x_{t+1} &= \rho_x x_t+\varphi_x\sigma_t\epsilon_{x,t+1} \\
\sigma^2_{t+1} &= \mu^2_{\sigma}+\rho_s\left(\sigma^2_t-\mu^2_s\right)+\sigma_w\epsilon_{w,t+1} \\
\epsilon_{c,t+1},\epsilon_{d,t+1},\epsilon_{x,t+1},\epsilon_{w,t+1} &\sim i.i.d.N\left(0,1\right).
\end{aligned}$$

In Bansal and Yaron's calibration, x_t is a highly persistent component, and σ_t is the stochastic volatility component representing time-varying economic uncertainty that is also highly persistent. The long-run risks literature assumes that x_t is fully observable and thus appears as a state variable in the wealth-consumption ratio and price-dividend ratio. However, this component is difficult to identify using empirically observed economic variables as documented by Bansal et al. (2007), Ma (2013), and Johannes et al. (2016), among others. The difficulty in estimating x_t gives rise to the agent's ambiguity about the mean consumption growth. As a result, we adopt a more plausible information structure by assuming that x_t is unobservable. Collard et al. (2017) provide ample theoretical support for this assumption.

In particular, we maintain that the agent observes the realizations of Δc_{t+1} and Δd_{t+1} contemporaneously but never observes the realization of x_t or $(\epsilon_{c,t}, \epsilon_{d,t}, \epsilon_{x,t})$. This feature of the model characterizes ambiguity, i.e., the agent's lack of confidence in estimating the conditional mean of consumption growth. Instead, the agent uses consumption and dividend growth realizations to filter the unobserved long-run risk component x_t. To make the model tractable and comparable to the long-run risks model, we assume that the conditional volatility of consumption growth σ_t is observable. We also assume that values of the parameter vector $(\mu_c, \mu_d, \varphi_c, \varphi_d, \varphi_x, \rho_x, \lambda, \mu_s, \rho_s, \sigma_w)$ are known to the agent.

Suppose that x_0 has a Gaussian distribution. The standard Kalman filter implies that the agent updates beliefs according to Bayes' rule conditional on the history of realizations of Δc_{t+1} and Δd_{t+1} given the Gaussian prior. The updated belief is also Gaussian with mean $\hat{x}_{t+1}$ and variance ν_{t+1}, i.e., $x_{t+1} \sim N(\hat{x}_{t+1}, \nu_{t+1})$. We define $\hat{x}_{t+1|t} = E[x_{t+1}|\mathcal{I}_t]$ and $\nu_{t+1|t} = E\left[\left(x_{t+1} - \hat{x}_{t+1|t}\right)^2 |\mathcal{I}_t\right]$. It immediately follows that

$$\hat{x}_{t+1|t} = \rho_x \hat{x}_t, \text{ and } \nu_{t+1|t} = \rho_x^2 \nu_t + \varphi_x^2 \sigma_t^2.$$

The Kalman filter implies the following updating equations

$$\begin{aligned}
\hat{x}_{t+1} &= \hat{x}_{t+1|t} + \nu_{t+1|t} \begin{bmatrix} 1 & \lambda \end{bmatrix} F_{t+1|t}^{-1} \begin{bmatrix} v_{t+1|t}^{c} \\ v_{t+1|t}^{d} \end{bmatrix} \\
\nu_{t+1} &= \nu_{t+1|t} - \nu_{t+1|t}^{2} \begin{bmatrix} 1 & \lambda \end{bmatrix} F_{t+1|t}^{-1} \begin{bmatrix} 1 & \lambda \end{bmatrix}'
\end{aligned}$$

where $F_{t+1|t}$ is given by

$$F_{t+1|t} = \begin{bmatrix} \nu_{t+1|t} + \sigma_t^2 & \lambda \nu_{t+1|t} \\ \lambda \nu_{t+1|t} & \lambda^2 \nu_{t+1|t} + \varphi_d^2 \sigma_t^2 \end{bmatrix}$$

and the innovation vector $\begin{bmatrix} v_{t+1|t}^{c} & v_{t+1|t}^{d} \end{bmatrix}$ is given by

$$\begin{bmatrix} v_{t+1|t}^{c} \\ v_{t+1|t}^{d} \end{bmatrix} = \begin{bmatrix} \Delta c_{t+1} - \mu_c - \rho_x \hat{x}_t \\ \Delta d_{t+1} - \mu_d - \lambda \rho_x \hat{x}_t \end{bmatrix}.$$

This model has three state variables $(\hat{x}_t, \nu_t, \sigma_t)$. The value function under smooth ambiguity utility $V_t = V_t(C; \hat{x}_t, \nu_t, \sigma_t)$ satisfies

$$\begin{aligned}
V_t &= \left[(1-\beta) C_t^{1-1/\psi} + \beta \left\{\mathcal{R}_t\left(V_{t+1}\right)\right\}^{1-1/\psi}\right]^{\frac{1}{1-1/\psi}}, \\
\mathcal{R}_t\left(V_{t+1}\right) &= \left(\mathbb{E}_{\{\hat{x}_t, \nu_t\}}\left[\left(\mathbb{E}_{\{x_t, \sigma_t, t\}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1-\eta}{1-\gamma}}\right]\right)^{\frac{1}{1-\eta}}.
\end{aligned}$$

The certainty equivalent $\mathcal{R}_t\left(V_{t+1}\right)$ reflects the agent's aversion toward ambiguity in estimating the long-run risk component x_t. The agent lacks confidence in the Gaussian posterior of x_t and thus applies pessimistic distortion to the posterior. This distortion is visible in Figure 1. In what follows, we describe the mechanism of how ambiguity aversion leads to distortion in the posterior.

The SDF in this model is

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}} \left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\frac{1}{\psi}-\gamma} \left(\frac{\left(\mathbb{E}_{\{x_t, \sigma_t, t\}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}.$$

We solve the price-dividend ratio, $\frac{P_t}{D_t} = \Phi\left(\hat{x}_t, \nu_t, \sigma_t\right)$, from the Euler equation

$$\Phi\left(\hat{x}_t, \nu_t, \sigma_t\right) = \mathbb{E}_t\left[M_{t,t+1}\left(1 + \Phi\left(\hat{x}_{t+1}, \nu_{t+1}, \sigma_{t+1}\right)\right)\exp\left(\Delta d_{t+1}\right)\right].$$

Given the Gaussian posterior obtained according to Bayes' rule, $x_t \sim N\left(\hat{x}_t, \nu_t\right)$, we derive the distorted density of x_t due to ambiguity aversion. The SDF $M_{t,t+1}$ can be decomposed as $M_{t,t+1} = M_{t,t+1}^{EZ} M_t^{AA}$ in which $M_{t,t+1}^{EZ}$ and M_t^{AA} are given respectively by

$$M_{t,t+1}^{EZ} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\frac{1}{\psi}-\gamma}, M_t^{AA} = \left(\frac{\left(\mathbb{E}_{\{x_t,\sigma_t,t\}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}.$$

The Euler equation can be rewritten as

$$0 = \mathbb{E}_t\left[M_{t,t+1}^{EZ}\left(R_{t+1} - R_t^f\right)\left(\frac{\left(\mathbb{E}_{\{x_t,\sigma_t,t\}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}\right].$$

By the law of iterated expectations, we obtain

$$0 = \int \mathbb{E}_t\left[M_{t,t+1}^{EZ}\left(R_{t+1} - R_t^f\right)|x_t\right] \frac{\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}|x_t\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}} f\left(x_t|\hat{x}_t, \nu_t\right)}{\int\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}|x_t\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}} f\left(x_t|\hat{x}_t, \nu_t\right) dx_t} dx_t \tag{3}$$

where $f\left(x_t|\hat{x}_t, \nu_t\right)$ denotes the Bayesian density of x_t given $\hat{x}_t$ and ν_t. It is clear from (3) that the distorted density driven by ambiguity, $\tilde{f}\left(x_t|\hat{x}_t, \nu_t, t\right)$, is given by

$$\tilde{f}\left(x_t|\hat{x}_t, \nu_t, t\right) = \frac{\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}|x_t\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\int\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}|x_t\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}} f\left(x_t|\hat{x}_t, \nu_t\right) dx_t} f\left(x_t|\hat{x}_t, \nu_t\right)$$

2.2 Alternative Models Featuring Ambiguity Neutral Preferences

The recursive utility function of Epstein and Zin (1989) takes the form

$$V_t(C) = \left[(1-\beta) C_t^{1-1/\psi} + \beta \left\{ \mathbb{E}_t \left(V_{t+1}(C)^{1-\gamma} \right) \right\}^{1-1/\psi} \right]^{\frac{1}{1-1/\psi}},$$

As usual, the SDF under recursive utility, denoted by M_{t+1}^{EZ}, is

$$M_{t,t+1}^{EZ} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left(\frac{V_{t+1}}{\mathbb{E}_t (V_{t+1})} \right)^{\frac{1}{\psi} - \gamma}. \tag{4}$$

By setting $\eta = \gamma$ in the generalized recursive smooth ambiguity utility function, we suppress ambiguity aversion and obtain Epstein-Zin's recursive utility model as a special case. We impose this parametric restriction to obtain model "EZMS" as the ambiguity-neutral version of model AAMS.

The second alternative model is the long-run risk model of Bansal et al. (2012), which we label as "EZLRRSV". The model specification is

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \epsilon_{c,t+1} \\
\Delta d_{t+1} &= \mu_d + \lambda x_{t+1} + \varphi_d \sigma_t \epsilon_{d,t+1} + \varphi_c \sigma_t \epsilon_{c,t+1} \\
x_{t+1} &= \rho_x x_t + \varphi_x \sigma_t \epsilon_{x,t+1} \\
\sigma_{t+1}^2 &= \mu_\sigma^2 + \rho_\sigma \left(\sigma_t^2 - \mu_\sigma^2 \right) + \sigma_w \epsilon_{w,t+1} \\
\epsilon_{c,t+1}, \epsilon_{d,t+1}, \epsilon_{x,t+1}, \epsilon_{w,t+1} &\sim i.i.d.N(0,1).
\end{aligned}$$

with notations defined in the same way as in model AALRRSV. The two state variables are x_t and σ_t^2. The price-dividend ratio, $\frac{P_t}{D_t} = \Phi\left(x_t, \sigma_t^2\right)$, satisfies the Euler equation

$$\Phi\left(x_t, \sigma_t^2\right) = \mathbb{E}_t \left[M_{t,t+1} \left(1 + \Phi\left(x_t, \sigma_t^2\right) \right) \exp\left(\Delta d_{t+1}\right) \right].$$

We present the structural parameters to be estimated for each model in Table 2. In estimating models AALRRSV and EZLRRSV, we impose that $\mu_c = \mu_d$. We solve all the models examined in this paper using the collocation projection method with Chebyshev polynomials. Pohl et al. (2017)

show that this is a reliable solution method for nonlinear asset pricing models. The details of the implementation and numerical accuracy assessment are available in the Internet Appendix.

3 Data and the Estimation Method

3.1 Data

Throughout this paper, lower case denotes the natural logarithm of an upper case variable; e.g., $c_t = \ln(C_t)$, where C_t is the observed consumption in period t, and $d_t = \ln(D_t)$, where D_t is dividends paid in period t. Similarly, we use logarithmic risk-free interest rate (r_t^f) and aggregate equity market return inclusive of dividends ($r_t = \ln(P_t + D_t) - \ln P_{t-1}$) in the analysis, where P_t is the stock price in period t.

We use real annual data from 1941 to 2015. The sample period 1941–1949 provides initial lags for the recursive parts of our estimation and the sample period 1950–2015 yields estimation results and diagnostics. Our measure for the risk-free rate is the one-year U.S. Treasury Bill rate. To construct the real risk-free rate, we regress the ex-post real one-year Treasury Bill yield on the nominal rate and past annual inflation, available from Wharton Research Data Services (WRDS) Treasury and Inflation database. The fitted values from this regression are the proxy for the ex-ante real interest rate. Using other estimates of expected inflation to construct the real rate does not lead to significant changes in our results. Our proxy for risky assets is the value-weighted returns (including dividends) on the aggregate stock market portfolio of the NYSE/AMEX/NASDAQ, which is obtained from the Center for Research in Security Prices (CRSP) and deflated using the CPI data. We use the sum of real nondurable and services consumption, items 16 and 17 on the NIPA Table 7.1 "Selected Per Capita Product and Income Series in Current and Chained Dollars," published by the Bureau of Economic Analysis (BEA) as our measure of real consumption. These values are reported in chained 2009 U.S. Dollars and constructed using mid-year population data. We construct the dividend growth rate series by first computing the gross dividend level from the value-weighted returns including and excluding dividends and lagged index levels. We then obtain the real dividend growth rate by deflating the nominal growth rate.

Table 1 presents the summary statistics of the data used in estimation. The p-values of Jarque and Bera (1980) test of normality imply that the assumption of normality is not rejected for the consumption growth series, but it is rejected for all other variables. Real equity returns, interest rates, and dividend growth rates all exhibit negative skewness. In addition, both real interest rates and dividend growth rates show significant excess kurtosis. Figure 2 plots the data.

3.2 GSM: Estimation of the structural model

To estimate model parameters we use a Bayesian method proposed by Gallant and McCulloch (2009), abbreviated GM hereafter, which they termed General Scientific Models (GSM). The GSM methodology was refined in Aldrich and Gallant (2011), abbreviated AG hereafter.[8] The discussion here incorporates those refinements and is to a considerable extent a paraphrase of AG.

GSM is a Bayesian counterpart of the indirect inference method of Gouriéroux, Monfort, and Renault (1993) and the efficient method of moments of Gallant and Tauchen (1996, 1998). As such, implementing this estimation method requires fitting the data with an over-parameterized auxiliary model (not rooted in theory) and then recovering parameter estimates from the structural model (founded on theory) by computing the mapping linking the parameter spaces of these two models. We discuss the estimation method pertaining to the structural model and the map in detail, and then discuss the auxiliary model and its estimation briefly.

Let the transition density of a structural model be denoted by

$$p(y_t|z_{t-1}, \theta), \quad \theta \in \Theta,$$

where y_t is the vector of observable variables, $z_{t-1} = (y_{t-1}, \ldots, y_{t-L})$ if Markovian and $z_{t-1} = (y_{t-1}, \ldots, y_1)$ if not, and Θ is the structural parameter space. As a result, z_{t-1} serves as a shorthand for lag-lengths that are generally greater than 1. Thus, transition densities may depend on L-lags of the data (if Markovian) or the entire history of observations (if non-Markovian). There are five structural models under consideration in this application: the three models featuring smooth ambiguity and the two alternative models with Epstein-Zin's recursive utility, all of which are Markovian and described in Section 2.

[8] The code implementing the method with AG refinements, together with a User's Guide, is in the public domain at http://www.aronaldg.org/webfiles/gsm.

We presume that there is no straightforward algorithm for computing the likelihood but that we can simulate data from $p(\cdot|\cdot,\theta)$ for a given $\theta \in \Theta$. We presume that simulations from the structural model are ergodic. We assume that there is a transition density f (called the auxiliary model)

$$f(y_t|z_{t-1},\omega), \quad \omega \in \Omega$$

and Ω is the auxiliary model parameter space. In addition, we assume that a map exists

$$g : \theta \mapsto \omega$$

such that

$$p(y_t|z_{t-1},\theta) = f(y_t|z_{t-1},g(\theta)), \qquad \theta \in \Theta. \tag{5}$$

We assume that $f(y_t|z_{t-1},\omega)$ and its gradient $(\partial/\partial\omega)f(y_t|z_{t-1},\omega)$ are fairly easy to evaluate. Then g is called the "implied map".[9] When Equation (5) holds, f is said to "nest" p. Whenever we need the likelihood $\prod_{t=1}^{n} p(y_t|z_{t-1},\theta)$, we use

$$\mathcal{L}(\theta) = \prod_{t=1}^{n} f(y_t|z_{t-1},g(\theta)), \tag{6}$$

where $\{y_t, z_{t-1}\}_{t=1}^{n}$ are the data and n is the sample size. After substituting $\mathcal{L}(\theta)$ for $\prod_{t=1}^{n} p(y_t|z_{t-1},\theta)$, standard Bayesian MCMC methods become applicable. That is, we have a likelihood $\mathcal{L}(\theta)$ from Equation (6) and a prior $\xi(\theta)$ from Subsection 3.5 that are sufficient for us to implement Bayesian methods by means of MCMC. A good introduction to these methods is Gamerman and Lopes (2006).

The difficulty in implementing GM's proposal is to compute the implied map g accurately enough that the accept/reject decision in an MCMC chain (Step 5 in the algorithm below) is correct when f is a nonlinear model. The algorithm proposed by AG to address this difficulty is described next.

[9] Gouriéroux, Monfort, and Renault (1993), Gallant and Tauchen (1996), Gallant and McCulloch (2009), and Gallant and Tauchen (2010) provide rigorous support for conditions ensuring that the auxiliary model f is a good approximation for the structural model p.

Given θ, $\omega = g(\theta)$ is computed by minimizing Kullback-Leibler divergence

$$d(f,p) = \iint [\log p(y|z,\theta) - \log f(y|z,\omega)]\; p(y|z,\theta)\, dy\, p(z|\theta)\, dz$$

with respect to ω. The advantage of Kullback-Leibler divergence over other distance measures is that the part that depends on the unknown $p(\cdot|\cdot,\theta)$, $\iint \log p(y|z,\theta)\, p(y|z,\theta)\, dy\, p(z|\theta)\, dz$, does not have to be computed to solve the minimization problem. We approximate the integral that must be computed by

$$\iint \log f(y|z,\omega)\, p(y|z,\theta)\, dy\, p(z|\theta)\, dx \approx \frac{1}{N}\sum_{t=1}^{N} \log f(\hat{y}_t|\hat{z}_{t-1},\omega),$$

where $\{\hat{y}_t, \hat{z}_{t-1}\}_{t=1}^{N}$ is a simulation of length N from $p(\cdot|\cdot,\theta)$. Upon dropping the division by N, the implied map is computed as

$$g : \theta \mapsto \operatorname*{argmax}_{\omega} \sum_{t=1}^{N} \log f(\hat{y}_t \mid \hat{z}_{t-1}, \omega). \tag{7}$$

We use $N = 1000$ in the estimation of all the five models. Results (posterior means, posterior standard deviations, etc.) are not sensitive to N; doubling N makes no difference other than doubling computational time. It is essential that the same seed be used to start these simulations so that the same θ always produces the same simulation.

GM run a Markov chain $\{\omega_t\}_{t=1}^{K}$ of length K to compute $\hat{\omega}$ that solves expression (7). There are two other Markov chains discussed below and so this chain is called the ω-subchain to distinguish among them. While the ω-subchain must be run to provide the scaling for the model assessment method that GM propose, the $\hat{\omega}$ that corresponds to the maximum of $\sum_{t=1}^{N} \log f(\hat{y}_t \mid \hat{z}_{t-1}, \omega)$ over the ω-subchain is not a sufficiently accurate evaluation of $g(\theta)$ for our auxiliary model. This is mainly because our auxiliary model is a multivariate GARCH specification of Bollerslev (1986) that Engle and Kroner (1995) call BEKK. Likelihoods incorporating BEKK are notoriously difficult to optimize. AG use $\hat{\omega}$ as a starting value and maximize the expression (7) using the BFGS algorithm, see Fletcher (1987). This also is not a sufficiently accurate evaluation of $g(\theta)$. A second refinement is necessary. The second refinement is embedded within the MCMC chain $\{\theta_t\}_{t-1}^{H}$ of length H

that is used to compute the posterior distribution of θ. It is called the θ-chain. The θ-chain is generated using the Metropolis algorithm. The Metropolis algorithm is an iterative scheme that generates a Markov chain whose stationary distribution is the posterior of θ. To implement it, we require a likelihood, a prior, and transition density in θ called the proposal density. The likelihood is Equation (6) and the prior, $\xi(\theta)$, is described in Section 3.5.

The prior may require quantities computed from the simulation $\{\hat{y}_t, \hat{z}_{t-1}\}_{t-1}^{N}$ that are used in computing Equation (6). In particular, quantities computed in this fashion can be viewed as the evaluation of a functional of the structural model of the form $p(\cdot|\cdot,\theta) \mapsto \varrho$, where $\varrho \in \mathbf{P}$ and $\mathbf{P}$ is the space of functionals of the form $\theta \mapsto p(\cdot|\cdot,\theta) \mapsto \varrho$. Thus, the prior is a function of the form $\xi(\theta,\varrho)$. But since the functional ϱ is a composite function with $\theta \mapsto p(\cdot|\cdot,\theta) \mapsto \varrho$, $\xi(\theta,\varrho)$ is essentially a function of θ alone. Thus, we only use $\xi(\theta,\varrho)$ notation when attention to the subsidiary computation $p(\cdot|\cdot,\theta) \mapsto \varrho$ is required.

Let q denote the proposal density. For a given θ, $q(\theta,\theta^*)$ defines a distribution of potential new values θ^*. We use a move-one-at-a-time, random-walk, proposal density that puts its mass on discrete, separated points, proportional to a normal density. Two aspects of the proposal scheme are worth noting. The first is that the wider the separation between the points in the support of q the less accurately $g(\theta)$ needs to be computed for α at step 5 of the algorithm below to be correct. A practical constraint is that the separation cannot be much more than a standard deviation of the proposal density or the chain will eventually stick at some value of θ. Our separations are typically 1/2 of a standard deviation of the proposal density. In turn, the standard deviations of the proposal density are typically no more than the standard deviations of the prior distributions of structural parameters shown in Tables 3 to 7 and no less than one order of magnitude smaller. The second aspect worth noting is that the prior is putting mass on these discrete points in proportion to $\xi(\theta)$. Because one does not have to normalize either the likelihood or the prior in an MCMC chain, normalization of densities does not matter for the computation of the chain and similarly for the joint distribution $f(y|z,g(\theta))\xi(\theta)$ considered as a function of θ. However, $f(y|z,\omega)$ must be normalized such that $\int f(y|x,\omega)\,dy = 1$ to ensure that the implied map expressed in (7) is computed correctly.

The algorithm for the θ-chain is as follows. Given a current θ^o and the corresponding $\omega^o = g(\theta^o)$, we obtain the next pair (θ', ω') as follows:

1. Draw θ^* according to $q(\theta^o, \theta^*)$.

2. Draw $\{\hat{y}_t, \hat{z}_{t-1}\}_{t=1}^N$ according to $p(y_t|z_{t-1}, \theta^*)$.

3. Compute $\zeta^* = g(\theta^*)$ and the functional ϱ^* from the simulation $\{\hat{y}_t, \hat{z}_{t-1}\}_{t=1}^N$.

4. Compute $\alpha = \min\left(1, \frac{\mathcal{L}(\theta^*)\,\xi(\theta^*,\varrho^*)\,q(\theta^*,\theta^o)}{\mathcal{L}(\theta^o)\,\xi(\theta^o,\varrho^o)\,q(\theta^o,\theta^*)}\right)$.

5. With probability α, set $(\theta', \omega') = (\theta^*, \omega^*)$, otherwise set $(\theta', \omega') = (\theta^o, \omega^o)$.

It is at step 3 that AG made an important modification to the algorithm proposed by GM. At that point one has putative pairs (θ^*, ω^*) and (θ^o, ω^o) and corresponding simulations $\{\hat{y}_t^*, \hat{z}_{t-1}^*\}_{t=1}^N$ and $\{\hat{y}_t^o, \hat{z}_{t-1}^o\}_{t=1}^N$. AG use ω^* as a start and recompute ω^o using the BFGS algorithm, obtaining $\hat{\omega}^o$. If

$$\sum_{t=1}^{N} \log f(\hat{y}_t^o \mid \hat{z}_{t-1}^o, \hat{\omega}^o) > \sum_{t=1}^{N} \log f(\hat{y}_t^o \mid \hat{z}_{t-1}^o, \omega^o),$$

then $\hat{\omega}^o$ replaces ω^o. In the same fashion, ω^* is recomputed using ω^o as a start. Once computed, a (θ, ω) pair is never discarded. Neither are the corresponding $\mathcal{L}(\theta)$ and $\xi(\theta, \varrho)$. Because the support of the proposal density is discrete, points in the θ-chain will often recur, in which case $g(\theta)$, $\mathcal{L}(\theta)$, and $\xi(\theta, \varrho)$ are retrieved from storage rather than computed afresh. If the modification just described results in an improved (θ^o, ω^o), that pair and corresponding $\mathcal{L}(\theta^o)$ and $\xi(\theta^o, \varrho^o)$ replace the values in storage; similarly for (θ^*, ω^*). The upshot is that the values for $g(\theta)$ used at step 4 will be optima computed from many different random starts after the chain has run awhile.

3.3 GSM: Estimation of the auxiliary model

The observed data are y_t for $t = 1, \ldots, n$, where y_t is a vector of dimension M. The vector of observable variables used in estimation has four components: real equity returns, real interest rates, real per capita consumption growth rates, and real dividend growth rates. The symbols P, Q, V, etc. that appear in this section are general vectors (matrices) of statistical parameters and are not instances of the model parameters or functionals in Section 2.

The data are modeled as

$$y_t = \mu_{z_{t-1}} + U_{z_{t-1}} \varepsilon_t$$

where

$$\mu_{z_{t-1}} = b_0 + Bz_{t-1}, \tag{8}$$

which is the location function of a k-lag vector auto-regressive (VAR(k)) specification, obtained by letting columns of B past the first kM be zero. In this formulation, $U_{z_{t-1}}$ is the Cholesky factor of

$$\begin{aligned}
\Sigma_{z_{t-1}} &= U_0 U_0' && (9)\\
&\quad + Q\Sigma_{z_{t-2}} Q' && (10)\\
&\quad + P(y_{t-1} - \mu_{z_{t-2}})(y_{t-1} - \mu_{z_{t-2}})' P' && (11)\\
&\quad + \max[0, \tilde{V}(y_{t-1} - \mu_{z_{t-2}})] \max[0, \tilde{V}(y_{t-1} - \mu_{z_{t-2}})]', && (12)
\end{aligned}$$

where, as with B, the lag length is determined by letting the trailing columns of P and $\tilde{V}$ be zeros. In this application, the auxiliary model is not Markovian due to the recursion in expression (10).[10] As in Gallant and Tauchen (2014), the last term in the model above captures the leverage effect. In computations, $\max(0, x)$ in expression (12), which is applied element-wise, is replaced by a twice differentiable cubic spline approximation that plots slightly above $\max(0, x)$ over (0.00,0.10) and coincides elsewhere.

The density $h(\varepsilon)$ of the *i.i.d.* ε_t is the square of a Hermite polynomial times a normal density, the idea being that the class of such h is dense in Hellenger norm and can therefore approximate a density to within arbitrary accuracy in Kullback-Leibler distance, see Gallant and Nychka (1987). Such approximations are often called *sieves*; Gallant and Nychka term this particular sieve semi-nonparametric maximum likelihood estimator, or SNP.[11] The density $h(\varepsilon)$ is the normal when the degree of the Hermite polynomial is zero. In addition, the constant term of the Hermite polynomial can be a linear function of z_{t-1}. This has the effect of adding a nonlinear term to the location function (8) and the variance function (9). It also causes the higher moments of $h(\varepsilon)$ to depend on z_{t-1} as well. The SNP auxiliary model is determined statistically by adding terms as indicated by the BIC protocol for selecting the terms that comprise a sieve, see Schwarz (1978).

In our specification, U_0 is an upper triangular matrix, P and $\tilde{V}$ are diagonal matrices, and Q

[10] See Gallant and Long (1997) for the properties of estimators of the form used in this section when the model is not Markovian.

[11] See Gallant and Tauchen (2014) for an introduction and implementation of the SNP estimation.

is scalar. The degree of the SNP $h(\varepsilon)$ density is four. We specify that the constant term of the SNP density does not depend on the past. The auxiliary model chosen for our analysis, based on the BIC, has 1 lag in the conditional mean component, 1 lag in each of ARCH and GARCH terms. Although the univariate analysis of stock price dynamics generally incorporates a leverage term, we find in our SNP estimation with four variables that this term is not necessary according to the BIC.

The auxiliary model in the SNP estimation has 51 parameters of which 50 are estimated and one determined by a normalization rule. The error distributions implied by the auxiliary model differ significantly from the distributions of innovation shocks assumed in those structural models in Section 2. We numerically solve the structural models assuming normally distributed innovation shocks to consumption and dividend growth rates. The error distributions of simulations from these models are markedly non-Gaussian. For example, in addition to GARCH effects, the four-dimensional error distribution implied by the AAMS model is skewed in all four components and has fat-tails for consumption growth, dividend growth and stock returns and thin tails for bond returns.

3.4 Relative model comparison

Relative model comparison is standard Bayesian inference. The posterior probabilities of the five structural models may be computed using the Newton and Raftery (1994) $\hat{p}^4$ method for computing the marginal likelihood from an MCMC chain when assigning equal prior probability to each model. An alternative is method f_5 of Gamerman and Lopes (2006), Section 7.2.1. The advantage of these methods is that knowledge of the normalizing constants of the likelihood $\mathcal{L}(\theta)$ and the prior $\xi(\theta)$ are not required. We do not know these normalizing constants due to the imposition of support conditions. It is important, however, that the auxiliary model be the same for all models. Otherwise the normalizing constant of $\mathcal{L}(\theta)$ would be required. One divides the marginal density for each model by the sum for all models to get the posterior probabilities for relative model assessment.

Unfortunately, these and similar methods require that the range of the likelihoods that occur in the MCMC be within the float limits of the computing equipment employed. This can be remedied by left truncating the MCMC draws, which can be interpreted as a modification to the prior.

However, not only is it hard to interpret a truncation prior of this sort, but also we found that the implied ordering of the models is sensitive to the truncation for both the $\hat{p}^4$ and f_5 methods. Therefore, in the results reported below we used the BIC for model selection.

3.5 The prior and its support

All structural models considered in this paper are richly parameterized. We represent the parameter vector by θ. Table 2 summarizes structural parameters of all asset pricing models in Section 2. The prior of any structural parameter vector is the combination of the product of independent normal density functions and support conditions. The product of independent normal density functions is given by

$$\xi(\theta) = \prod_{i=1}^{\tilde{n}} N\left[\theta_i | \left(\theta_i^*, \sigma_\theta^2\right)\right]$$

where $\tilde{n}$ denotes the number of parameters. The complete set of location and scale parameters for independent normal priors as well as support conditions are available in the Internet Appendix. We set the location parameter values such that the asset pricing models generate mean risk-free rate that is not too high and mean equity premium that is not close to zero. For all models' parameters, we set the scale parameter values to be sufficiently large and use wide support intervals. This allows a wide range of parameter values of any model to be explored in the estimation, which in turn, provides ample room for asset pricing models to contribute to the identification of estimated parameters. Due to the support conditions, the effective prior is not an independence prior. For some values of θ^* proposed in Step 1 of the θ-chain described in Section 3, a model solution at Step 2 may not exist. In such cases, α at Step 5 is set to zero.

The prior support of the subjective discount factor (β), the coefficient of risk aversion (γ), and the EIS (ψ) parameter are set to $0.9 < \beta < 0.995$, $0.1 < \gamma < 100$, and $0.1 < \psi < 10$, respectively. The subjective discount factor must be high enough to imply a reasonably low risk-free rate. The range $0.9 < \beta < 0.995$ is wide compared to the prior on this parameter in Schorfheide et al. (2017). The support interval for γ that we use is much wider than the reasonable range $1 < \gamma < 10$ suggested by Mehra and Prescott (1985). Different from calibration studies on long-run risks, we do not impose $\psi > 1$ but allow for possibilities of $\psi < 1$ and a preference for late resolution of uncertainty. For the ambiguity aversion parameter η, the support interval is $\gamma < \eta < 200$. Again, this interval is wide given calibrated studies such as Ju and Miao (2012), Jahan-Parvar and Liu (2014) and Altug et al. (2017). Because the agent is ambiguity averse when $\eta > \gamma$, we impose this condition in estimating models with smooth ambiguity utility. The location parameters for β, γ, ψ and η in the prior are set at values consistent with the extant calibration studies. The scale parameters for these preference parameters are set to large values to deliver loose priors.

For models EZMS, AAMS and AAMSSV, we use the parameter estimates and the associated standard errors reported in Cecchetti, Lam, and Mark (2000) to determine the location and scale parameter values for parameters $\mu_h, \mu_l, \sigma_c, p_{hh}$ and p_{ll} in the Markov-switching model of consumption growth. In the AAMSSV model with time-varying volatility, our parameter choices for location and scale of p^{σ}_{hh}, $p^{\sigma}_{ll}, \sigma_h$ and σ_l rely on estimates of Lettau et al. (2008) and Boguth and Kuehn (2013). The location values of the dividend volatility parameter σ_d and the leverage parameter λ are determined by the calibration of Ju and Miao (2012). Following Abel (1999), we impose $\lambda \geq 1$ in the estimation. Estimation results of Bansal et al. (2007), Aldrich and Gallant (2011), and Schorfheide et al. (2017) lead to values of λ in the $[1.5, 4.5]$ range. We choose $1 \leq \lambda \leq 6$ as the support interval.

For models AALRRSV and EZLRRSV, we use the calibrated parameter values in Bansal et al. (2012) and priors postulated in Schorfheide et al. (2017) to choose the location and scale parameter values, and support intervals as well. For example, the location of the unconditional mean of consumption growth, μ_c, is set at 0.02 with a small scale parameter value. The location of the persistence parameter of the long-run risk component, ρ_x, is set at 0.95 with a large scale parameter value of 0.2. The support interval for ρ_x is $-0.99 < \rho_x < 0.99$. Similarly, other model parameters also have loose priors and wide support intervals as in Schorfheide et al. (2017).

3.6 Estimation results

We summarize estimation results using the GSM method in Tables 3 to 7.[12] We plot the prior and posterior densities of the estimated structural parameters in Figures 3 to 7. These plots show considerable shifts in both location and scale between priors and posteriors, suggesting that the estimation procedure and data have a significant impact on the estimation results. The impact of

[12] For each asset pricing model, we run the standard MCMC chain with the likelihood put to 1 at every draw to obtain the prior distribution of model parameters presented in Tables 3 to 7 and Figures 3 to 7.

priors and support conditions is notable, but of second order of importance.

Estimation results show that the posterior estimates of β are tightly bounded in all models and generally imply low risk-free rates. There is an ongoing debate about the value of the EIS parameter (ψ) in the macro-finance literature. This parameter is crucial for equilibrium asset pricing models to match macroeconomic and financial moments in the data, see Bansal and Yaron (2004), Croce (2014) and Liu and Miao (2015) among others. Some studies (e.g., Hall, 1988 and Ludvigson, 1999) find that the EIS estimate is less than 1, based on aggregate consumption and asset returns data. Other studies find higher values using cohort- or household-level data (e.g., Attanasio and Weber, 1993 and Vissing-Jorgensen, 2002). Our estimation strongly suggests an EIS greater than 1 and thus a preference for early resolution of uncertainty. As shown in Tables 3 to 7, the posterior mean, median, 5 and 95 percentiles of ψ estimates are all above 1 in all models presented in Section 2. The plots of the posterior density for ψ in Figures 3 to 7 also reveal that the posterior dispersion of this parameter over the MCMC chain is small. Jeong et al. (2015) estimate the recursive multiple prior utility model using asset prices data and obtain estimates of ψ greater than 10. High estimates of ψ generated from our estimation imply low and stable risk-free rates (see Section 4). In a DSGE analysis with broader scope, Bianchi et al. (2016) rely on the mechanism of time-varying ambiguity on operating costs to ease the tension between excess equity volatility and smooth risk-free rates.

The posterior estimates of ψ for models AAMS and EZMS are high and comparable to the estimates in the long-run risk literature. The posterior mean, median and 5 and 95 percentiles of ψ estimates are moderately higher in the EZMS model than in the AAMS model, with the posterior mean and median being above 2. The ψ estimates in the EZMS model are close to results obtained by Schorfheide et al. (2017) and Bansal, Kiku, and Yaron (2016). Our estimation results suggest that incorporating ambiguity in the model leads to lower estimates of ψ. This is also evident from a comparison of estimates in the EZLRRSV model and in the AALRRSV model. The posterior estimates of ψ are significantly lower in a long-run risk model with ambiguity than in a pure long-run risk model. Nevertheless, our estimates of ψ in the long-run risk model are still lower than those reported by Schorfheide et al. (2017). The discrepancy arises because 1) we use the projection method rather than log-linear approximation to solve models, 2) we use the GSM Bayesian method for model estimation, and 3) we use a different sample of data.

Our estimation results strongly support asset pricing models with smooth ambiguity. The posterior estimates of the ambiguity aversion parameter η are significantly large in models AAMS, AAMSSV and AALRRSV. Not surprisingly, the estimates obtained for the AAMS model are close to the calibrated value in Ju and Miao (2012) ($\eta = 8.864$). In addition, the estimates of η are modestly higher when regime-switching volatility in consumption growth is incorporated in the estimation. We observe that the posterior mean and median of η are about 10 in the AAMSSV model while about 7 in the AAMS model. In the long-run risk setting, the GSM Bayesian estimation generates high posterior estimates of η with mean and median of about 23. These results suggest that empirical support for models with smooth ambiguity is robust to different specifications of consumption dynamics and that the extent of ambiguity aversion largely depend on other preference parameters and primitive parameters in the consumption and dividend growth processes. While the estimated degree of ambiguity aversion varies across several models, these estimates are all reasonable from the perspective of decision-making. One could conduct thought experiments as in Halevy (2007) and Ju and Miao (2012) to gauge reasonable values of the ambiguity aversion parameter.

Estimates of the coefficients of risk aversion γ importantly hinge on the presence of ambiguity aversion. Estimation results of models EZMS and EZLRRSV show that the posterior mean and median of γ are high and the 5 and 95 percentiles imply tight bounds for the estimate. In particular, the posterior estimates of γ in the estimated long-run risk model EZLRRSV are close to the results reported by Schorfheide et al. (2017) and Bansal et al. (2016). The posterior mean of γ is 8.4, and the associated 95 percentile is 10.4. These values are also close the the calibrated values in Bansal and Yaron (2004) and Bansal et al. (2012). On the other hand, the γ estimate is more dispersed in models with smooth ambiguity, i.e., models AAMS, AAMSSV and AALRRSV, as is evident from wide (5%,95%) intervals. In a related work, Chen, Favilukis, and Ludvigson (2013) estimate preference parameters of recursive utility using a semiparametric technique. Their estimated relative risk aversion parameter ranges from 17 to 60.

In the GSM Bayesian estimation, primitive parameters in the consumption and dividend growth processes are jointly estimated with preference parameters. Models AAMS, AAMSSV and EZMS have Markov-switching consumption growth while models AALRRSV and EZLRRSV feature long-run risks. In the Markov-switching environment, our estimation method identifies a normal regime and a contraction regime for mean consumption growth. The posterior estimates of μ_h are largely in line with the historical average annual consumption growth. For instance, the posterior mean and median of μ_h in the AAMS model are about 2%. In addition, the posterior estimates of the transition probability p_{hh} (p^{μ}_{hh} in model AAMSSV) are close to 1 and thus indicate that this regime is very persistent. Furthermore, the estimates of low mean growth regime for these models indicate a relatively transitory contraction regime with lower estimates of the transition probability p_{ll} (p^{μ}_{ll} in model AAMSSV).

Note that we obtain these estimates from structural estimation of asset pricing models using data on both fundamentals and asset returns. The GSM Bayesian estimation takes into account equilibrium asset prices and yields estimated consumption dynamics that corresponds to the agent's subjective belief. Compared with estimates of the parameters of the Markov-switching model reported by calibration studies (e.g., Cecchetti et al. (2000) and Ju and Miao (2012)), our estimates imply a "peso" version of the model. That is, the severe contraction state rarely realizes in the observed data or simulations due to its low likelihood $(1 - p_{hh})$. However, because an agent cannot observe the mean growth state and is also aware of severity (μ_l) and persistence (p_{ll}) of the contraction regime, the agent is always concerned about state uncertainty and moreover, ambiguity aversion magnifies the impact of this concern. In addition, the posterior estimates of the low mean regime μ_l seem too low given the post-war experience of the economy, and the estimated persistence of this regime varies significantly across different models. These results suggest that apart from ambiguity on the mean growth state, extra sources of ambiguity about parameters of the Markov-switching model may co-exist.

In estimating the AAMSSV model, we find two distinct volatility regimes, both of which are persistent. This result is consistent with the findings of Lettau et al. (2008) and Boguth and Kuehn (2013). However, the posterior estimates of the high volatility regime σ_h are too high to be reconciled with the post-war consumption data. The estimates of μ_l are even more negative than the estimates for the AAMS model. Nevertheless, these estimates are more consistent with the long sample of Shiller's data.[13] Again, extra sources of ambiguity may arise due to learning

[13] We thank Robert Shiller for making the data available at http://www.econ.yale.edu/~shiller/data/chapt26.xlsx.

from past experiences or parameter uncertainty.[14] For models AAMS and AAMSSV, the leverage parameter λ and the dividend growth volatility σ_d estimates are reasonably close to the calibrated values considered by Abel (1999), Bansal and Yaron (2004) and Ju and Miao (2012). The posterior estimates of λ are roughly between 2 and 4 with a posterior mean of about 3 for both models. The estimates of λ and σ_d for the EZMS model are significantly higher than those for models AAMS and AAMSSV.

Turning to estimation results of models featuring long-run risks, we find that the estimated models AALRRSV and EZLRRSV both provide support to the presence of a persistent component in the consumption growth process. This empirical support is evident even when ambiguity about conditional mean growth is incorporated in the model. The posterior estimates of the persistence parameter ρ_x are close to 1 with narrow (5%, 95%) intervals. Converted into estimates at a monthly frequency, our results are similar to those reported by Schorfheide et al. (2017). In addition, the stochastic volatility component is also persistent in our estimation, a result consistent with Schorfheide et al. (2017).[15] Other parameter estimates including μ_c, μ_s, σ_w λ, ϕ_d and ϕ_c are similar to the estimates reported by the studies on long-run risks such as Bansal et al. (2012), Bansal et al. (2016) and Schorfheide et al. (2017).

We present results of relative model comparison in Tables 3 to 7, based on the maximum of the log likelihood and the BIC for all estimated models. We use the auxiliary model presented in Section 3.3 and the MCMC chain of structural parameters of each asset pricing model to compute the maximum of the log likelihood and the BIC of the model. According to these two criteria, among all five estimated models the AAMSSV best characterizes the data in that the model provides the best fit of the SNP density given the observed data. The log likelihood computation leads to the model ranking AAMSSV$\succ$AALRRSV$\succ$EZLRRSV$\succ$EZMS$\succ$AAMS. The BIC gives us the same ranking except that EZMS$\succ$EZLRRSV because the number of model parameters is also taken into account. Based on the BIC ranking the AALRRSV model is close to AAMSSV, but the remainder are more than 40 orders of magnitude distant. These findings suggest that 1) time-

[14] A full-fledged analysis of modeling multiple sources of ambiguity requires development of new models that have parameter uncertainty, state uncertainty and learning. Estimating this class of models is beyond the scope of our current study.

[15] Applying the GSM Bayesian estimation, we find that the parameter value of ρ_s in the MCMC chain remains stagnant at a high level ($\rho_s = 0.95$).

varying volatility in consumption is important for asset pricing models to deliver the SNP densities that fit the data well, because according to the log likelihood criterion priority is given to models AAMSSV, AALRRSV and EZLRRSV, all of which feature time-varying volatility, and 2) asset pricing models (AAMSSV and AALRRSV) with ambiguity, learning and time-varying volatility are preferred to the long-run risk model EZLRRSV in the statistical model comparison. Although the model of Ju and Miao (2012), AAMS, receives less statistical support than other models do, it can match key financial moments well, as shown in the next section.

4 Asset Pricing Implications

4.1 Variance risk premium

The moments of equity returns are naturally defined under the physical measure implied by fundamentals and the state variables in any asset pricing model. Furthermore, we can study the dynamics of the risk-neutral variance and variance risk premium (henceforth, VRP) generated from models considered above. As noted in Bollerslev, Tauchen, and Zhou (2009), the market variance risk premium is defined as the difference between the expected equity return variances under the risk-neutral and physical measures, and it measures the risk premium compensation for investors bearing the variance risk. Several studies show that the mean and volatility of the market variance risk premium are high, which poses a serious challenge to many existing asset pricing models, for example, see the discussion in Drechsler (2013). In a calibration study, Miao, Wei, and Zhou (2012) find that the AAMS model can roughly match the mean and volatility of the VRP in the data. Here, we take a different stance in that we do not calibrate any model to target moments of the VRP. Instead, we examine whether our estimated models produce empirically reasonable dynamics of the VRP.

In the literature, a commonly used empirical proxy for the risk-neutral volatility is the Chicago Board Options Exchange (CBOE)'s volatility index (VIX). In the empirical analysis, we measure the market variance risk premium as the difference between the model-free implied variance and the conditional projection of realized variance. Our empirical estimation of the VRP closely follows the study of Liu and Zhang (2015), which applies the CBOE's methodology of constructing the

VIX to index options with 90 days maturity. To estimate the variance of equity returns under the physical measure, we first compute realized returns and then take a linear projection to obtain the conditional variance, which denoted by VOL_t^2. The variance risk premium is defined as

$$VRP_t = VIX_t^2 - VOL_t^2.$$

In the model, the risk-neutral variance VIX_t^2 takes the form

$$VIX_t^2 = \mathbb{E}_t^{\mathcal{Q}}\left[\sigma_{r,t+1}^2\right] = \frac{\mathbb{E}_t\left[M_{t,t+1}\sigma_{r,t+1}^2\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]}$$

where $\mathcal{Q}$ denotes the risk-neutral measure, and the expected variance under the physical measure is given by

$$VOL_t^2 = \mathbb{E}_t\left[\sigma_{r,t+1}^2\right]$$

where $\sigma_{r,t}^2 = \mathbb{E}_t\left[r_{t+1}^2\right] - \left(\mathbb{E}_t\left[r_{t+1}\right]\right)^2$.

4.2 Impulse responses

We perform impulse responses analyses for the estimated asset pricing models by investigating key financial variables including the SDF, price-dividend ratio, conditional equity premium, equity volatility and variance risk premium. We use mean estimates reported in Tables 3 to 7 to parameterize models and compute impulse responses functions. Results for models AAMS, AAMSSV, AALRRSV and EZLRRSV are plotted in Figures 8 and 9. We assume that the shock to mean growth rate of consumption occurs in the third period and lasts only one period.

Figure 8 shows that when the mean consumption growth regime shifts from “high” (μ_h) to “low” (μ_l), Bayesian updating leads to a lower level of belief π_t. Veronesi (1999) has shown that with CRRA utility, the impact will be an increase in conditional equity volatility and equity premium. This effect is amplified under ambiguity aversion. The plotted ambiguity-distorted belief manifests endogenous pessimism that implies a sharp increase in the SDF and a decrease in the price-dividend ratio. As a result, the conditional equity volatility and equity premium rise significantly. Since conditional volatility rises in states where the SDF is high, the risk-neutral variance increases more than the physical return variance does, leading to an increase in the VRP. Figure 8 displays qualitatively similar impulse responses of beliefs and financial variables for the AAMSSV model where the consumption volatility state is assumed to be σ_h throughout the response periods.[16] The notable discrepancies in the magnitude of responses between the AAMS model and the AAMSSV model are largely due to the inclusion of time-varying volatility in the AAMSSV model and different parameter estimates as discussed in Section 3.6.

Figure 9 displays the responses of key variables in models AALRRSV and EZLRRSV when a negative shock of size $-4\varphi_x\mu_s$ hits the long-run risk component x_t, which is assumed to be zero initially. Different from the AAMS model with Markov-switching growth rates, in the AALRRSV model Bayesian filtering of x_t implies persistent movements in financial variables because of its long-run risk feature. Again, the plotted ambiguity-distorted belief reflects the agent's pessimistic view about the conditional mean growth rate of consumption. In line with the long-run risk model, learning about x_t produces a SDF and a price-dividend ratio that move in the opposite directions upon the impact of the shock. Thus, in the AALRRSV model the long-run risk component carries a positive risk premium. Because the conditional volatility of consumption growth is assumed to be constant in this analysis, the conditional equity volatility decreases on impact and rises slowly afterwards. The conditional equity premium exhibits a similar response as a consequence. The VRP falls at first and rises afterwards, due to the response of the conditional equity volatility. Figure 9 shows similar impulse responses for the EZLRRSV model in which the long-run risk component is fully observable. In both models, the response of the VRP is negligible compared to the results for models AAMS and AAMSSV.

4.3 Financial moments

We investigate the ability of all estimated models in replicating unconditional moments of key macroeconomic and financial variables. Unlike calibration studies, our aim is not to match unconditional moments of asset returns in the data as closely as possible. Instead, we assess the impact of ambiguity aversion on financial variables based on estimated parameter values. In addition, we examine how well our estimated models can match moments of asset returns, given that our estima-

[16] The impulse responses plot for the EZMS model is similar and thus omitted here for the sake of brevity.

tion strategy is designed not to match moments but to fit the SNP densities of asset pricing models given the observed data. If any estimated model is reasonably successful in reproducing unconditional moments of consumption growth and asset returns, we view this outcome as confirmation that the model characterizes the underlying data generating process well. This analysis makes our structural estimation more relevant from an alternative empirical perspective. By examining asset pricing implications of estimated models, our analysis supersedes previous studies on structural estimations such as Bansal et al. (2007), Aldrich and Gallant (2011) and Jeong et al. (2015).

Table 8 presents unconditional moments of asset returns simulated from all asset pricing models considered in this paper. For each model, we compute these moments on a MCMC chain of 12,000 estimates and report mean, median, standard deviation, 5th and 95th percentiles of simulation results. To facilitate comparison, we present moments computed from the historical U.S. data. Due to the high EIS estimates in all models and resulting intertemporal substitution effect, the mean and volatility of the risk-free rate are low across these models. All models produce simulations on their chains of estimates that contain the historical equity premium and return volatility in the $(5\%, 95\%)$ intervals.

Table 8 shows that among all models, the AAMS model can best match moments of returns. The estimated AAMS model delivers mean and volatility of the risk-free rate, equity premium and return volatility, and mean and volatility of the VRP close to the moments computed from the data. In addition, the 5th and 95th percentiles of simulated moments are sufficiently tight to include the data moments except for the volatilities of the risk-free rate and VRP. The intuition of the impact of ambiguity on asset returns has been illustrated in previous studies, for example see Ju and Miao (2012) and Collard et al. (2017). That is, the precautionary savings motive driven by ambiguity aversion reduces the risk-free rate, and in addition to the standard risk premium the agent also demands an uncertainty premium for being ambiguous about the data-generating process. The latter mechanism is evident from inspecting the market price of risk, which is defined as $\sigma(M_{t,t+1})/\mathbb{E}(M_{t,t+1})$. According to the conditional version of the Euler equation:

$$\mathbb{E}_t\left(R_{t+1}\right) - R_{f,t} = -\frac{\sigma_t\left(M_{t,t+1}\right)}{\mathbb{E}_t\left(M_{t,t+1}\right)}\sigma_t\left(R_{t+1}\right)\rho_t\left(M_{t,t+1}, R_{t+1}\right),$$

the high market price of risk implied by the AAMS model leads to a high equity premium. Since the estimated model also produces volatility of dividend growth close to the data and the leverage parameter consistent with previous calibration studies, the model can naturally match the volatility of equity returns in the data.

The AAMS model also generates a high VRP close to the data. This is a remarkable result, since we do not use the risk-neutral variance data to aid estimation. The implied high VRP is a consequence of strong co-movement of the SDF and the return volatility when the economy shifts to a bad state. The co-movement therefore leads to a substantial wedge between the risk-neutral variance and the objective variance. On the other hand, the estimated EZMS model (ambiguity neutral case) shows poor performance in matching the moments. The mean of simulated equity premium in this model is only half of the historical equity premium whereas the moments of the VRP are much higher than the data.

It is evident in Table 8 that incorporating time-varying consumption volatility in the Markov-switching model does not yield significantly better asset pricing results, though the GSM Bayesian estimation provides statistical support to this model relative to the more parsimonious model AAMS. The model predicts mean values of equity premium and VRP moderately higher than the data. The range of the 5th−95th percentile is wider than that in the AAMS model both for the simulated equity premium and the VRP. The mean of the market price of risk increases greatly with the addition of regime-switching conditional volatility.

In the long-run risk setting, the equity premium and market price of risk implied by the AALRRSV model is higher than those in the EZLRRSV model due to the significant impact of ambiguity. However, neither model is able to match moments of the VRP in the data. Both models generate mean and volatility of the VRP close to zero. This is in contrast to models AAMS and AAMSSV that can match both equity premium and mean VRP well. In fact, one must introduce jumps in state processes to generate a high and volatile VRP in the long-run risk setting, for example see Drechsler (2013). We leave structural estimation of this class of models for future research.

As the AAMS model can best match unconditional moments of financial variables, we next study conditional financial moments generated by this model. Because the AAMSSV model is an extension of AAMS and a statistically preferred model as suggested by the model comparison, we also investigate conditional financial moments in AAMSSV. Figure 10 shows simulated conditional equity premium, return volatility, market price of risk and the VRP plotted against the state variable π_t in model AAMS. The conditional moments are drawn from the 5th to 95th percentile of the simulations implied by 12,000 MCMC estimates of structural parameters of the model. We also show conditional moments generated from Ju and Miao (2012)'s calibration for comparison. We observe that the simulated 90% region of conditional moments does not include the calibration results of Ju and Miao (2012). This is because Ju and Miao (2012) use a long sample for their calibration. Figure 11 plots simulated conditional moments for model AAMSSV, where in each simulation the expectation with respect to volatility states is computed using stationary probabilities of the two volatility regimes. For both models, we observe that the key conditional financial moments exhibit a hump-shape when plotted against π_t, and that conditional equity premium, market price of risk and VRP peak close to high values of π_t. This is due to that our estimation implies a very persistent normal regime of consumption growth with p_{hh} close to 1. Suppose that the economy initially stays in the normal regime. A negative shock to consumption prompts the agent to update his belief π_t downward, leading to enhanced state uncertainty. Ambiguity aversion further exacerbates the scenario by inducing endogenous pessimism and thus implies a significant increase in conditional equity premium, market price of risk and VRP.

5 Conclusion

We have estimated a series of consumption-based asset pricing models with and without smooth ambiguity preferences. We use the GSM Bayesian estimation method developed by Gallant and McCulloch (2009) and an encompassing and flexible auxiliary model to jointly estimate preference parameters and dynamic models of consumption and dividend growth postulated in asset pricing models. We employ the semi-nonparametric method to estimate the auxiliary model and the GSM Bayesian method to obtain posterior estimates of structural parameters of asset pricing models. Our structural estimation with macro-finance data provides statistical support for asset pricing models with smooth ambiguity. Based on our estimation results, the quantitative effects of smooth ambiguity on asset returns are significant, both in the Markov-switching and long-run risk environments.

Our main findings are: (1) the distinction between risk aversion and ambiguity aversion is robust to the intertemporal substitution effect, i.e., our estimation provides statistical support to an ambiguity-averse representative agent who also prefers early resolution of uncertainty (or has a high EIS), (2) the statistical support for smooth ambiguity is robust to specifications of consumption and dividend processes, (3) a model comparison shows that models with ambiguity, learning and time-varying volatility are preferred to the long-run risk model, and (4) in the Markov-switching environment our estimation identifies a normal regime and a contraction regime for the mean growth rate of consumption as well as two distinct volatility regimes; in the long-run risk environment our estimation identifies the long-run risk component.

References

Abel, A. B., 1999. Risk premia and term premia in general equilibrium. Journal of Monetary Economics 43 (1), 3–33.

Aldrich, E. M., Gallant, A. R., 2011. Habit, long-run risks, prospect? a statistical inquiry. Journal of Financial Econometrics 9, 589–618.

Altug, S., Collard, F., Çakmakli, C., Mukerji, S., Özsöylev, H., 2017. Ambiguous business cycles: A quantitative assessment. Working Paper.

Attanasio, O. P., Weber, G., 1993. Consumption growth, the interest rate and aggregation. Review of Economic Studies 60, 631–649.

Backus, D., Ferriere, A., Zin, S., 2015. Risk and ambiguity in models of business cycles. Journal of Monetary Economics 69, 42–63.

Bansal, R., Gallant, R., Tauchen, G., 2007. Rational pessimism and exuberance. Review of Economic Studies 74, 1005–1033.

Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2014. Volatility, the macroeconomy, and asset prices. Journal of Finance 69, 2471–2511.

Bansal, R., Kiku, D., Yaron, A., 2012. An empirical evaluation of the long-run risks model for asset prices. Critical Finance Review 1, 183–221.

Bansal, R., Kiku, D., Yaron, A., 2016. Risks for the long run: Estimation with time aggregation. Journal of Monetary Economics 82, 52–69.

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

Bianchi, F., Ilut, C., Schneider, M., 2016. Uncertainty shocks, asset supply and pricing over the business cycle. Review of Economic Studies forthcoming.

Bidder, R., Dew-Becker, I., 2016. Long-run risk is the worst-case scenario. American Economic Review 106 (9), 2494–2527.

Boguth, O., Kuehn, L.-A., 2013. Consumption volatility risk. Journal of Finance 68, 2589–2615.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22 (11), 4463–4492.

Bryzgalova, S., Julliard, C., 2015. The consumption risk of bonds and stocks. Working Paper.

Cecchetti, S. G., Lam, P.-S., Mark, N. C., 2000. Asset pricing with distorted beliefs: Are equity returns too good to be true? American Economic Review 90, 787–805.

Chen, X., Favilukis, J., Ludvigson, S., 2013. An estimation of economic models with recursive preferences. Quantitative Economics 4, 39–83.

Chen, Z., Epstein, L., 2002. Ambiguity, risk, and asset returns in continuous time. Econometrica 70 (4), 1403–1443.

Cogley, T., Sargent, T. J., 2008. The market price of risk and the equity premium: A legacy of the great depression? Journal of Monetary Economics 55, 454–476.

Collard, F., Mukerji, S., Sheppard, K., Tallon, J.-M., 2017. Ambiguity and the historical equity premium. Working Paper.

Collin-Dufresne, P., Johannes, M., Lochstoer, L. A., 2016. Parameter learning in general equilibrium: The asset pricing implications. American Economic Review 106, 664–698.

Croce, M. M., 2014. Long-run productivity risk: A new hope for production-based asset pricing? Journal of Monetary Economics 66, 13–31.

Drechsler, I., 2013. Uncertainty, time-varying fear, and asset prices. Journal of Finance 68, 1843–1889.

Engle, R. F., Kroner, K. F., 1995. Multivariate Simulteneous Generalized ARCH. Econometric Theory 11 (1), 122–150.

Epstein, L., Schneider, M., 2008. Ambiguity, information quality and asset pricing. Journal of Finance 63, 197–228.

Epstein, L. G., Schneider, M., 2007. Learning under ambiguity. Review of Economic Studies 74 (4), 1275–1303.

Epstein, L. G., Wang, T., 1994. Intertemporal asset pricing under Knightian uncertainty. Econometrica 62 (3), 283–322.

Epstein, L. G., Zin, S. E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. Econometrica 57 (4), 937–969.

Fama, E. F., French, K. R., 1988a. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–25.

Fama, E. F., French, K. R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–49.

Fletcher, R., 1987. Practical Methods of Optimization, 2nd Edition. John Wiley and Sons, New York, NY.

Gallant, A. R., Long, J. R., 1997. Estimating stochastic differential equations efficiently by minimum chi-squared. Biometrica 84, 125–141.

Gallant, A. R., McCulloch, R. E., 2009. On the determination of general scientific models with application to asset pricing. Journal of the American Statistical Association 104, 117–131.

Gallant, A. R., Nychka, D. W., 1987. Seminonparametric maximum likelihood estimation. Econometrica 55, 363–390.

Gallant, A. R., Tauchen, G., 1996. Which moments to match? Econometric Theory 12, 657–681.

Gallant, A. R., Tauchen, G., 1998. Reprojecting partially observed systems with application to interest rate diffusions. Journal of the American Statistical Association 93, 10–24.

Gallant, A. R., Tauchen, G., 2010. Handbook of Financial Econometrics. Elsevier/North-Holland, Amsterdam, Ch. Simulated Score Methods and Indirect Inference for Continuous-time Models, pp. 427–478.

Gallant, A. R., Tauchen, G., April 2014. SNP: A Program for Nonparametric Time Series Analysis - User's Guide. Penn State University and Duke University, version 9.1 Edition.

Gamerman, D., Lopes, H. F., 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd Edition. No. 86 in CRC Texts in Statistical Science. Chapman & Hall, New York, NY.

Gouriéroux, C., Monfort, A., Renault, E. M., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.

Halevy, Y., 2007. Ellsberg revisited: An experimental study. Econometrica 75, 503–536.

Hall, R., 1988. Intertemporal substitution in consumption. Journal of Political Economy 96, 339–357.

Hansen, L. P., 2007. Beliefs, doubts and learning: The valuation of macroeconomic risk. American Economic Review 97, 1–30.

Hansen, L. P., Sargent, T. J., 2010. Fragile beliefs and the price of uncertainty. Quantitative Economics 1, 129–162.

Hayashi, T., Miao, J., 2011. Intertemporal substitution and recursive smooth ambiguity preferences. Theoretical Economics 6.

Ilut, C., Schneider, M., 2014. Ambiguous business cycles. American Economic Review 104, 2368–2399.

Jahan-Parvar, M. R., Liu, H., 2014. Ambiguity aversion and asset prices in production economies. Review of Financial Studies 27 (10), 3060–3097.

Jarque, C. M., Bera, A. K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6 (3), 255–259.

Jeong, D., Kim, H., Park, J. Y., 2015. Does ambiguity matter? Estimating asset pricing models with a multiple-priors recursive utility. Journal of Financial Economics 115 (2), 361–382.

Johannes, M., Lochstoer, L. A., Mou, Y., 2016. Learning about consumption dynamics. Journal of Finance 71, 551–600.

Ju, N., Miao, J., 2012. Ambiguity, learning, and asset returns. Econometrica 80, 559–591.

Klibanoff, P., Marinacci, M., Mukerji, S., 2005. A smooth model of decision making under ambiguity. Econometrica 73 (6), 1849–1892.

Klibanoff, P., Marinacci, M., Mukerji, S., 2009. Recursive smooth ambiguity preferences. Journal of Economic Theory 144 (3), 930–976.

Leippold, M., Trojani, F., Vanini, P., 2008. Learning and asset prices under ambiguous information. Review of Financial Studies 21 (6), 2565–2597.

Lettau, M., Ludvigson, S. C., Wachter, J. A., 2008. The declining equity premium: What role does macroeconomic risk play? Review of Financial Studies 21, 1653–1687.

Liu, H., Miao, J., 2015. Growth uncertainty, generalized disappointment aversion and production-based asset pricing. Journal of Monetary Economics 69, 70–89.

Liu, H., Zhang, Y., 2015. Ambiguity and financial uncertainty in a real business cycle model. Working Paper.

Ludvigson, S. C., 1999. Consumption and credit: A model of time-varying liquidity constraints. Review of Economics and Statistics 81, 434–447.

Ma, J., 2013. Long-run risk and its implications for the equity premium puzzle: New evidence from a multivariate framework. Journal of Money, Credit and Banking 45 (1), 121–145.

McConnell, M. M., Perez-Quiros, G., 2000. Output fluctuations in the united states: What has changed since the early 1980s? American Economic Review 90, 1464–1476.

Mehra, R., Prescott, E. C., 1985. The equity premium: A puzzle. Journal of Monetary Economics 15, 145–161.

Miao, J., Wei, B., Zhou, H., 2012. Ambiguity aversion and variance premium. Working Paper.

Newton, M. A., Raftery, A. E., 1994. Approximate bayesian inference with the weighted likelihood bootstrap. Journal of the Royal Statistical Society - Series B 56, 3–48.

Pohl, W., Schmedders, K., Wilms, O., 2017. Higher-order effects in asset pricing models with long-run risks. Journal of Finance forthcoming.

Schorfheide, F., Song, D., Yaron, A., 2017. Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach. Econometrica forthcoming.

Schwarz, G., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461–464.

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

Shiller, R. J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421–36.

Veronesi, P., 1999. Stock market overreaction to bad news in good times: A rational expectations equilibrium model. Review of Financial Studies 12, 975–1007.

Vissing-Jorgensen, A., 2002. Limited asset market participation and the elasticity of intertemporal substitution. Journal of Political Economy 110, 825–853.

Table 1: Summary Statistics of the Data

	r_t^e	r_t^f	$r_t^e - r_t^f$	Δc_t	Δd_t
Mean	5.98	0.96	5.03	1.83	1.56
St. Dev.	19.70	2.47	19.96	2.14	14.08
Skewness	-0.8193	-1.4763	-0.6988	0.1079	-0.8716
Kurtosis	0.5926	5.0291	0.4457	0.0360	2.8810
J-B Test	0.0135	0.0010	0.0263	0.5000	0.0010

This table reports summary statistics for annual U.S. data (1941–2015). Mean and standard deviations of aggregate equity returns (r_t), one-year Treasury Bill rate (r_t^f), excess returns ($r_t - r_t^f$), real per capita log consumption growth (Δc_t), and real log dividend growth (Δd_t) are expressed in percentages. "J-B test" reports the p-values of Jarque and Bera (1980) test of normality, where the null hypothesis is that the time series is normally distributed.

Table 2: **Model Summary**

Model	State variables	Parameters
AAMS	π_t	$\{\beta, \gamma, \psi, \eta, \mu_h, \mu_l, p_{hh}, p_{ll}, \sigma, \lambda, \sigma_d\}$
AAMSSV	$(\pi_t,\ s_t^\sigma)$	$\{\beta, \gamma, \psi, \eta, \mu_h, \mu_l, p_{hh}^\mu, p_{ll}^\mu, \sigma_h, \sigma_l, p_{hh}^\sigma, p_{ll}^\sigma, \lambda, \sigma_d\}$
AALRRSV	$(\hat{x}_t,\ \nu_t,\ \sigma_t)$	$\{\beta, \gamma, \psi, \eta, \mu_c, \rho_x, \varphi_x, \lambda, \varphi_d, \mu_\sigma, \rho_\sigma, \sigma_w\}$
EZMS	π_t	$\{\beta, \gamma, \psi, \mu_h, \mu_l, p_{hh}, p_{ll}, \sigma, \lambda, \sigma_d\}$
EZLRRSV	(x_t,σ_t^2)	$\{\beta, \gamma, \psi, \mu_c, \rho_x, \varphi_x, \lambda, \varphi_d, \varphi_c, \mu_\sigma, \rho_\sigma, \sigma_w\}$

This table summarizes relevant state variables and structural parameters for each asset pricing model described in Section 2.

Table 3: **GSM Estimation Results: the AAMS Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.985	0.985	0.978	0.993	0.975	0.974	0.969	0.985
γ	4.908	4.750	3.250	6.750	2.841	3.063	0.563	4.563
ψ	1.512	1.563	1.188	1.813	2.040	2.031	1.781	2.406
η	9.109	9.500	6.500	12.500	6.959	6.938	5.063	8.938
p_{ll}	0.543	0.531	0.344	0.781	0.835	0.839	0.786	0.888
p_{hh}	0.783	0.813	0.563	0.938	0.996	0.997	0.994	0.997
μ_l	-0.059	-0.059	-0.074	-0.035	-0.039	-0.039	-0.048	-0.031
μ_h	0.022	0.021	0.014	0.033	0.022	0.022	0.016	0.029
λ	2.598	2.750	1.250	3.750	3.420	3.422	2.703	4.203
σ_c	0.028	0.029	0.018	0.041	0.019	0.019	0.015	0.022
σ_d	0.137	0.133	0.086	0.180	0.137	0.137	0.113	0.168
BIC						832.14		
Log likelihood						-392.32		
MCMC repetitions		10,000				12,000		

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AAMS model. "BIC" represents the Bayesian information criteria, see Schwarz (1978). "Log likelihood" represents the maximum of the log likelihood of the encompassing model over the MCMC chain of estimates. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for the U.S. annual data 1941–2015.

Table 4: **GSM Estimation Results: the AAMSSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.984	0.983	0.978	0.991	0.982	0.984	0.972	0.991
γ	4.723	4.750	3.250	6.250	1.167	0.875	0.125	4.125
ψ	1.483	1.438	1.188	1.813	1.357	1.348	1.090	1.668
η	9.235	9.500	6.500	12.500	10.252	10.125	6.875	13.625
p_{ll}^{μ}	0.508	0.531	0.281	0.719	0.668	0.686	0.504	0.746
p_{hh}^{μ}	0.806	0.813	0.563	0.938	0.996	0.998	0.984	0.999
μ_l	-0.059	-0.059	-0.074	-0.027	-0.056	-0.057	-0.068	-0.042
μ_h	0.022	0.021	0.014	0.029	0.023	0.023	0.014	0.033
p_{ll}^{σ}	0.849	0.859	0.734	0.953	0.986	0.990	0.948	0.996
p_{hh}^{σ}	0.841	0.859	0.734	0.953	0.982	0.984	0.957	0.995
σ_l	0.015	0.015	0.009	0.021	0.013	0.012	0.004	0.022
σ_h	0.030	0.029	0.018	0.041	0.038	0.038	0.029	0.050
λ	2.881	2.750	1.750	3.750	2.739	2.641	1.953	4.016
σ_d	0.131	0.133	0.086	0.180	0.159	0.157	0.122	0.210
BIC					746.31			
Log likelihood					-342.93			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AAMSSV model. "BIC" represents the Bayesian information criteria, see Schwarz (1978). "Log likelihood" represents the maximum of the log likelihood of the encompassing model over the MCMC chain of estimates. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for the U.S. annual data 1941–2015.

Table 5: **GSM Estimation Results: the AALRRSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.982	0.981	0.976	0.989	0.986	0.987	0.979	0.992
γ	4.791	4.875	2.625	6.625	4.683	5.031	1.719	6.406
ψ	1.482	1.469	1.156	1.781	1.225	1.113	1.012	1.785
η	24.926	25.000	17.000	33.000	23.371	23.500	10.500	35.500
μ_c	0.019	0.019	0.017	0.021	0.019	0.019	0.017	0.020
ρ_x	0.772	0.781	0.531	0.969	0.941	0.941	0.926	0.957
ϕ_x	0.157	0.148	0.086	0.227	0.248	0.248	0.197	0.295
λ	2.905	2.875	1.625	4.375	3.555	3.453	2.953	4.672
ϕ_d	2.754	2.625	1.375	4.125	4.877	4.906	3.844	5.844
μ_s	0.020	0.021	0.011	0.028	0.020	0.020	0.019	0.021
ρ_s	0.969	0.969	0.969	0.969	0.950	0.950	0.950	0.950
σ_w	2.37E-04	2.29E-04	1.37E-04	3.51E-04	2.57E-04	2.54E-04	2.37E-04	2.79E-04
BIC					765.09			
Log likelihood					-356.64			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AALRRSV model. “BIC” represents the Bayesian information criteria, see Schwarz (1978). “Log likelihood” represents the maximum of the log likelihood of the encompassing model over the MCMC chain of estimates. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for the U.S. annual data 1941–2015.

Table 6: **GSM Estimation Results: the EZMS Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.985	0.985	0.978	0.991	0.976	0.976	0.970	0.986
γ	4.771	4.750	3.250	6.250	2.909	2.906	2.344	3.484
ψ	1.488	1.438	1.188	1.813	2.400	2.281	1.844	3.656
p_{ll}	0.530	0.531	0.281	0.781	0.972	0.972	0.943	0.989
p_{hh}	0.774	0.813	0.563	0.938	0.993	0.993	0.987	0.999
μ_l	-0.059	-0.059	-0.074	-0.035	-0.030	-0.029	-0.042	-0.017
μ_h	0.022	0.021	0.014	0.029	0.030	0.030	0.020	0.041
λ	2.647	2.750	1.750	3.750	4.974	5.109	3.391	5.859
σ_c	0.028	0.029	0.018	0.037	0.021	0.022	0.010	0.029
σ_d	0.134	0.133	0.086	0.180	0.181	0.184	0.137	0.223
BIC					812.98			
Log likelihood					-384.90			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the EZMS model. "BIC" represents the Bayesian information criteria, see Schwarz (1978). "Log likelihood" represents the maximum of the log likelihood of the encompassing model over the MCMC chain of estimates. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for the U.S. annual data 1941–2015.

Table 7: **GSM Estimation Results: the EZLRRSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.984	0.985	0.978	0.991	0.982	0.982	0.977	0.989
γ	4.978	4.750	3.250	6.750	8.431	8.531	6.219	10.438
ψ	1.448	1.438	1.063	1.813	1.732	1.758	1.227	2.117
μ_c	0.019	0.019	0.017	0.020	0.019	0.019	0.018	0.021
ρ_x	0.762	0.813	0.438	0.938	0.908	0.918	0.863	0.962
ϕ_x	0.151	0.148	0.086	0.227	0.189	0.184	0.145	0.245
λ	2.977	3.250	1.750	4.250	3.167	3.141	2.547	3.922
ϕ_d	2.846	2.750	1.750	4.250	4.602	4.594	3.891	5.344
ϕ_c	2.925	2.750	1.750	4.250	2.355	2.297	1.328	3.547
μ_s	0.021	0.021	0.013	0.028	0.021	0.021	0.020	0.022
ρ_s	0.938	0.938	0.938	0.938	0.950	0.950	0.950	0.950
σ_w	2.55E-04	2.59E-04	1.68E-04	3.51E-04	2.28E-04	2.31E-04	2.13E-04	2.46E-04
BIC						815.02		
Log likelihood						-381.61		
MCMC repetitions		10,000				12,000		

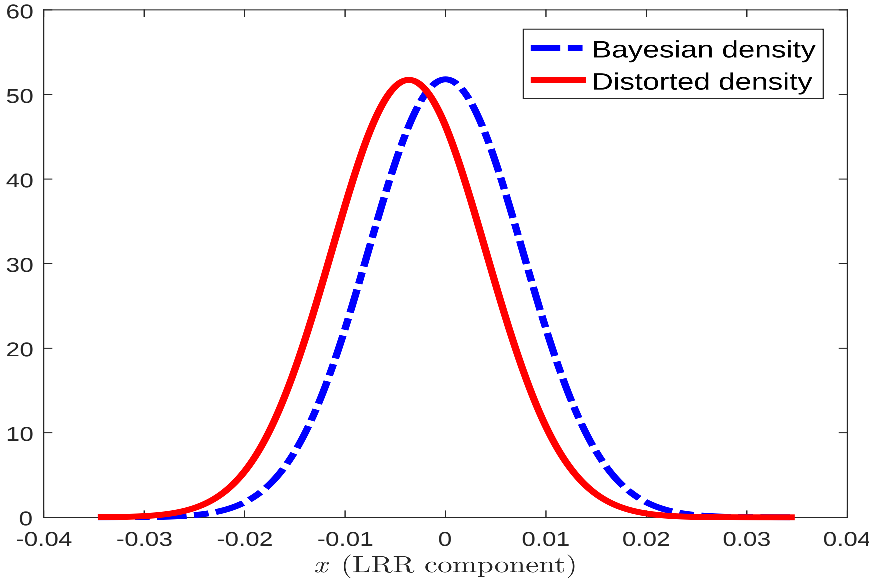
This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the EZLRRSV model. "BIC" represents the Bayesian information criteria, see Schwarz (1978). "Log likelihood" represents the maximum of the log likelihood of the encompassing model over the MCMC chain of estimates. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for the U.S. annual data 1941–2015.

Table 8: Financial Moments

	$\mathbb{E}(r_t^f)$	$\sigma(r_t^f)$	$\mathbb{E}(r_t - r_t^f)$	$\sigma(r_t - r_t^f)$	$\mathbb{E}(VRP_t)$	$\sigma(VRP_t)$	MPR
Data	1.41	2.82	5.32	17.77	11.07	24.94	N.A.
				AAMS			
Mean	1.595	1.541	5.812	18.441	12.955	9.300	1.280
Median	1.399	1.598	6.349	18.266	12.644	8.833	1.308
Std	0.800	0.286	1.738	2.220	2.967	2.720	0.341
95%	2.941	1.951	8.033	22.729	18.860	14.495	1.819
5%	0.444	1.951	3.010	15.378	8.332	5.411	0.758
				AAMSSV			
Mean	1.183	1.680	6.371	22.818	17.090	14.007	2.987
Median	1.267	1.673	5.910	22.579	14.678	10.430	2.790
Std	0.959	0.580	3.542	4.469	12.092	13.418	1.414
95%	2.465	2.685	14.446	30.242	35.407	39.337	5.970
5%	-0.622	0.800	1.650	15.633	4.835	3.878	1.205
				AALRRSV			
Mean	1.079	1.555	7.841	21.134	-0.744	1.752	1.312
Median	1.095	1.581	7.817	21.103	-0.288	1.086	1.142
Std	0.473	0.259	1.853	3.629	1.263	1.782	0.772
95%	1.818	1.946	10.838	28.345	0.406	6.490	2.976
5%	0.226	1.128	4.863	16.106	-3.218	0.438	0.489
				EZMS			
Mean	1.282	2.100	2.919	40.489	58.704	67.854	0.698
Median	1.278	2.035	2.727	41.471	54.164	65.354	0.647
Std	0.468	0.394	1.634	8.211	26.877	35.106	0.208
95%	2.107	2.902	6.007	52.464	107.602	128.264	1.128
5%	0.511	1.579	0.706	28.070	21.067	21.228	0.454
				EZLRRSV			
Mean	1.708	0.973	4.318	17.633	1.436	0.549	0.569
Median	1.686	0.856	4.458	17.560	1.398	0.524	0.573
Std	0.336	0.385	1.150	1.906	0.396	0.280	0.078
95%	2.289	1.589	5.999	20.903	2.137	1.130	0.692
5%	1.169	0.718	2.337	14.413	0.858	0.200	0.428

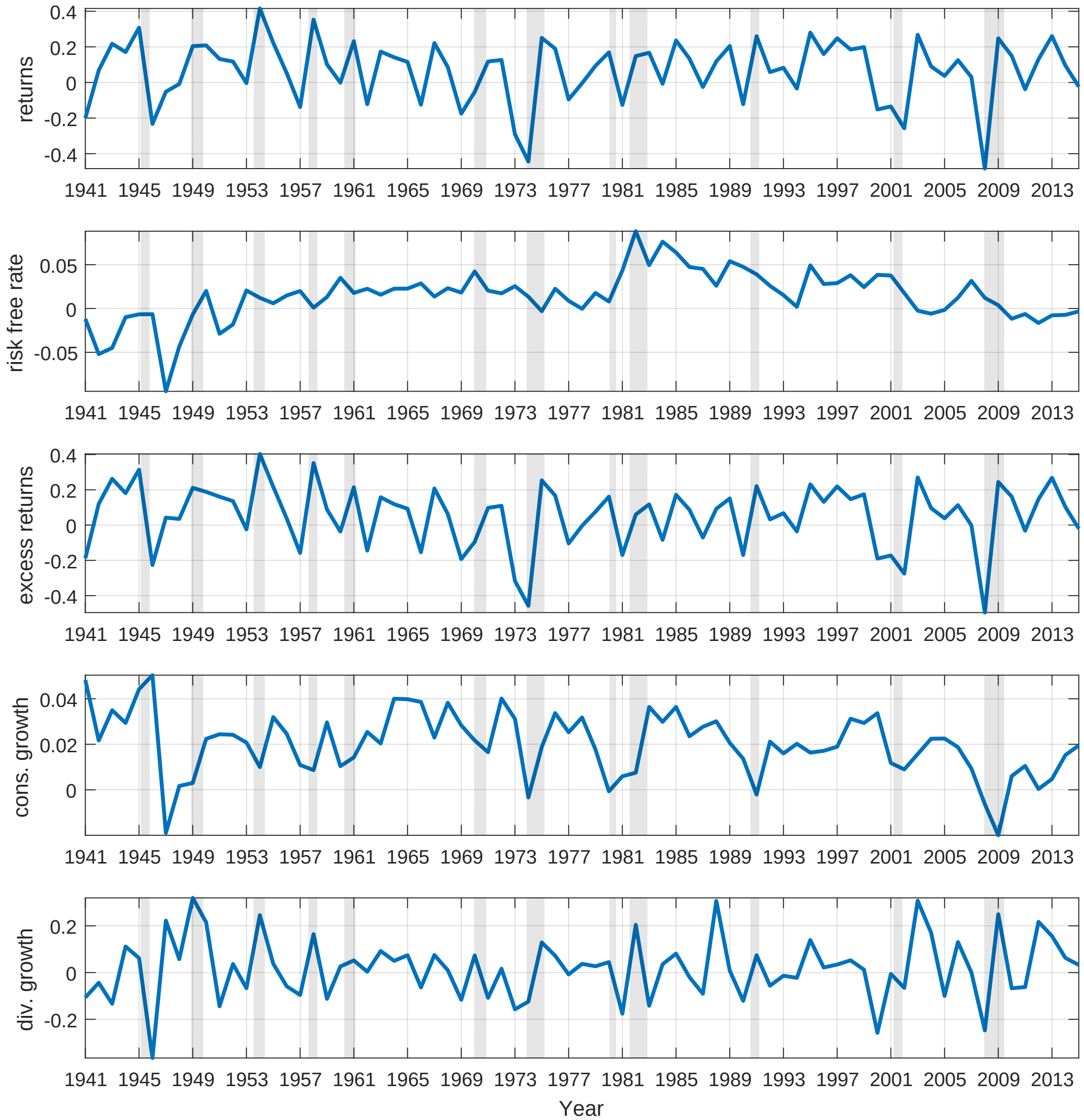
This table presents unconditional financial moments generated from the estimated models. These quantities are computed from simulated variables paths based on 12,000 Bayesian MCMC estimates of the structural parameters. $\mathbb{E}(r_t^f)$ and $\mathbb{E}(r_t - r_t^f)$ are mean risk-free rate and mean equity premium respectively (in percentage). $\sigma(r_t^f)$ and $\sigma(r_t - r_t^f)$ are volatilities of risk-free rates and excess returns respectively (in percentage). Moments of asset returns are computed based on annual data for the period 1941–2015. Variance risk premium (VRP) data covers the period 1996–2015. $\sigma(M_t)/\mathbb{E}(M_t)$ is the market price of risk.

Figure 1: **Model AALRRSV: Bayesian and distorted densities of x**



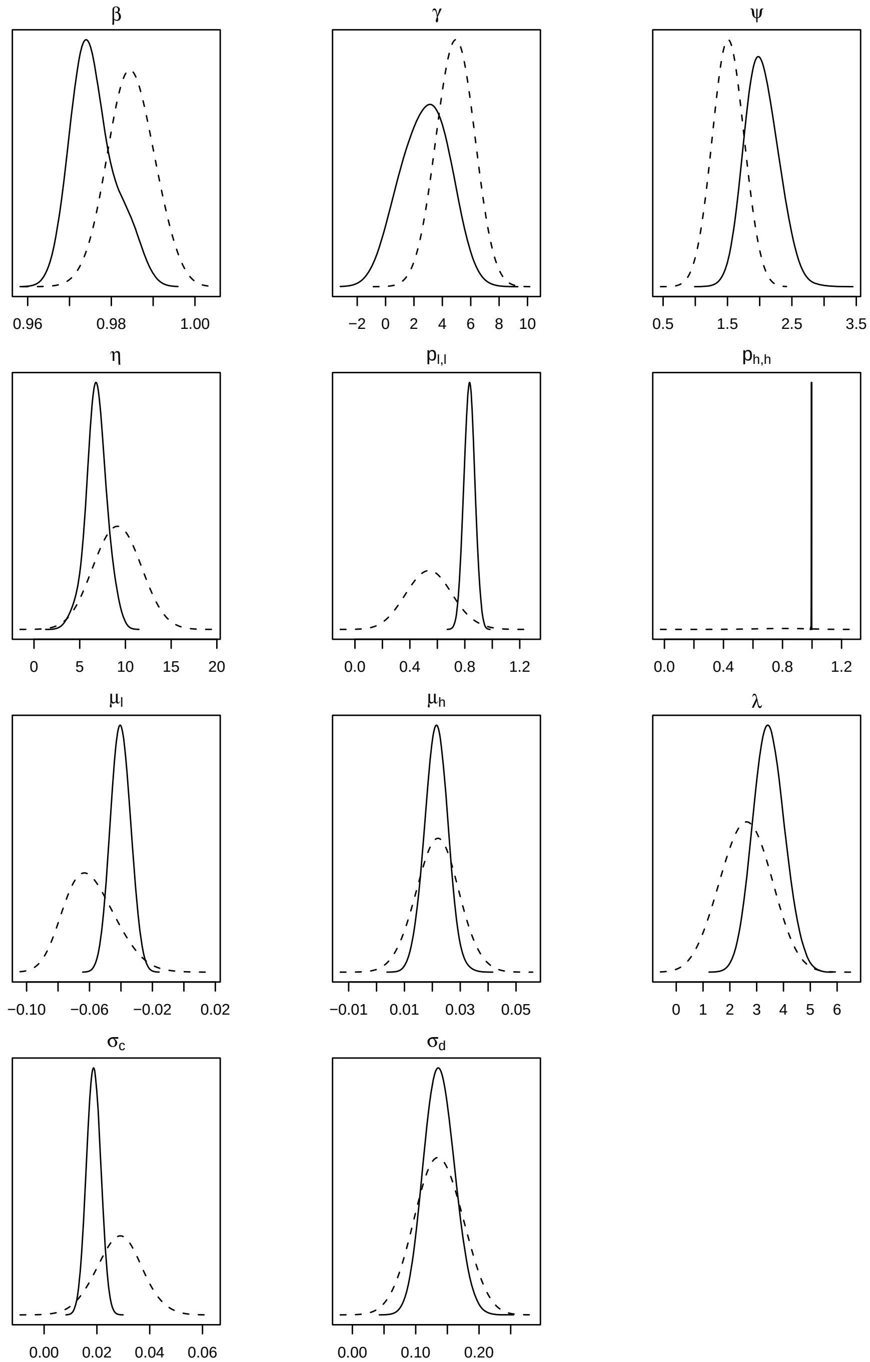
Notes: This figure plots Bayesian density and distorted density of the long-run risk component x for the AALRRSV model. The Bayesian density is $x_t \sim N(\hat{x}_t, \nu_t)$, and the distorted density is $\tilde{f}(x_t|\hat{x}_t, \nu_t, t)$. The distorted density is generated from solving the model. The state vector is assumed to take the value ($\hat{x}_t = 0$, $\nu_t = \bar{\nu}$ (steady-state) and $\sigma_t = \mu_s$). Model parameters are set at posterior mean estimates presented in Table 5.

Figure 2: Time-series of variables



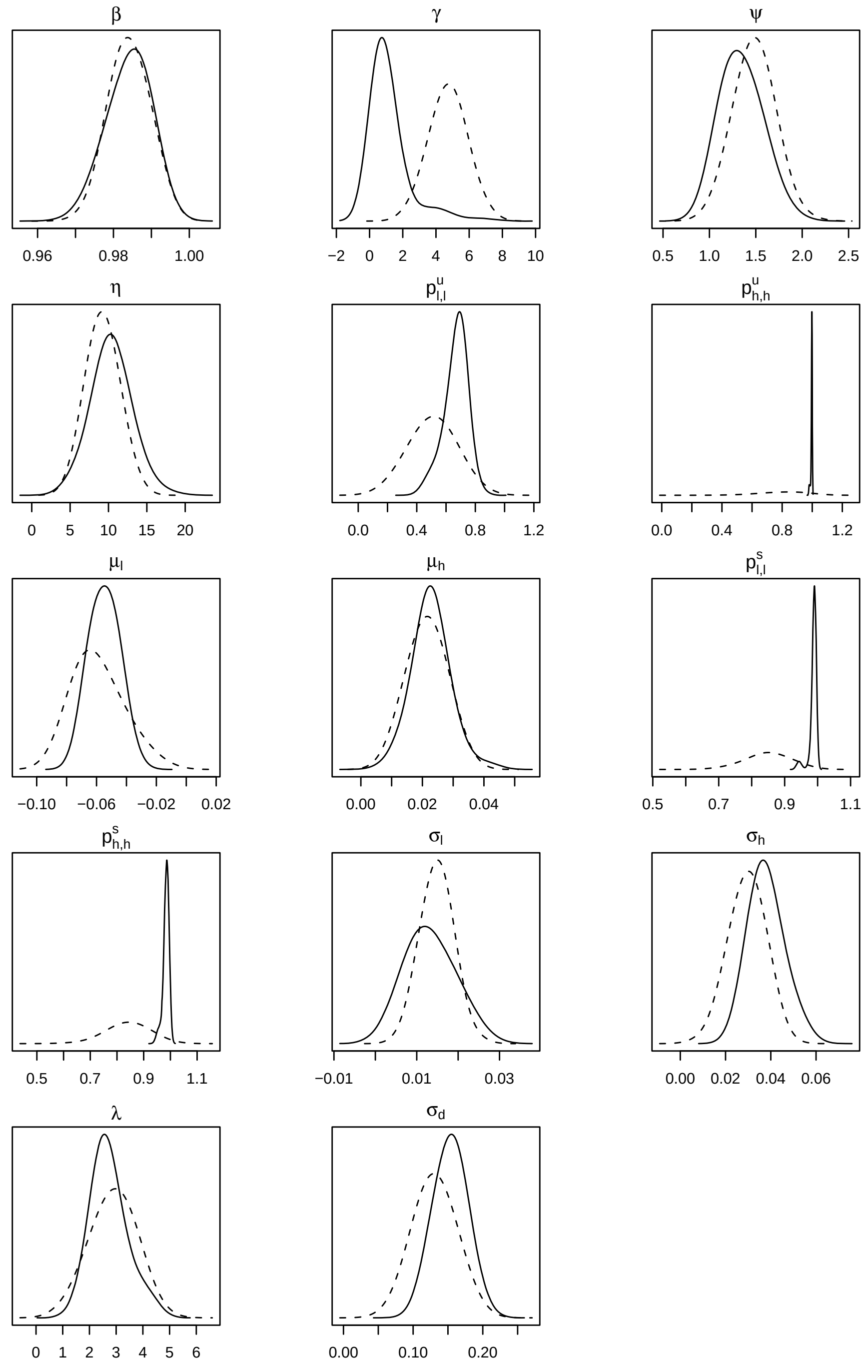
The figure shows CRSP value-weighted index returns, one-year Treasury Bill rates, excess returns, per-capita log consumption growth, and log dividend growth rates for the 1941–2015 period. All series plotted are at an annual frequency and in real terms. Shaded areas represent NBER recessions.

Figure 3: Prior and Posterior Densities of Estimated Parameters of AAMS Model



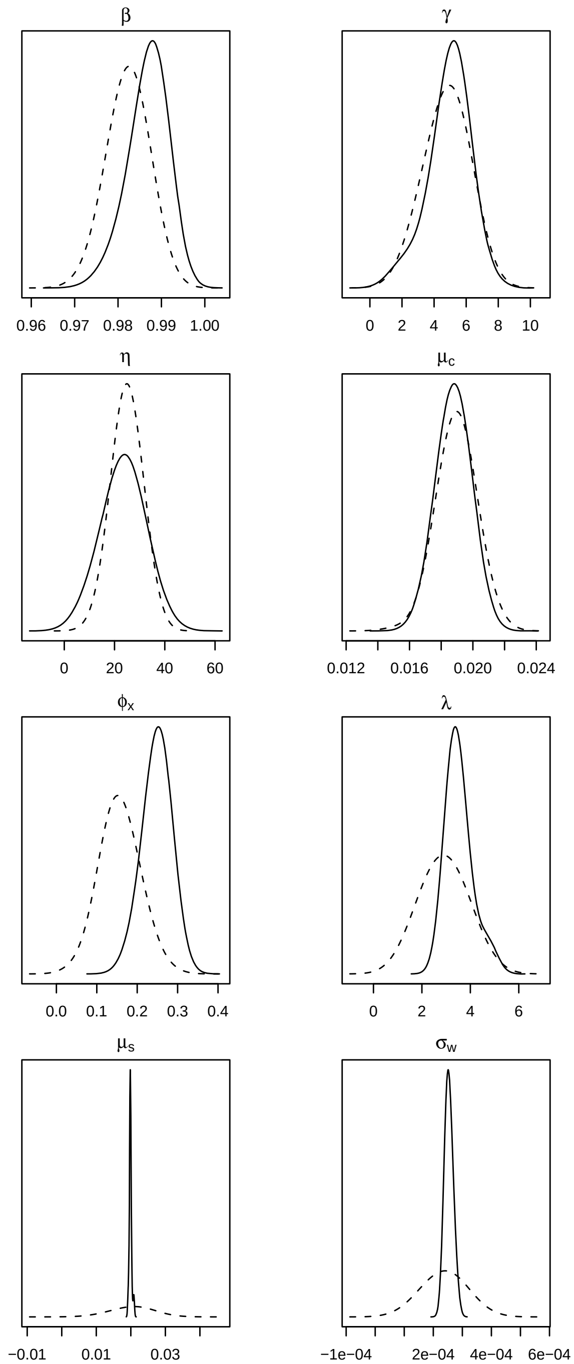
This figure plots prior and posterior densities of the Ju and Miao (2012) model parameters. The solid lines depict posterior densities and dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 4: Prior and Posterior Densities of Estimated Parameters of AAMSSV Model



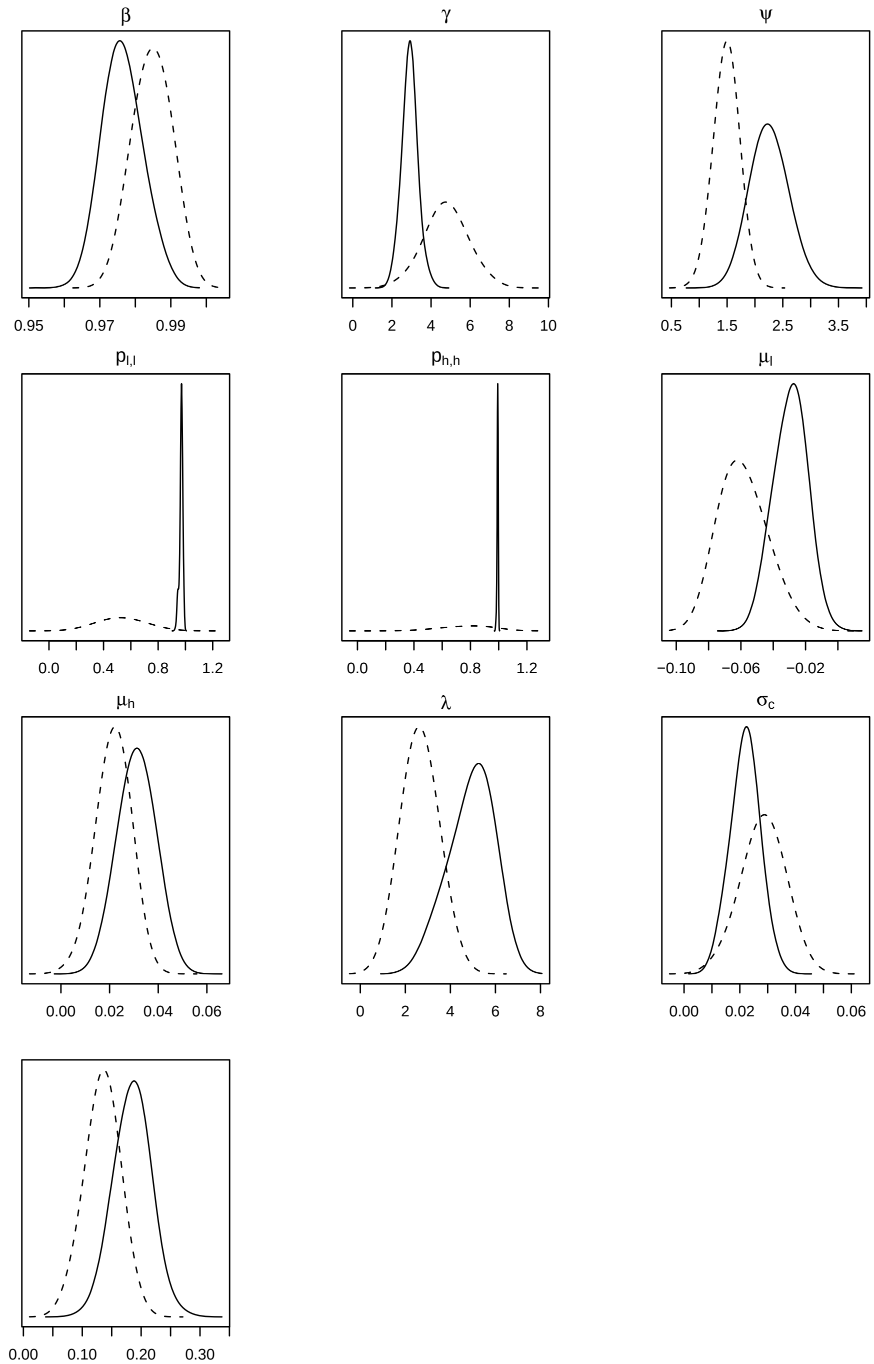
This figure plots prior and posterior densities of the AAMSSV model, featuring ambiguity aversion, Markov switching in both conditional mean and volatility of the consumption process. The solid lines depict posterior densities and dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 5: Prior and Posterior Densities of Estimated Parameters of AALRRSV Model



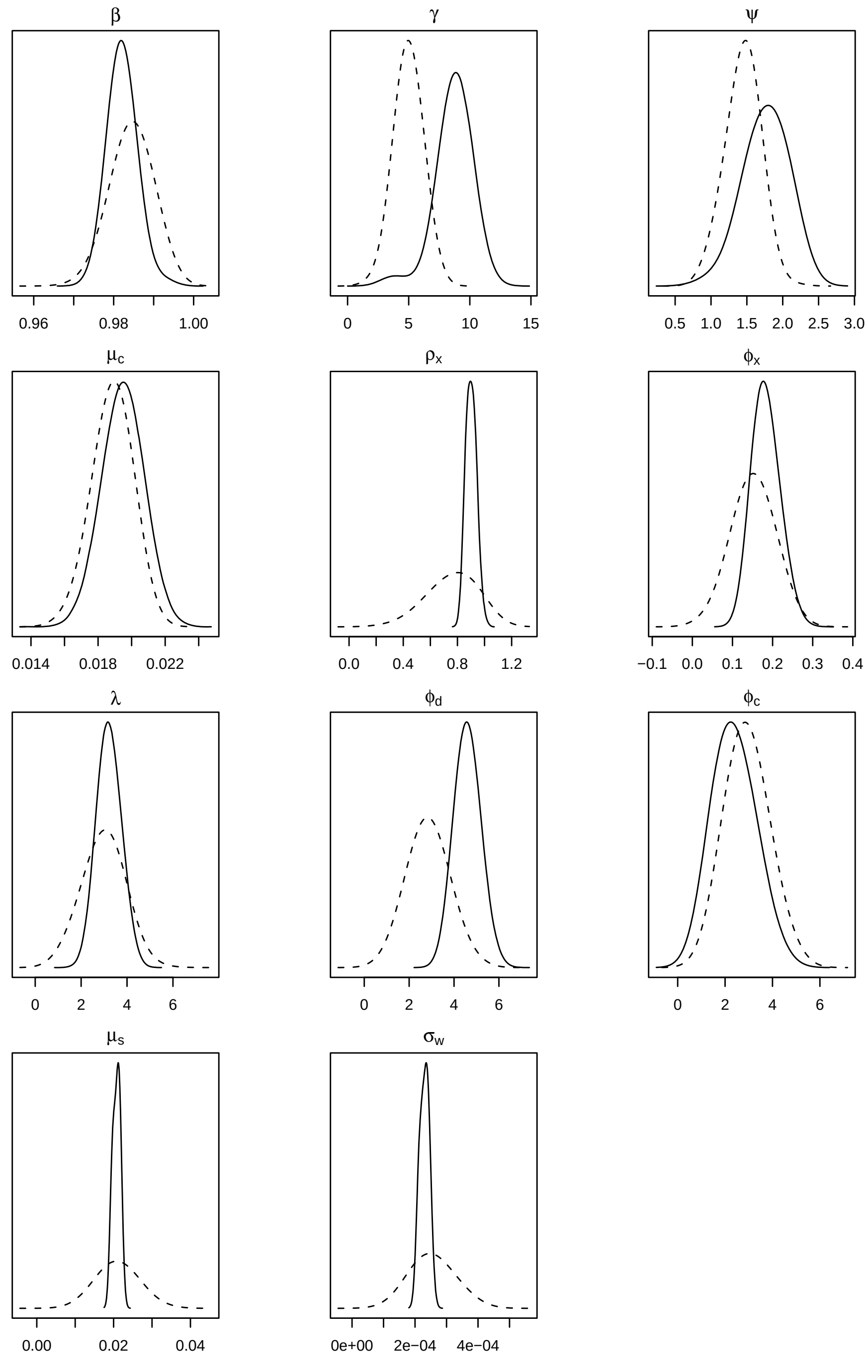
This figure plots prior and posterior densities of the AALRRSV model, featuring ambiguity aversion, Kalman learning, stochastic volatility, and long-run risks in the conditional mean of the consumption process. The solid lines depict posterior densities and dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 6: Prior and Posterior Densities of Estimated Parameters of EZMS Model



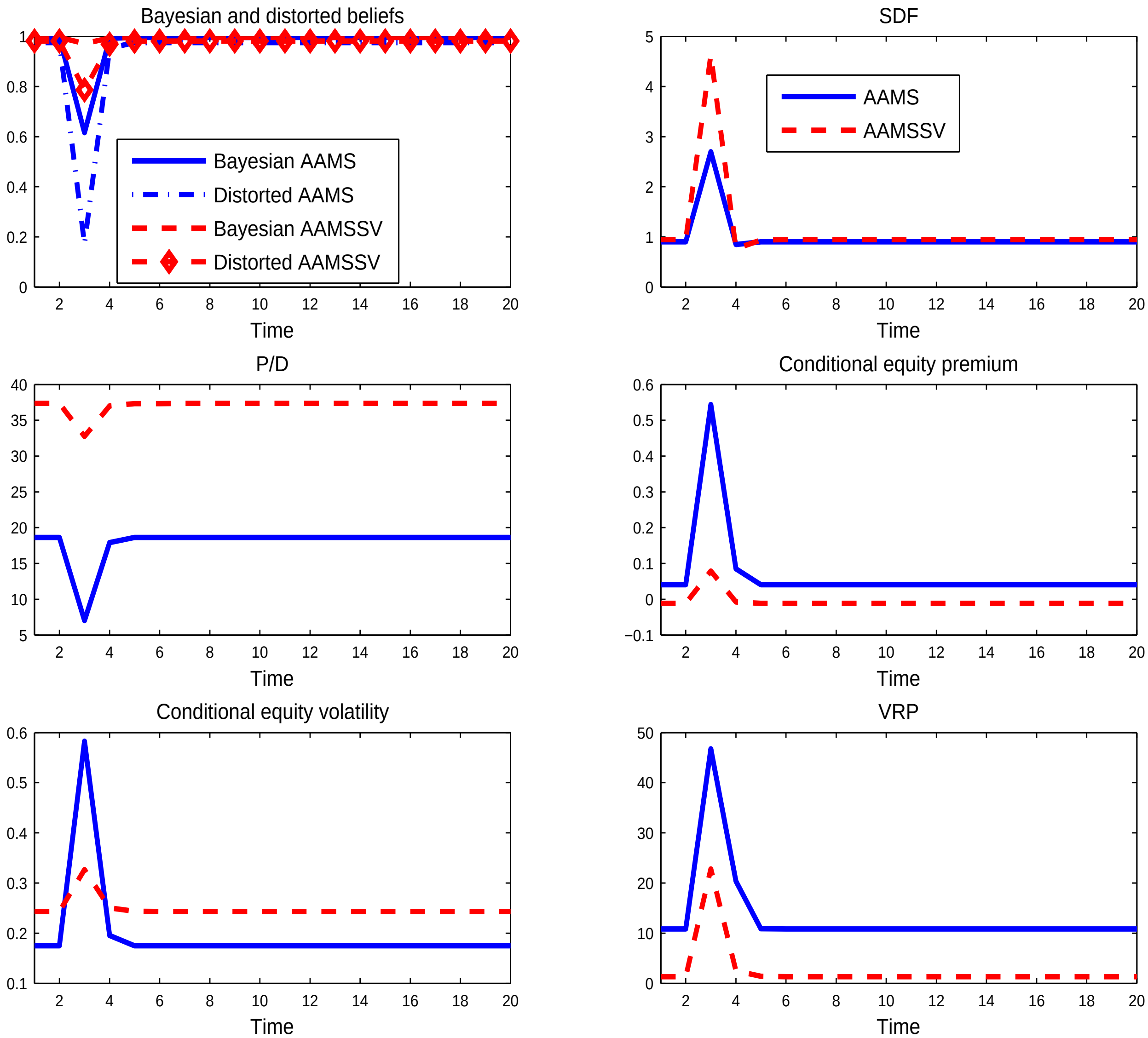
This figure plots prior and posterior densities of the EZMS model parameters, featuring Epstein and Zin preferences and Markov regimes in the conditional mean of the consumption growth process. The solid lines depict posterior densities and dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 7: Prior and Posterior Densities of Estimated Parameters of EZLRRSV BKY Model



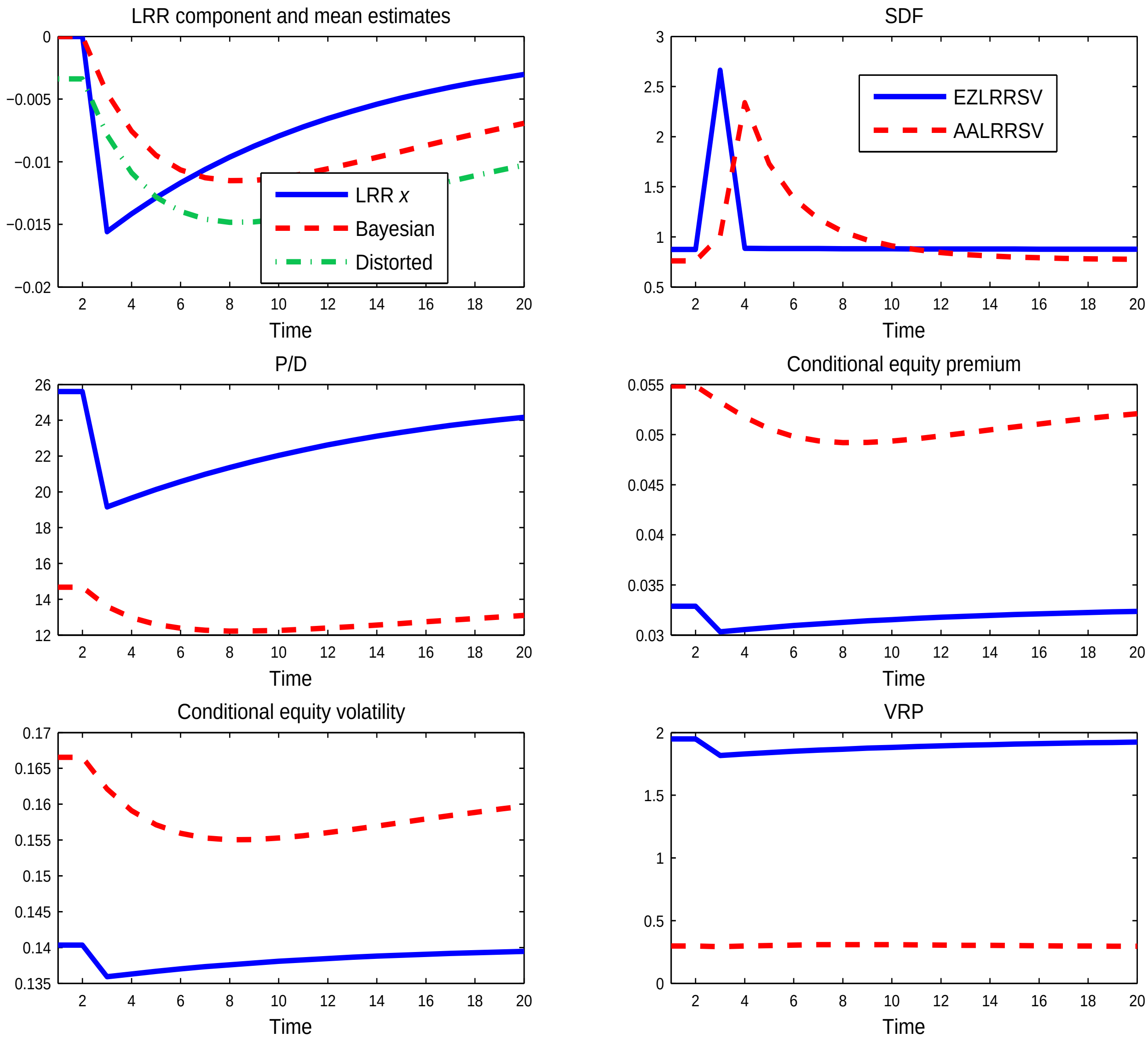
This figure plots prior and posterior densities of the EZLRRSV BKY model of Bansal et al. (2012), featuring Epstein-Zin preferences, stochastic volatility and long-run risks in the conditional mean of the consumption process. The solid lines depict posterior densities and dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 8: **Impulse responses: AAMSSV**



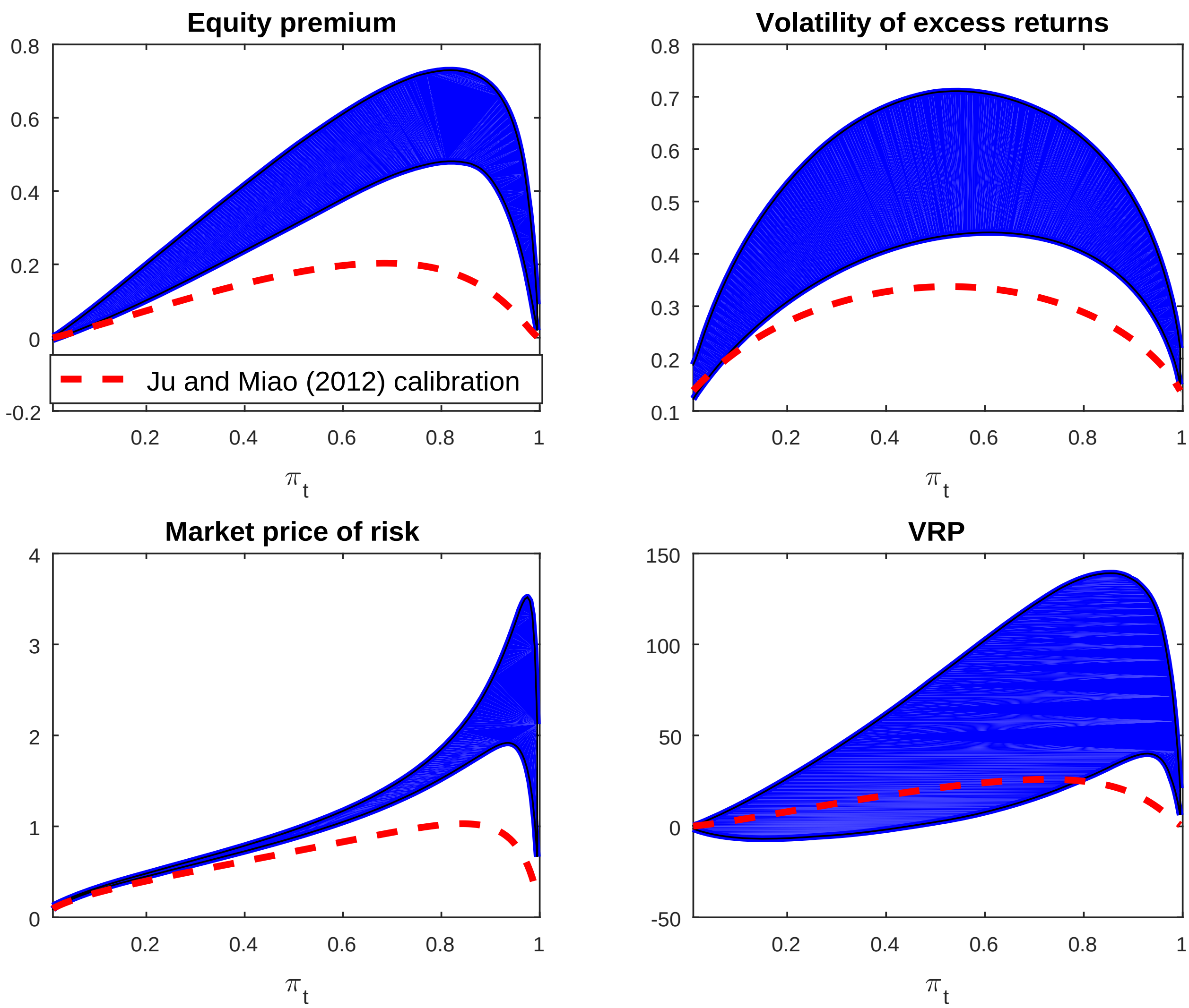
This figure plots the impulse response functions for models AAMS and AAMSSV when the mean consumption growth state shifts from μ_h to μ_l in the third period. Before the realization of the shock, mean consumption growth is assumed to stay in state μ_h without the impact of innovation shocks. For the AAMSSV model, the volatility state is assumed to be σ_h throughout all periods. The results plotted are for model parameters set at posterior means of Bayesian MCMC estimates.

Figure 9: **Impulse responses: AALRRSV**



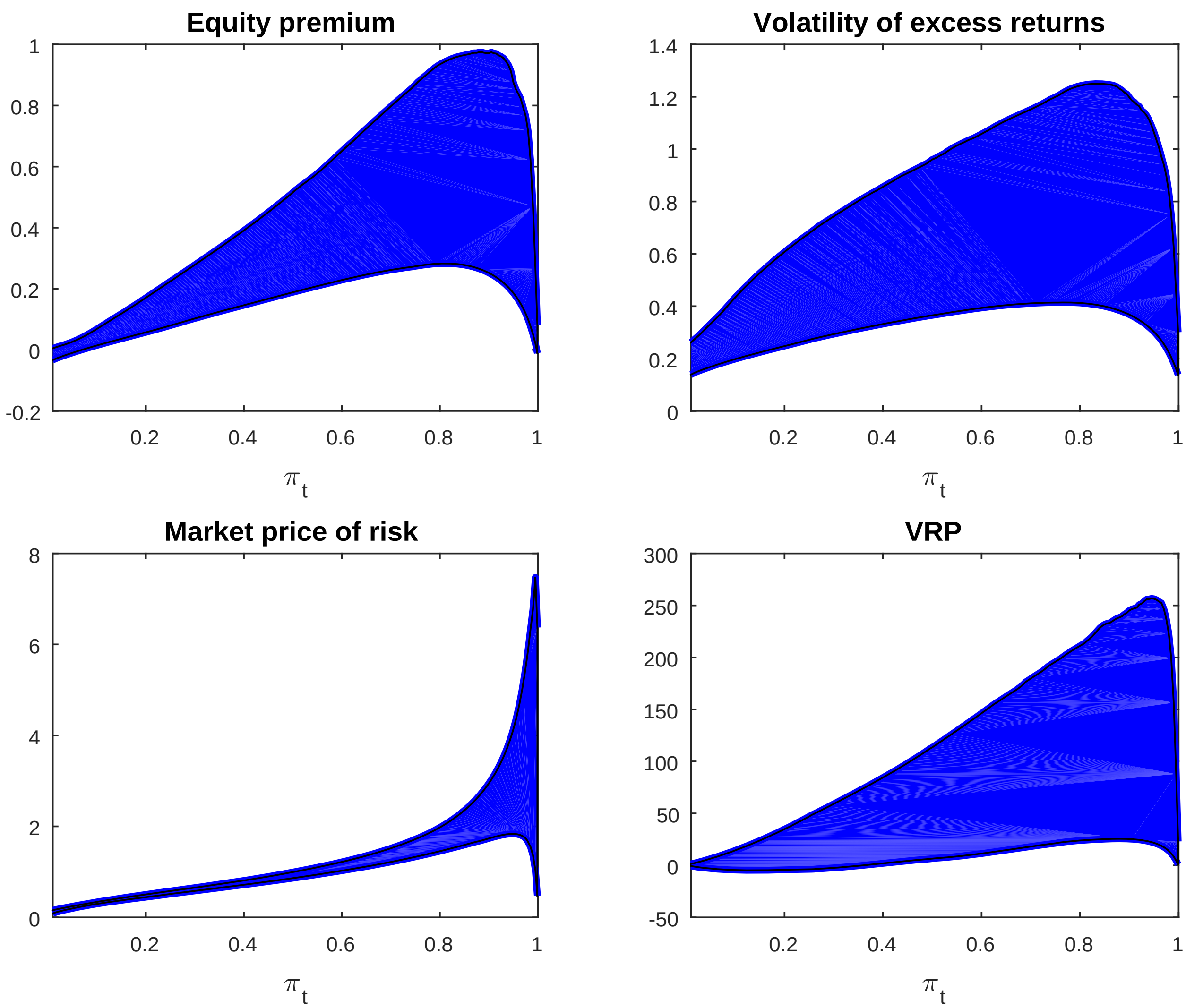
This figure plots the impulse response functions for models AALRRSV and EZLRRSV when a shock of size $-4\varphi_x\mu_s$ to x_t occurs in the third period. Before the realization of the shock, the AALRRSV economy is assumed to stay in state $(\hat{x}_t, \nu_t, \sigma_t)$ for which $\Delta c_t = \mu_c, \Delta d_t = \mu_d, x_t = 0, \sigma_t = \mu_s$ and $\nu_t = \bar{\nu}$ (steady-state) without the impact of innovation shocks. The distorted mean estimate is computed by applying the rejection sampling method and simulations. Before the realization of the shock, the EZLRRSV economy is assumed to stay in state $(x_t = 0, \sigma_t^2 = \mu_\sigma^2)$ without the impact of innovation shocks. The results plotted are for model parameters set at posterior means of Bayesian MCMC estimates.

Figure 10: **AAMS model: Conditional financial moments**



This figure plots conditional financial moments ranging from 5 to 95 percentile of simulated conditional moments for the AAMS model. The simulation is based on 12,000 Bayesian MCMC estimates of structural parameters. The dashed line plots the conditional moments calculated based on Ju and Miao's calibration.

Figure 11: **AAMSSV model: Conditional financial moments**



This figure plots conditional financial moments ranging from 5 to 95 percentile of simulated conditional moments for the AAMSSV model. The simulation is based on 12,000 Bayesian MCMC estimates of structural parameters.

1 Numerical methods

We use the collocation projection method with Chebyshev polynomials to solve asset pricing models in the paper. See Judd (1992) for an introduction to projection methods and Pohl et al. (2017) for applications to solving models with long-run risks.

We solve each model in two steps. In the first step, we use the projection method to solve the functional equation for the value function $V_t(C)$ to obtain the wealth-consumption ratio. Suppose that the vector of state variables for a model is denoted by z_t (e.g., $z_t = \{\pi_t\}$ in model AAMS). By homogeneity, we have $V_t(C) = C_t G(z_t)$ where $G(z_t)$ is a function to be determined. As shown by Epstein and Zin (1989), the wealth-consumption ratio W_t/C_t is given by

$$\frac{W_t}{C_t} = \frac{1}{1-\beta}\left(\frac{V_t}{C_t}\right)^{1-\frac{1}{\psi}}.$$

In the second step, we apply the projection method to solve the Euler equation to obtain the price-dividend ratio, given that we can determine the SDF $M_{t,t+1}$ from the solution in the first step. We denote the current state of the economy by z and the next period's state by z'.

1.1 Solving the AAMS Model

This model is developed by Ju and Miao (2012). See "Ambiguity, Learning, and Asset Returns: Technical Appendix" for details about the numerical method.

The functional equation for $G(\pi)$ implied by the generalized recursive smooth ambiguity utility function is given by

$$G(\pi) = \left[(1-\beta) + \beta\left(\mathbb{E}\left[\left(\mathbb{E}\left[G(\pi')^{1-\gamma}\exp\left((1-\gamma)\Delta c(s')\right)\Big| s'\right]\right)^{\frac{1-\eta}{1-\gamma}}\Big|\pi\right]\right)^{\frac{1-1/\psi}{1-\eta}}\right]^{\frac{1}{1-1/\psi}}. \tag{1}$$

The intertemporal marginal rate of substitution (or stochastic discount factor) is given by

$$\begin{aligned} M(\pi', s'|\pi) &= \beta\exp\left(-\frac{1}{\psi}\Delta c(s')\right)\left(\frac{G(\pi')\exp(\Delta c(s'))}{\mathcal{R}(G(\pi')\exp(\Delta c(s'))|\pi)}\right)^{\frac{1}{\psi}-\gamma} \\ &\times\left(\frac{\left(\mathbb{E}\left[G(\pi')^{1-\gamma}\exp((1-\gamma)\Delta c(s'))\Big| s',\pi\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}(G(\pi')\exp(\Delta c(s'))|\pi)}\right)^{-(\eta-\gamma)}. \end{aligned}$$

The price-dividend ratio $\varphi(\pi)$ satisfies the Euler equation

$$\varphi(\pi) = \mathbb{E}\left[M(\pi', s'|\pi)(1+\varphi(\pi'))\exp(\Delta d(s'))\big|\pi\right]. \tag{2}$$

The laws of motion of consumption and dividend growth are

$$\begin{aligned}\Delta c(s) &= \mu(s)+\sigma_c\epsilon_c, \quad \epsilon_c \sim N(0,1)\\ \Delta d(s) &= \lambda\Delta c(s)+g_d+\tilde{\sigma}_d\epsilon_d,\ \epsilon_d\sim N(0,1)\end{aligned}$$

where the transition probabilities are

$$\Pr\left(s'=l|s=l\right)=p_{ll},\ \ \Pr\left(s'=h|s=h\right)=p_{hh}$$

and ϵ_c and ϵ_d are two independent innovation shocks.

The (nonlinear) law of motion of the state variable π is

$$\pi'=\frac{p_{hh}f\left(\Delta c\left(s'\right)|s'=h\right)\pi+\left(1-p_{ll}\right)f\left(\Delta c\left(s'\right)|s'=l\right)(1-\pi)}{f\left(\Delta c\left(s'\right)|s'=h\right)\pi+f\left(\Delta c\left(s'\right)|s'=l\right)(1-\pi)}.$$

We approximate the solution functions $G(\pi)$ and $\varphi(\pi)$ by Chebyshev polynomials, namely,

$$\hat{G}\left(\pi;\boldsymbol{a}^G\right)=\sum_{k=0}^{n_\pi}a_k^G T_k\left(t_\pi\right),\ \ \hat{\varphi}\left(\pi;\boldsymbol{a}^\varphi\right)=\sum_{k=0}^{n_\pi}a_k^\varphi T_k\left(t_\pi\right)$$

where $T_k:[-1,1]\to\mathbb{R},\ k=0,1,...,n_\pi$ are Chebyshev polynomials and the transformation of the argument for the polynomial is given by

$$t_\pi=2\left(\frac{\pi-\pi_{\min}}{\pi_{\max}-\pi_{\min}}\right)-1$$

with $\pi_{\min}=0$ and $\pi_{\max}=1$. To implement the collocation method, we solve the two functional equations (1) and (2) on a grid of π obtained by applying the inverse of the transformation to the $n_\pi+1$ zeros of the Chebyshev polynomial $T_{n_\pi+1}$.

Equations (1) and (2) define two residual functions that are to be minimized sequentially by choosing the coefficients $\boldsymbol{a}^G$ and $\boldsymbol{a}^\varphi$. The collocation projection method leads to two square systems of nonlinear equations, which can be solved with a nonlinear equations solver (e.g., Powell's hybrid algorithm). Because the underlying innovation shocks are Gaussian, we use Gauss-Hermite quadrature to calculate conditional expectations in the residual functions.

1.2 Solving the AAMSSV Model

Compared to AAMS, the AAMSSV model has one additional state variable s_t^σ indicating the volatility state. It follows that $G(\pi,s^\sigma)$ satisfies the equation

$$G\left(\pi,s^\sigma\right)=\left[(1-\beta)+\beta\left(\mathbb{E}\left[\left(\mathbb{E}\left[G\left(\pi',s^{\sigma\prime}\right)^{1-\gamma}\exp\left((1-\gamma)\Delta c\left(s^{\mu\prime},s^{\sigma\prime}\right)\right)\middle|s^{\mu\prime},s^\sigma\right]\right)^{\frac{1-\eta}{1-\gamma}}\middle|\pi\right]\right)^{\frac{1-1/\psi}{1-\eta}}\right]^{\frac{1}{1-1/\psi}}.$$

The SDF and Euler equation are given by

$$
\begin{aligned}
M\left(\pi', s^{\sigma\prime}, s^{\mu\prime}|\pi, s^{\sigma}\right) &= \beta \exp\left(-\frac{1}{\psi}\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)\left(\frac{G\left(\pi', s^{\sigma\prime}\right)\exp\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)}{\mathcal{R}\left(\left.G\left(\pi', s^{\sigma\prime}\right)\exp\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)\right|\pi, s^{\sigma}\right)}\right)^{\frac{1}{\psi}-\gamma} \\
&\times\left(\frac{\left(\mathbb{E}\left[\left.G\left(\pi', s^{\sigma\prime}\right)^{1-\gamma}\exp\left(\left(1-\gamma\right)\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)\right| s^{\mu\prime}, \pi, s^{\sigma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}\left(\left.G\left(\pi', s^{\sigma\prime}\right)\exp\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)\right|\pi, s^{\sigma}\right)}\right)^{-(\eta-\gamma)}
\end{aligned}
$$

and

$$
\varphi\left(\pi, s^{\sigma}\right) = \mathbb{E}\left[\left.M\left(\pi', s^{\sigma\prime}, s^{\mu\prime}|\pi, s^{\sigma}\right)\left(1+\varphi\left(\pi', s^{\sigma\prime}\right)\right)\exp\left(\Delta d\left(s^{\mu\prime}, s^{\sigma\prime}\right)\right)\right|\pi, s^{\sigma}\right].
$$

The laws of motions for consumption and dividend growth are

$$
\begin{aligned}
\Delta c\left(s^{\mu}, s^{\sigma}\right) &= \mu\left(s^{\mu}\right) + \sigma\left(s^{\sigma}\right)\epsilon_c, \quad \epsilon_c \sim N\left(0,1\right) \\
\Delta d\left(s^{\mu}, s^{\sigma}\right) &= \lambda\Delta c\left(s^{\mu}, s^{\sigma}\right) + g_d + \tilde{\sigma}_d\epsilon_d, \ \epsilon_d \sim N\left(0,1\right)
\end{aligned}
$$

where the transition probabilities for the two independent Markov chains of s^{μ} and s^{σ} are given by

$$
\begin{aligned}
\Pr\left(s^{\sigma\prime} = l|s^{\sigma} = l\right) &= p_{ll}^{\sigma}, \ \Pr\left(s^{\sigma\prime} = h|s^{\sigma} = h\right) = p_{hh}^{\sigma} \\
\Pr\left(s^{\mu\prime} = l|s^{\mu} = l\right) &= p_{ll}^{\mu}, \ \Pr\left(s^{\mu\prime} = h|s^{\mu} = h\right) = p_{hh}^{\mu}
\end{aligned}
$$

The law of motion of the state variable π is given by the Bayes' rule

$$
\pi' = \frac{p_{hh}^{\mu} f\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)|s^{\mu\prime} = h, s^{\sigma\prime}\right)\pi + \left(1-p_{ll}^{\mu}\right) f\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)|s^{\mu\prime} = l, s^{\sigma\prime}\right)\left(1-\pi\right)}{f\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)|s^{\mu\prime} = h, s^{\sigma\prime}\right)\pi + f\left(\Delta c\left(s^{\mu\prime}, s^{\sigma\prime}\right)|s^{\mu\prime} = l, s^{\sigma\prime}\right)\left(1-\pi\right)}
$$

We approximate the solutions to $G\left(\pi, s^{\sigma}\right)$ and $\varphi\left(\pi, s^{\sigma}\right)$ by Chebyshev polynomials as

$$
\begin{aligned}
\hat{G}\left(\pi, s^{\sigma} = l; \boldsymbol{a}_l^G\right) &= \sum_{k=0}^{n_\pi} a_{k,l}^G T_k\left(t_\pi\right), \ \ \hat{G}\left(\pi, s^{\sigma} = h; \boldsymbol{a}_h^G\right) = \sum_{k=0}^{n_\pi} a_{k,h}^G T_k\left(t_\pi\right) \\
\hat{\varphi}\left(\pi, s^{\sigma} = l; \boldsymbol{a}_l^{\varphi}\right) &= \sum_{k=0}^{n_\pi} a_{k,l}^{\varphi} T_k\left(t_\pi\right), \ \ \hat{\varphi}\left(\pi, s^{\sigma} = h; \boldsymbol{a}_h^{\varphi}\right) = \sum_{k=0}^{n_\pi} a_{k,h}^{\varphi} T_k\left(t_\pi\right)
\end{aligned}
$$

i.e., we seek four sets of coefficients $\left(\boldsymbol{a}_l^G, \boldsymbol{a}_h^G, \boldsymbol{a}_l^{\varphi}, \boldsymbol{a}_h^{\varphi}\right)$ that minimize the residual functions.

1.3 Solving the AALRRSV Model

We consider the long-run risk model

$$
\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_{t+1} + \sigma_t \epsilon_{c,t+1} \\
\Delta d_{t+1} &= \mu_d + \lambda x_{t+1} + \varphi_d \sigma_t \epsilon_{d,t+1} \\
x_{t+1} &= \rho_x x_t + \varphi_x \sigma_t \epsilon_{x,t+1} \\
\sigma_{t+1}^2 &= \mu_\sigma^2 + \rho_s \left(\sigma_t^2 - \mu_s^2\right) + \sigma_w \epsilon_{w,t+1} \\
\epsilon_{c,t+1}, \epsilon_{d,t+1}, \epsilon_{x,t+1}, \epsilon_{w,t+1} &\sim i.i.d. N(0,1).
\end{aligned}
$$

where the long-run risk component x_t is unobservable. We define $\hat{x}_{t+1|t} = E\left[x_{t+1}|\mathcal{I}_t\right]$ and $\nu_{t+1|t} = E\left[\left(x_{t+1} - \hat{x}_{t+1|t}\right)^2 |\mathcal{I}_t\right]$ where $\mathcal{I}_t$ denotes available information at time t. It immediately follows that

$$\hat{x}_{t+1|t} = \rho_x \hat{x}_t, \text{ and } \nu_{t+1|t} = \rho_x^2 \nu_t + \varphi_x^2 \sigma_t^2.$$

The Kalman filter implies the following updating equations

$$
\begin{aligned}
\hat{x}_{t+1} &= \hat{x}_{t+1|t} + \nu_{t+1|t} \begin{bmatrix} 1 & \lambda \end{bmatrix} F_{t+1|t}^{-1} \begin{bmatrix} v_{t+1|t}^c \\ v_{t+1|t}^d \end{bmatrix} \\
\nu_{t+1} &= \nu_{t+1|t} - \nu_{t+1|t}^2 \begin{bmatrix} 1 & \lambda \end{bmatrix} F_{t+1|t}^{-1} \begin{bmatrix} 1 & \lambda \end{bmatrix}'
\end{aligned}
$$

where $F_{t+1|t}$ is given by

$$F_{t+1|t} = \begin{bmatrix} \nu_{t+1|t} + \sigma_t^2 & \lambda \nu_{t+1|t} \\ \lambda \nu_{t+1|t} & \lambda^2 \nu_{t+1|t} + \varphi_d^2 \sigma_t^2 \end{bmatrix}$$

and the innovation vector $\begin{bmatrix} v_{t+1|t}^c & v_{t+1|t}^d \end{bmatrix}$ is given by

$$\begin{bmatrix} v_{t+1|t}^c \\ v_{t+1|t}^d \end{bmatrix} = \begin{bmatrix} \Delta c_{t+1} - \mu_c - \rho_x \hat{x}_t \\ \Delta d_{t+1} - \mu_d - \lambda \rho_x \hat{x}_t \end{bmatrix}.$$

Expressed as an intertemporal equation, the solution function $G(\hat{x}, \nu, \sigma)$ satisfies

$$G(\hat{x}, \nu, \sigma) = \left[(1-\beta) + \beta \left(\mathbb{E}\left[\left(\mathbb{E}\left[G(\hat{x}', \nu', \sigma')^{1-\gamma} \exp\left((1-\gamma)\Delta c(x')\right) \Big| x, \sigma \right] \right)^{\frac{1-\eta}{1-\gamma}} \Big| \hat{x}, \nu \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\eta}} \right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

The SDF and Euler equation are given by

$$
\begin{aligned}
M\left(x', x, \hat{x}', \nu', \sigma' | \hat{x}, \nu, \sigma\right) &= \beta \exp\left(-\frac{1}{\psi}\Delta c\left(x'\right)\right)\left(\frac{G\left(\hat{x}', \nu', \sigma'\right)\exp\left(\Delta c\left(x'\right)\right)}{\mathcal{R}\left(G\left(\hat{x}', \nu', \sigma'\right)\exp\left(\Delta c\left(x'\right)\right)\middle|\, \hat{x}, \nu, \sigma\right)}\right)^{\frac{1}{\psi}-\gamma} \\
&\quad\times\left(\frac{\left(\mathbb{E}\left[G\left(\hat{x}', \nu', \sigma'\right)^{1-\gamma}\exp\left(\left(1-\gamma\right)\Delta c\left(x'\right)\right)\middle|\, x, \hat{x}, \nu, \sigma\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}\left(G\left(\hat{x}', \nu', \sigma'\right)\exp\left(\Delta c\left(x'\right)\right)\middle|\, \hat{x}, \nu, \sigma\right)}\right)^{-(\eta-\gamma)}
\end{aligned}
$$

and

$$
\varphi\left(\hat{x}, \nu, \sigma\right) = \mathbb{E}\left[M\left(x', x, \hat{x}', \nu', \sigma' | \hat{x}, \nu, \sigma\right)\left(1+\varphi\left(\hat{x}', \nu', \sigma'\right)\right)\exp\left(\Delta d\left(x'\right)\right)\middle|\, \hat{x}, \nu, \sigma\right].
$$

We approximate the solution functions $G\left(\hat{x}, \nu, \sigma\right)$ and $\varphi\left(\hat{x}, \nu, \sigma\right)$ by three-dimensional product Chebyshev polynomials, namely,

$$
\begin{aligned}
\hat{G}\left(\hat{x}, \nu, \sigma; \boldsymbol{a}^G\right) &= \sum_{k_{\hat{x}}=0}^{n_{\hat{x}}}\sum_{k_\nu=0}^{n_\nu}\sum_{k_\sigma=0}^{n_\sigma} a^G_{k_{\hat{x}}} a^G_{k_\nu} a^G_{k_\sigma} T_{k_{\hat{x}}}\left(t_{\hat{x}}\right) T_{k_\nu}\left(t_\nu\right) T_{k_\sigma}\left(t_\sigma\right) \\
\hat{\varphi}\left(\hat{x}, \nu, \sigma; \boldsymbol{a}^\varphi\right) &= \sum_{k_{\hat{x}}=0}^{n_{\hat{x}}}\sum_{k_\nu=0}^{n_\nu}\sum_{k_\sigma=0}^{n_\sigma} a^\varphi_{k_{\hat{x}}} a^\varphi_{k_\nu} a^\varphi_{k_\sigma} T_{k_{\hat{x}}}\left(t_{\hat{x}}\right) T_{k_\nu}\left(t_\nu\right) T_{k_\sigma}\left(t_\sigma\right).
\end{aligned}
$$

In constructing Chebyshev polynomials as basis functions, we obtain the lower and upper bounds for each state variable by simulations. Because $x_t \sim N\left(\hat{x}_t, \nu_t\right)$, we use Gauss-Hermite quadrature to compute the conditional expectation involving state x_t. To compute conditional expectations with respect to the underlying shocks $\left(\epsilon_c, \epsilon_d, \epsilon_x, \epsilon_\sigma\right)$, we apply the monomial method with degree 5, see Judd (1999) for details of the monomial method. If the dimension of underlying shocks is d, the monomial method requires $2d^2+1$ points to compute an expectation, whereas Gauss-Hermite quadrature requires N^d nodes with N being the number of nodes in one dimension. When the dimension of underlying shocks is large, the monomial method is much more efficient than quadrature methods. This gain in efficiency is particularly important for our structural estimation. A number of simulations suggest that for our model the monomial method yields accurate results compared with Gauss-Hermite quadrature.

To implement the collocation method, we solve the two square systems of nonlinear equations derived from equilibrium conditions on a grid of dimension $\left(n_{\hat{x}}+1\right)\times\left(n_\nu+1\right)\times\left(n_\sigma+1\right)$ for the state variables. The grid is constructed from zeros of Chebyshev polynomials of all state variables.

An alternative approach is to discretize the AR(1) process of σ_t^2 into a $n-$state Markov chain by the method developed in Tauchen (1986). Caldara et al. (2012) adopt this approach to solve DSGE models with recursive preferences and stochastic volatility. To avoid negative volatility states in the Markov chain, we keep positive values only and normalize transition probabilities accordingly. As such, given each volatility state σ_i, the solution functions $G\left(\hat{x}, \nu, \sigma_i\right)$ and $\varphi\left(\hat{x}, \nu, \sigma_i\right)$ can be approximated by two-dimensional product Chebyshev polynomials in $\hat{x}$ and ν. Through simulations, we find that this approach yields results that are close to the approximation with three-dimensional product Chebyshev polynomials.

1.4 Solving the EZLRRSV Model

The laws of motion of Δc, Δd, x and σ are given by the long-run risk model

$$\begin{aligned}
\Delta c &= \mu_c + x + \sigma\epsilon_c \\
\Delta d &= \mu_d + \lambda x + \varphi_d \sigma \epsilon_d + \varphi_c \sigma \epsilon_c \\
x' &= \rho_x x + \varphi_x \sigma \epsilon_x \\
\sigma^{2\prime} &= \mu_\sigma^2 + \rho_\sigma \left(\sigma^2 - \mu_\sigma^2\right) + \sigma_w \epsilon_w \\
\epsilon_c, \epsilon_d, \epsilon_x, \epsilon_w &\sim i.i.d.N\left(0,1\right).
\end{aligned}$$

The solution function $G\left(x,\sigma^2\right)$ satisfies

$$G\left(x,\sigma^2\right) = \left[\left(1-\beta\right) + \beta\left(\mathbb{E}\left[G\left(x',\sigma^{2\prime}\right)^{1-\gamma}\exp\left(\left(1-\gamma\right)\Delta c\right)\Big| x,\sigma^2\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

The SDF and Euler equation are given by

$$M\left(x',\sigma^{2\prime}|x,\sigma^2\right) = \beta\exp\left(-\frac{1}{\psi}\Delta c\right)\left(\frac{G\left(x',\sigma^{2\prime}\right)\exp\left(\Delta c\right)}{\mathcal{R}\left(G\left(x',\sigma^{2\prime}\right)\exp\left(\Delta c\right)|\, x,\sigma^2\right)}\right)^{\frac{1}{\psi}-\gamma}$$

and

$$\varphi\left(x,\sigma^2\right) = \mathbb{E}\left[M\left(x',\sigma^{2\prime}|x,\sigma^2\right)\left(1+\varphi\left(x,\sigma^2\right)\right)\exp\left(\Delta d\right)\big| x,\sigma^2\right]$$

We approximate the solution functions $G\left(x,\sigma^2\right)$ and $\varphi\left(x,\sigma^2\right)$ by two-dimensional product Chebyshev polynomials in x and σ^2:

$$\begin{aligned}
\hat{G}\left(x,\sigma^2;\boldsymbol{a}^G\right) &= \sum_{k_x=0}^{n_x}\sum_{k_\sigma=0}^{n_\sigma} a_{k_x}^G a_{k_\sigma}^G T_{k_x}\left(t_x\right) T_{k_\sigma}\left(t_\sigma\right) \\
\hat{\varphi}\left(x,\sigma^2;\boldsymbol{a}^\varphi\right) &= \sum_{k_x=0}^{n_{\hat{x}}}\sum_{k_\sigma=0}^{n_\sigma} a_{k_x}^\varphi a_{k_\sigma}^\varphi T_{k_x}\left(t_x\right) T_{k_\sigma}\left(t_\sigma\right).
\end{aligned}$$

2 Numerical accuracy of the solution method

We use the method proposed by Judd (1992) to assess numerical accuracy of our numerical solutions. The numerical accuracy check is through computing the Euler equation error. Previous studies such as Guerrieri and Iacoviello (2015) and Collard, Mukerji, Sheppard, and Tallon (2017) rely on this approach to assess the accuracy of their numerical solutions. Note that instead of computing the Euler equation error implied by calibrated parameters as previous studies do, we compute the error based on the MCMC chain of parameter estimates for each asset pricing model. For each model, we compute several metrics of the error on a chain of estimates (12,000 sets of estimates) obtained from the GSM Bayesian estimation.

For the AAMS model, the Euler equation errors defined on the dividend claim and consumption claim are respectively given by

$$
\begin{aligned}
EulerErr_t^D &= \frac{-C_t + \left(\mathbb{E}_t\left[\beta C_{t+1}^{-1/\psi}\left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{1/\psi-\gamma}\left(\frac{\left(\mathbb{E}_{\{s_{t+1},t\}}[V_{t+1}^{1-\gamma}]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t(V_{t+1})}\right)^{-(\eta-\gamma)}\frac{\frac{P_{t+1}}{D_{t+1}}+1}{\frac{P_t}{D_t}}\frac{D_{t+1}}{D_t}\right]\right)^{-\psi}}{C_t} \\
EulerErr_t^C &= \frac{-C_t + \left(\mathbb{E}_t\left[\beta C_{t+1}^{-1/\psi}\left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{1/\psi-\gamma}\left(\frac{\left(\mathbb{E}_{\{s_{t+1},t\}}[V_{t+1}^{1-\gamma}]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t(V_{t+1})}\right)^{-(\eta-\gamma)}\frac{\frac{W_{t+1}}{C_{t+1}}}{\frac{W_t}{C_t}-1}\frac{C_{t+1}}{C_t}\right]\right)^{-\psi}}{C_t}.
\end{aligned}
$$

The errors are defined in a similar way for other models including AAMSSV, AALRRSV, EZLRRSV and EZMS. The differences are only with regard to the SDF and conditioning state variables. This measure is expressed as a fraction of consumption goods, namely the residual of the Euler equation normalized by consumption. $EulerErr_t^D$ ($EulerErr_t^C$) quantifies the error the agent would commit if he use the approximate solution for the price of the dividend (consumption) claim to decide on marginal investment.

Following Judd (1992), we consider several metrics of the error to evaluate numerical accuracy:

$$
\begin{aligned}
\mathcal{E}_1^D &= \log_{10}\left(\mathbb{E}\left[\left|EulerErr_t^D\right|\right]\right),\ \mathcal{E}_2^D = \log_{10}\left(\mathbb{E}\left[\left(EulerErr_t^D\right)^2\right]\right) \\
\mathcal{E}_1^C &= \log_{10}\left(\mathbb{E}\left[\left|EulerErr_t^C\right|\right]\right),\ \mathcal{E}_2^C = \log_{10}\left(\mathbb{E}\left[\left(EulerErr_t^C\right)^2\right]\right).
\end{aligned}
$$

We report the mean, 5 percentile and 95 percentile of each metric evaluated on the MCMC chains of estimates. It is important to note that we compute the Euler equation error outside the grid that we use to implement the collocation projection method. This is done because we want to assess whether our approximate solutions perform well for simulated data under each model, and because in the GSM Bayesian estimation we use the simulated data to find the mapping recovery from structural parameters to the auxiliary model parameters. We report all measures in $\log_{10}$ terms. For example, a value of $\mathcal{E}_1^D$ equal to -3 suggests that if an agent relies on the approximate solution of the price of the dividend claim, he would expect to make a mistake of 1 dollar for each \$1000 risky investment. The economic interpretation is similar for $\mathcal{E}_1^C$. The metric $\mathcal{E}_2^{D(C)}$ measures the quadratic average of the error. Results reported in Table 1 show that our approximate solutions are accurate.

Table 1: **Numerical accuracy: Euler errors**

Model	$\mathcal{E}_1^D$	$\mathcal{E}_2^D$	$\mathcal{E}_1^C$	$\mathcal{E}_2^C$
AAMS				
Mean	-2.654	-5.281	-3.656	-7.282
95 percentile	-2.179	-4.334	-3.230	-6.437
5 percentile	-3.256	-6.480	-4.233	-8.420
AAMSSV				
Mean	-2.594	-4.940	-4.100	-7.985
95 percentile	-1.826	-3.282	-3.165	-6.120
5 percentile	-3.918	-7.679	-5.612	-11.072
AALRRSV				
Mean	-2.207	-4.083	-4.093	-7.550
95 percentile	-1.633	-2.956	-2.983	-5.551
5 percentile	-2.626	-4.951	-4.932	-9.297
EZLRRSV				
Mean	-2.877	-5.387	-2.820	-5.255
95 percentile	-2.724	-5.066	-2.690	-4.966
5 percentile	-3.126	-5.880	-3.044	-5.708
EZMS				
Mean	-3.751	-7.335	-4.572	-8.987
95 percentile	-3.251	-6.026	-3.979	-7.834
5 percentile	-4.621	-9.077	-5.444	-10.737

Table 2: **Prior: AAMS**

Parameter	Min	Max	μ	σ
β	0.9	0.995	0.985	0.005
γ	0.1	100	5	1
ψ	0.1	10	1.5	0.2
η	γ	200	8.87	2
p_{ll}	0.2	0.999	0.516	0.13
p_{hh}	0.2	0.999	0.978	0.24
μ_l	-0.08	0.00	-0.0678	0.017
μ_h	0.00	0.08	0.022	0.006
λ	1	6	2.74	0.8
σ_c	0.004	0.06	0.03	0.0075
σ_d	0.03	0.3	0.13	0.03

3 Priors on structural parameters

We report support conditions (Min and Max), prior location and scale parameters for structural parameters in models AAMS, AAMSSV, AALRRSV and EZLRRSV.[1] For each model, the prior is the combination of the product of independent normal density functions and support conditions. The product of independent normal density functions is given by

$$\xi(\theta) = \prod_{i=1}^{\tilde{n}} N\left[\theta_i | \left(\theta_i^*, \sigma_\theta^2\right)\right]$$

where $\tilde{n}$ denotes the number of parameters. Because this prior is intersected with support conditions that are not all of product form, and because a support condition that rejects parameter values in the MCMC chain implies extreme parameter values such that the solution method fails, this is not an independence prior.

[1] The EZMS model has the same prior on parameters as the AAMS model does except for the absence of the ambiguity aversion parameter η.

Table 3: **Prior: AAMSSV**

Parameter	Min	Max	μ	σ
β	0.9	0.995	0.985	0.005
γ	0.1	100	5	1
ψ	0.1	10	1.5	0.2
η	γ	200	8.87	2
p_{ll}	0.2	0.999	0.516	0.13
p_{hh}	0.2	0.999	0.978	0.24
μ_l	-0.08	0.00	-0.0678	0.017
μ_h	0.00	0.08	0.022	0.006
p^{σ}_{ll}	0.2	0.999	0.85	0.07
p^{σ}_{hh}	0.2	0.999	0.85	0.07
σ_l	0.004	0.06	0.015	0.0038
σ_h	0.004	0.06	0.03	0.0075
λ	1	6	2.74	0.8
σ_d	0.03	0.3	0.13	0.03

Table 4: **Prior: AALRRSV**

Parameter	Min	Max	μ	σ
β	0.9	0.995	0.985	0.005
γ	0.1	100	5	1
ψ	0.1	10	1.5	0.2
η	γ	200	25	5
μ_c	0.012	0.025	0.02	0.001
ρ_x	-0.99	0.99	0.8	0.2
ϕ_x	0.01	0.5	0.15	0.04
λ	1	10	3	0.8
ϕ_d	0.5	10	3	0.8
μ_s	0.001	0.1	0.02	0.005
ρ_s	0.3	0.99	0.8	0.2
σ_w	1e-5	0.001	2.5e-4	6.25e-5

Table 5: **Prior: EZLRRSV**

Parameter	Min	Max	μ	σ
β	0.9	0.995	0.985	0.005
γ	0.1	100	5	1
ψ	0.1	10	1.5	0.2
μ_c	0.012	0.025	0.019	0.001
ρ_x	-0.99	0.99	0.80	0.20
ϕ_x	0.01	0.5	0.15	0.04
λ	1	10	3	0.8
ϕ_d	0.5	10	3	0.8
ϕ_c	1	10	3	0.8
μ_s	0.001	0.10	0.02	0.005
ρ_s	0.30	0.99	0.8	0.2
σ_w	1e-5	0.001	2.5e-4	6.25e-5

4 GSM estimation results with augmented priors

We also perform the GSM Bayesian estimation with augmented priors taking into account moments of asset returns and consumption and dividend growth. The aim of this estimation is to examine whether our GSM estimation results reported in the paper are robust to the augmented priors. The augmented prior on moments is specified to be the product of independent normal density functions as

$$\bar{\xi}\left(\boldsymbol{m}\right) = \prod_{k=1}^{\bar{n}} N\left[m_k \middle| \left(m_k^*, \sigma_{m_k}^2\right)\right]$$

where $\boldsymbol{m} \equiv (m_1, m_2, ..., m_{\bar{n}})$ is a vector of moments under consideration. The location and scale parameters for the moment m_k are m_k^* and σ_{m_k} respectively. We use the following location parameter values for eight moments to form the prior.

$$\mathbb{E}(r_t^f) = 0.014, \quad \sigma(r_t^f) = 0.028, \quad \mathbb{E}(r_t) = 0.068, \quad \sigma(r_t) = 0.18$$
$$\mathbb{E}(\Delta c_t) = 0.018, \quad \sigma(\Delta c_t) = 0.021, \quad \mathbb{E}(\Delta d_t) = 0.018, \quad \sigma(\Delta d_t) = 0.14$$

The scale parameters are set at values such that the prior put 95% of its mass on being within 10% of its location parameter. We simulate these moments from asset pricing models in the GSM Bayesian estimation. The results reported below show that the GSM estimation with the augmented priors yield similar results to those reported in the paper.

Table 6: **GSM Estimation Results: the AAMS Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.991	0.991	0.989	0.991	0.980	0.980	0.977	0.983
γ	4.847	4.750	2.750	6.750	2.219	2.281	0.766	3.531
ψ	0.616	0.563	0.563	0.813	2.202	2.180	1.836	2.680
η	13.155	13.500	11.500	13.500	5.557	5.219	4.594	7.922
p_{ll}	0.405	0.406	0.406	0.406	0.860	0.866	0.764	0.923
p_{hh}	0.812	0.813	0.813	0.813	0.997	0.997	0.996	0.997
μ_l	-0.043	-0.043	-0.043	-0.043	-0.048	-0.048	-0.054	-0.041
μ_h	0.033	0.033	0.033	0.033	0.020	0.020	0.018	0.021
λ	2.803	2.750	2.750	3.250	2.791	2.734	2.391	3.547
σ_c	0.006	0.006	0.006	0.006	0.020	0.020	0.018	0.024
σ_d	0.130	0.133	0.117	0.133	0.135	0.136	0.124	0.146
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AAMS model. The GSM estimation imposes the augmented prior on moments of asset returns and consumption and dividend growth. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for annual data 1941–2015.

Table 7: **GSM Estimation Results: the AAMSSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.989	0.989	0.989	0.989	0.978	0.978	0.974	0.982
γ	4.982	5.250	3.250	6.250	0.848	0.844	0.219	1.531
ψ	0.450	0.438	0.438	0.438	1.779	1.715	1.434	2.199
η	9.496	9.500	9.500	9.500	10.232	9.281	8.531	15.500
p^{μ}_{ll}	0.282	0.281	0.281	0.281	0.706	0.728	0.611	0.774
p^{μ}_{hh}	0.812	0.813	0.813	0.813	0.998	0.999	0.997	0.999
μ_l	-0.066	-0.066	-0.066	-0.066	-0.055	-0.054	-0.060	-0.050
μ_h	0.033	0.033	0.033	0.033	0.018	0.018	0.016	0.019
p^{σ}_{ll}	0.863	0.859	0.734	0.984	0.989	0.989	0.982	0.993
p^{σ}_{hh}	0.840	0.859	0.703	0.953	0.989	0.990	0.984	0.993
σ_l	0.006	0.005	0.005	0.009	0.013	0.013	0.006	0.021
σ_h	0.006	0.006	0.006	0.010	0.029	0.029	0.026	0.032
λ	3.064	3.250	2.750	3.250	2.570	2.547	1.984	3.172
σ_d	0.107	0.102	0.086	0.117	0.134	0.134	0.122	0.146
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AAMSSV model. The GSM estimation imposes the augmented prior on moments of asset returns and consumption and dividend growth. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for annual data 1941–2015.

Table 8: **GSM Estimation Results: the AALRRSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.989	0.989	0.985	0.993	0.992	0.992	0.990	0.995
γ	8.967	8.875	7.125	10.875	8.013	8.031	7.156	8.906
ψ	1.155	1.156	0.906	1.406	0.899	0.895	0.855	0.957
η	33.125	33.000	25.000	41.000	31.085	31.500	19.500	41.500
μ_c	0.019	0.019	0.018	0.020	0.020	0.020	0.019	0.022
ρ_x	0.844	0.844	0.844	0.844	0.940	0.940	0.927	0.953
ϕ_x	0.311	0.305	0.273	0.367	0.209	0.210	0.181	0.233
λ	3.239	3.125	2.875	3.875	2.784	2.734	2.453	3.203
ϕ_d	4.900	4.875	4.375	5.375	4.995	4.984	4.609	5.359
μ_s	0.011	0.010	0.010	0.011	0.020	0.020	0.019	0.020
ρ_s	0.969	0.969	0.969	0.969	0.950	0.950	0.950	0.950
σ_w	1.84E-04	1.68E-04	1.68E-04	1.98E-04	2.08E-04	2.09E-04	1.93E-04	2.18E-04
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the AALRRSV model. The GSM estimation imposes the augmented prior on moments of asset returns and consumption and dividend growth. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for annual data 1941–2015.

Table 9: **GSM Estimation Results: the EZMS Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.986	0.985	0.985	0.987	0.981	0.981	0.978	0.984
γ	10.221	10.250	10.250	10.250	4.013	3.953	3.469	4.672
ψ	0.421	0.438	0.313	0.438	2.397	2.359	1.859	2.984
p_{ll}	0.343	0.344	0.344	0.344	0.904	0.897	0.860	0.951
p_{hh}	0.812	0.813	0.813	0.813	0.996	0.997	0.992	0.997
μ_l	-0.059	-0.059	-0.059	-0.059	-0.039	-0.041	-0.052	-0.018
μ_h	0.029	0.029	0.029	0.029	0.022	0.021	0.018	0.029
λ	3.252	3.250	3.250	3.250	3.192	3.172	2.641	3.766
σ_c	0.006	0.006	0.006	0.006	0.019	0.019	0.016	0.023
σ_d	0.105	0.102	0.102	0.117	0.132	0.133	0.117	0.145
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the EZMS model. The GSM estimation imposes the augmented prior on moments of asset returns and consumption and dividend growth. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for annual data 1941–2015.

Table 10: **GSM Estimation Results: the EZLRRSV Model**

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.993	0.993	0.989	0.993	0.984	0.984	0.981	0.986
γ	6.800	6.750	5.250	7.750	8.542	8.438	6.938	10.438
ψ	1.375	1.313	1.313	1.688	2.773	2.789	2.531	3.023
μ_c	0.020	0.020	0.019	0.021	0.022	0.023	0.021	0.024
ρ_x	0.937	0.938	0.938	0.938	0.985	0.989	0.971	0.990
ϕ_x	0.290	0.305	0.195	0.336	0.077	0.071	0.060	0.108
λ	2.582	2.250	2.250	3.750	4.518	4.469	3.688	5.406
ϕ_d	4.742	4.750	3.750	5.750	4.546	4.531	4.156	4.938
ϕ_c	4.002	4.250	2.750	5.250	1.217	1.156	1.016	1.656
μ_s	0.013	0.011	0.011	0.017	0.019	0.019	0.019	0.020
ρ_s	0.938	0.938	0.938	0.938	0.950	0.950	0.950	0.950
σ_w	1.50E-04	1.37E-04	1.37E-04	1.68E-04	1.93E-04	1.93E-04	1.87E-04	2.03E-04
MCMC repetitions		10,000				12,000		

This table presents prior and posterior marginal means, medians, 5 and 95 percentiles of model parameters for the EZLRRSV model. The GSM estimation imposes the augmented prior on moments of asset returns and consumption and dividend growth. MCMC repetitions after transients have dissipated are reported for both the prior and posterior. Estimation results are for annual data 1941–2015.

References

Caldara, D., Fernandez-Villaverde, J., Rubio-Ramirez, J. F., Yao, W., 2012. Computing dsge models with recursive preferences and stochastic volatility. Review of Economic Dynamics 15 (2), 188–206.

Collard, F., Mukerji, S., Sheppard, K., Tallon, J.-M., 2017. Ambiguity and the historical equity premium. Working Paper.

Epstein, L. G., Zin, S. E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. Econometrica 57 (4), 937–969.

Guerrieri, L., Iacoviello, M., 2015. Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. Journal of Monetary Economics 70, 22–38.

Ju, N., Miao, J., 2012. Ambiguity, learning, and asset returns. Econometrica 80, 559–591.

Judd, K. L., 1992. Projection methods for solving aggregate growth models. Journal of Economic Theory 58, 410–452.

Judd, K. L., 1999. Numerical Methods in Economics, 1st Edition. The MIT Press, Cambridge, MA.

Pohl, W., Schmedders, K., Wilms, O., 2017. Higher-order effects in asset pricing models with long-run risks. Journal of Finance forthcoming.

Tauchen, G., 1986. Finite state Markov-chain approximations to univariate and vector autoregressions. Economics Letters 20 (2), 177–181.