Does Smooth Ambiguity Matter for Asset Pricing?

DOI:
10.1093/rfs/hhy118

Document Version
Accepted author manuscript

Citation for published version (APA):
Liu, H., Gallant, A. R., & Jahan-Parvar, M. (2019). Does Smooth Ambiguity Matter for Asset Pricing? *Review of Financial Studies*, *32*, 3617–3666. https://doi.org/10.1093/rfs/hhy118

Published in:
Review of Financial Studies

Does Smooth Ambiguity Matter for Asset Pricing?

A. Ronald Gallant
Pennsylvania State University

Mohammad R. Jahan-Parvar
Federal Reserve Board

Hening Liu
University of Manchester

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 C61; D81; G11; G12.)

We thank Stijn Van Nieuwerburgh (the editor) and two anonymous referees for many helpful comments that substantially improved the paper. We also thank Yoosoon Chang, 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, Paul Schneider, Dongho Song, Michael Stutzer, Fabio Trojani, Toni Whited, seminar participants at the Federal Reserve Board, Cal Poly at San Luis Obispo, George Washington University, Georgetown University, Indiana University, North Carolina State University, University of Colorado, University of Maryland, Midwest Econometric Group 2014, CFE 2014, SoFiE 2015, China International Conference of Finance 2015 and 2018, FMA 2015, Econometric Society World Congress 2015, Winter Meeting of the Econometric Society 2016, SFS Cavalcade 2016, Zurich Workshop on Asset Pricing 2018, WFA 2018, and EFA 2018 for helpful comments and discussions. This paper was previously circulated under the title "Measuring Ambiguity Aversion." 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. Send correspondence to Mohammad R. Jahan-Parvar, Board of Governors of the Federal Reserve System, 20th St. NW and Constitution Ave., Washington, DC 20551. E-mail: Mohammad.Jahan-Parvar@frb.gov

A number of asset pricing puzzles pose remarkable challenges to the standard consumption-based model with a fully rational representative agent. Among them 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 volatility of returns, 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 a 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 Gallant and McCulloch's (2009) general scientific models (GSM), a simulation-based Bayesian estimation method, to estimate a set of three consumption-based asset pricing models with smooth ambiguity preferences and three models with recursive utility. Due to nonlinearities inherent in structural asset pricing models, the likelihood of these models is generally not available in closed form. In addition, similar to other nonlinear applications, we face sparsity of data. As has become standard in the macrofinance empirical literature, we use Bayesian methods (in our case, GSM) coupled with prior information to overcome data sparsity. We describe this estimation method in Section 2.2. We are interested in estimates of preferences parameters and parameters governing state dynamics jointly obtained from fitting structural models to data. We consider models with an ambiguity-averse representative agent who is uncertain about the con-

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

[2] See Epstein and Wang (1994), Z. Chen and Epstein (2002), Epstein and Schneider (2008), and Drechsler (2013) for applications of the multiple-priors preferences and Ju and Miao (2012), Collard et al. (2018), and Jahan-Parvar and Liu (2014) 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).

ditional 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). Our estimation suggests a clear distinction of the EIS from risk aversion for all models and statistical evidence supporting 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 et al. (2014), among others. By estimating a model with ambiguity, learning, and time-varying volatility, we aim to investigate whether (i) inclusion of time-varying volatility affects the estimated impact of ambiguity on asset prices, and (ii) the model with time-varying volatility represents an improvement over the original model. In addition, our estimation results allow us to disentangle the contribution of time-varying volatility and ambiguity aversion to improvements in fitting the data.

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. (2018). The motivation for examining this model is to study the impact of ambiguity in a long-run risk setting. This model features consumption growth dynamics similar to the model studied by Bansal, Kiku, and Yaron (2012). It features a long-run risk component and stochastic volatility in conditional mean and volatility of the consumption growth process, respectively. 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 stochastic discount factor (SDF) when the economy experiences a negative shock after staying at the "normal" growth rate for several periods. The pessimistic distortion to the SDF 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 three baseline models with Epstein and Zin's (1989) 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 long-run risk model allows for heteroscedas-

[3] Bidder and Dew-Becker (2016) study a related framework in which the agent estimates the consumption process nonparametrically and prices assets using a pessimistic model. They find that long-run risks arise endogenously as the worst-case outcome. Collard et al. (2018) 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.

ticity consumption growth and generates time-varying volatility in returns. The second baseline model is a special case of Ju and Miao's model that suppresses ambiguity aversion. In this model, the agent without endogenous pessimism uses Bayesian inference to evaluate mean consumption growth. The third model is an extension of the second and incorporates time-varying conditional volatility. By estimating a series of structural models with and without smooth ambiguity, we address two important questions: (i) Does a structural estimation with macrofinance data lend statistical support to the class of smooth ambiguity preferences that have a sound decision-theoretic basis? (ii) Based on a standard Bayesian model comparison between those featuring smooth ambiguity and Epstein-Zin's preferences, which estimated model is preferred?

We find a significant distinction between risk aversion and ambiguity aversion in the estimated models. This distinction is robust to different specifications of consumption dynamics. Incorporating smooth ambiguity aversion in consumption-based asset pricing models greatly reduces the burden on risk aversion, and yields posterior distributions of the risk-aversion parameter that are centered on values between 0.8 and 5. These values are noticeably smaller than median and mean values of the posterior distributions of the ambiguity aversion parameter, which range between 6.9 and 23.5. We find that incorporating smooth ambiguity in models lowers the estimated values of the EIS parameter. However, our estimation indicates that the EIS parameter is greater than 1 and thus lends support to the preference for early resolution of uncertainty. 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 Epstein-Zin's recursive utility model with regime-switching mean consumption growth, learning, and time-varying volatility. 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.[5] In addition, we find that the estimated smooth ambiguity models can match moments of asset returns better than the Epstein-Zin models do. The estimated Ju and Miao model, while receiving less statistical support than the long-run risk model, can match the equity premium and variance risk premium in the data well.

[5] In a Bayesian estimation study, A. Chen, Wasyk, and Winkler (2018) estimate asset pricing models with multiple risks including long-run growth, long-run volatility, habit, and a residual and find that the residual is important to explain variations in the price-dividend ratio.

We use the projection method to solve all models examined in this paper. The widely used log-linear approximation method in the long-run risk literature is not applicable to smooth ambiguity models. This is because learning is an important ingredient in smooth ambiguity models and 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 (2018) 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 (2018) 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.

This paper contributes to a growing body of literature on ambiguity, learning, and macrofinance. 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 Z. Chen 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. (2018) 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. Ai and Bansal (2018) show that a wide class of non-expected utility models, including the smooth ambiguity model, implies demands for risk compensation for news affecting continuation values and generates a premium related to macroeconomic announcements.

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. Thimme and Völkert (2015) use generalized method of moments (GMM) of Hansen (1982) to estimate a reduced-form, locally linearized version of the smooth ambiguity model without characterizing dynamics of beliefs. Their estimation crucially relies on an approximation of conditional expectation of future utility, a proxy for the wealth-consumption ratio, and omission of any specification for consumption dynamics. 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 (2018) 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 a 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.

1 Asset Pricing Models

1.1 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 (i) Ju and Miao's (2012) model in which the mean of consumption growth follows a hidden Markov chain with two states, abbreviated as AAMS, (ii) an extended version of Ju and Miao's model with time-varying conditional volatility, abbreviated as AAMSTV, and (iii) a long-run risk model featuring ambiguity in which the long-run risk component is assumed to be unobservable, abbreviated as AALRRSV. The latter model shares many features with the models introduced by Collard et al. (2018). 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.

1.1.1 The AAMS model

Aggregate consumption follows the process

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

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

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

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 is usually interpreted as the leverage parameter; see Abel (1999). We pin down the parameters g_d and $\tilde{\sigma}_d$ by the estimates of the 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 with consumption. As a result, given the prior belief π_0, 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 the conditional Gaussian 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

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

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

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

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 in Equation (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 other information at time t. The expectation operator $\mathbb{E}_{\pi_t}$ is taken with respect to the posterior belief about the unobservable state.

[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).

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-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

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

We rewrite the Euler equation as

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

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 $M_{t,t+1}^{EZ}$ as the SDF under recursive utility:

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

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

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

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 3.

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, $\frac{P_t}{D_t} = \Phi(\pi_t)$, the Euler equation becomes

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

1.1.2 The AAMSTV model

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

$$\Delta c_t = \mu(s_t^\mu) + \sigma(s_t^\sigma)\epsilon_{c,t}, \quad \epsilon_{c,t} \sim N(0,1)$$

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 (thus, no ambiguity about the volatility state). We make this simplifying assumption for three reasons. First, empirical studies such as Bryzgalova and Julliard (2015) have established that estimation and characterization of mean consumption growth is more difficult than consumption volatility. Second, according to the existing literature, the volatility state is very persistent, leading to filtered probabilities of the volatility state close to 1. These results suggest that ambiguity has limited room with respect to the consumption volatility.[7] Third, while Epstein and Ji (2013) and

[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 Markov chain Monte Carlo estimation infeasible. Lettau, Ludvigson, and Wachter (2008) also point out the convergence issue

Branger, Schlag, and Thimme (2016) argue that ambiguity about volatility may have certain asset pricing implications, Veronesi (1999), Ju and Miao (2012), and Miao, Wei, and Zhou (Forthcoming) show that learning about the conditional mean of fundamentals is sufficient to characterize salient features of macrofinancial variables. We confirm their findings based on structural estimation.

The agent updates beliefs according to Bayes' rule as

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

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

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

The value function is given by

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

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

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

where

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

for Epstein and Zin's recursive utility.

The SDF in this model is

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

The dividend growth process is specified in the same form as in the AAMS model, that is, in Equation (1). Stock returns and the risk-free rate are defined as usual. The price-dividend ratio ($\frac{P_t}{D_t} = \Phi\left(\pi_t, s^{\sigma}_t\right)$) satisfies the Euler equation

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

1.1.3 The AALRRSV model

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

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

In Bansal and Yaron's calibration, x_t is a highly persistent component, and σ_t is the highly persistent stochastic volatility component representing time-varying economic uncertainty. 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, Gallant, and Tauchen (2007), Ma (2013), and Johannes, Lochstoer, and Mou (2016), among others. The difficulty in estimating x_t gives rise to the agent's ambiguity about mean consumption growth. As a result, we adopt a more plausible information structure by assuming that x_t is unobservable. Collard et al. (2018) 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, that is, 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}, that is, $x_{t+1} \sim N(\hat{x}_{t+1}, \nu_{t+1})$. 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 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 Equation (3) that the distorted density driven by ambiguity, $\tilde{f}\left(x_t | \hat{x}_t, \nu_t, t\right)$, is given by

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

1.2 Alternative Models Featuring Ambiguity-Neutral Preferences

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

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

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

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

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 the EZMS model as the ambiguity-neutral version of the AAMS model.

The second alternative model, which we call EZMSTV, is the ambiguity-neutral version of AAMSTV where we suppress ambiguity aversion by setting $\eta = \gamma$, as in the derivation of EZMS. EZMSTV has the same consumption growth dynamics as the model studied by Lettau, Ludvigson, and Wachter (2008) and features Epstein-Zin preferences.

The third alternative model is the long-run risk model of Bansal, Kiku, and Yaron (2012), which we label as EZLRRSV. The model specification is

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

with notations defined in the same way as in the AALRRSV model. 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 structural parameters to be estimated for each model in Table 2. In estimating AALRRSV and EZLRRSV models, 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, Schmedders, and Wilms (2018) 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 Online Appendix.

2 Data and the Estimation Method

2.1 Data

Throughout this paper, lower case denotes the natural logarithm of an upper case variable; for example, $c_t = \ln(C_t)$, 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 ($r_t = \ln\left(P_t + D_t\right) - \ln P_{t-1}$) 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 the 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 consumer price index (CPI) data. We use the sum of real nondurable and services consumption, items 16 and 17 in national income and product accounts (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 the 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 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.

2.2 GSM: Estimation of the structural model

We use the Bayesian method proposed by Gallant and McCulloch (2009), which they termed general scientific models (GSM) to estimate the asset pricing models. The GSM methodology was refined in Aldrich and Gallant (2011), abbreviated AG hereafter.[8] The discussion here incorpo-

[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.

rates those refinements and is to a considerable extent a paraphrase of AG. GSM is a Bayesian simulation estimator. It is useful when a computationally tractable likelihood function is not available, data are sparse, but the structural model can be solved and simulated. It shares certain similarities with 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 GSM method 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 AG with Bansal, Gallant, and Tauchen (2007) displays the advantages of a Bayesian simulation approach relative to a frequentist EMM approach, particularly for the purpose of model comparison. GSM 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 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.[9] 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. Implementing GSM 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.

[9] 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.

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 six structural models under consideration in this application: the three models featuring smooth ambiguity and the three alternative models with Epstein-Zin's recursive utility, all of which are Markovian and described in Section 1.

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."[10] 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

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

[10] 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.

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 2.5 that are sufficient for us to implement Bayesian methods by means of Markov chain Monte Carlo (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 for the θ chain) is correct when f is a nonlinear model. The algorithm proposed by AG to address this difficulty is described next.

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

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

with respect to ω. The advantage of the 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 $N = 1,000$ in the estimation of all the six 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 of the random number generator 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 later 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 proposed by GM, 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 Expression (7) using the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) algorithm; see Fletcher (1987). This is also 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 2.5.

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 for the θ-chain 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 8 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 Equation (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 optimally computed from many different random starts after the chain has run awhile.

2.3 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 1.

The data are modeled as

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

where

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

which is the location function of a k-lag vector autoregressive (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).[11] 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.

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

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 Hellinger norm and can therefore approximate a density to within arbitrary accuracy in the Kullback-Leibler distance; see Gallant and Nychka (1987). Such approximations are often called sieves; Gallant and Nychka term this particular sieve semi-nonparametric or SNP.[12] 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 in Equation (8) and the variance function in Equation (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 constitute a sieve; see Schwarz (1978).

In our specification, U_0 is an upper triangular matrix, P and $\tilde{V}$ are diagonal matrices, and Q is scalar. The degree of the SNP $h(\varepsilon)$ density is four. The auxiliary model chosen for our analysis, based on the BIC, has one lag in the conditional mean component, one 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 1. We numerically solve the structural models assuming normally distributed innovation shocks to consumption and dividend growth rates, as discussed in the Online Appendix. 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.

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

2.4 Relative model comparison

Relative model comparison is standard Bayesian inference. The posterior probabilities of the six 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 deleting draws with exceedingly small computed likelihood, 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 in the next section we used the BIC for model selection.

2.5 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 1. 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 Online Appendix. Our estimation results do not crucially rely on the choice of values of location and scale parameters.

We set the location parameter values such that the asset pricing models generate a mean risk-free rate that is not too high and a 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 2, 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 with the prior on this parameter in Schorfheide, Song, and Yaron (2018). 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 the possibility of $\psi < 1$ and thus 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) and Jahan-Parvar and Liu (2014). 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 the EZMS, EZMSTV, AAMS, and AAMSTV models, 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 AAMSTV 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, Ludvigson, and Wachter (2008). 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, Gallant, and Tauchen (2007), Aldrich and Gallant (2011), and Schorfheide, Song, and Yaron (2018) lead to values of λ in the $[1.5, 4.5]$ range. We choose $1 \leq \lambda \leq 6$ as the support interval.

For the AALRRSV and EZLRRSV models, we use the calibrated parameter values in Bansal, Kiku, and Yaron (2012) and priors postulated in Schorfheide, Song, and Yaron (2018) 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, Song, and Yaron (2018).

2.6 Estimation results

We summarize estimation results in Tables 3 to 8.[13] We plot the prior and posterior densities of the estimated structural parameters in Figures 3 to 8. These figures show considerable shifts in both location and scale between priors and posteriors, suggesting that the estimation procedure and data have a significant impact on estimation results. The impact of 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 macrofinance 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. Our estimation strongly suggests an EIS greater than 1 and thus a preference for early resolution of uncertainty. As shown in Tables 3 to 8, the posterior mean, median, and 90% credible intervals of ψ estimates are uniformly above 1 for all models presented in Section 1. The plots of the posterior density for ψ in Figures 3 to 8 also reveal that the posterior dispersion of this parameter over the MCMC chain is small. Jeong, Kim, and Park (2015) estimate the recursive multiple-prior utility model using asset prices data and obtain implausibly high estimates of ψ that are greater than 10. High estimates of ψ generated from our

[13] 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 8 and Figures 3 to 8.

estimation imply low and stable risk-free rates (see Section 3). In a DSGE analysis with a broader scope, Bianchi, Ilut, and Schneider (2018) 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 the AAMS and EZMS models are high and comparable to the estimates in the long-run risk literature. The posterior mean, median, 5th, and 95th 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, Song, and Yaron (2018) 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 EZMSTV and EZLRRSV models and those in the AAMSTV and AALRRSV models. The posterior estimates of ψ are moderately lower in the AAMSTV model, and significantly lower in the AALRRSV than in the EZLRRSV model. Nevertheless, our estimates of ψ in the long-run risk model are still lower than those reported by Schorfheide, Song, and Yaron (2018). The discrepancy arises because (i) we use the projection method rather than log-linear approximation, (ii) we use the GSM Bayesian method for model estimation, and (iii) we use a different data sample.

Our estimation results strongly support asset pricing models with smooth ambiguity preferences. The posterior estimates of the ambiguity aversion parameter η are significantly large in the AAMS, AAMSTV, and AALRRSV models. Not surprisingly, the estimates obtained for the AAMS model are close to the calibrated value in Ju and Miao (2012) ($\eta = 8.864$). 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 AAMSTV model but only about 7 in the AAMS model. It is important to note that the difference of η estimates in the two models does not stem from a moments-matching exercise. The GSM Bayesian estimation delivers estimates of parameters in the utility function and consumption growth process jointly. In comparison with results for the AAMS model, the moderately higher estimates of η in the AAMSTV model are likely due to lower estimates of the transition probability p^{μ}_{ll}, which implies less persistence of the contraction regime for this model.[14] In the long-run risk setting, the

[14] In a comparative statics exercise, we find that incorporating exogenous time-varying volatility in the AAMS model while keeping original parameters of the model constant can raise equity premium significantly.

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 for the EZMS, EZMSTV, and EZLRRSV models show that the posterior mean and median of γ are high and the 5th and 95th percentiles imply fairly 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, Song, and Yaron (2018) and Bansal, Kiku, and Yaron (2016). The posterior mean of γ is 8.4, and the associated 95th percentile is 10.4. These values are also close to the calibrated values in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012). On the other hand, the γ estimate is more dispersed in models with smooth ambiguity, that is, the AAMS, AAMSTV, and AALRRSV models, as is evident from slightly wider 90% credible intervals. In a related work, X. 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. The AAMS, AAMSTV, EZMS, and EZMSTV models 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 the AAMSTV and EZMSTV models) 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 the AAMSTV and EZMSTV models).

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 correspond to the agent's subjective belief. Compared with estimates of the parameters of the Markov-switching model reported by calibration studies (for example, Cecchetti, Lam, and Mark (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 ($1 - p_{hh}$) implied by our estimation. 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 low given the postwar 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, additional sources of ambiguity about parameters of the Markov-switching model may coexist.

In estimating the AAMSTV and EZMSTV models, we find two distinct volatility regimes, both of which are persistent. This result is consistent with the findings of Lettau, Ludvigson, and Wachter (2008). However, the posterior estimates of the high volatility regime σ_h are somewhat high compared with the postwar consumption data. The estimates of μ_l are more negative than the estimates for the AAMS model. Nevertheless, these estimates are consistent with the long sample of Shiller's data.[15] Again, additional sources of ambiguity may arise due to learning from the past experience or parameter uncertainty.[16] For the AAMS, AAMSTV, and EZMSTV models, 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 1.5 and 4, with a posterior mean of about 3 for the AAMS and AAMSTV models, and around 2 for the EZMSTV model. The estimates of λ and

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

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

σ_d for the EZMS model are higher than those for the AAMS, AAMSTV, and EZMSTV models.

Turning to estimation results of models featuring long-run risks, we find that the estimated AALRRSV and EZLRRSV models both provide support for 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 90% credible intervals. Converted into estimates at a monthly frequency, our results are similar to those reported by Schorfheide, Song, and Yaron (2018). In addition, the stochastic volatility component is persistent in our estimation, a result consistent with Schorfheide, Song, and Yaron (2018).[17] 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, Kiku, and Yaron (2012, 2016) and Schorfheide, Song, and Yaron (2018).

We present results of relative model comparison in Tables 3 to 8, based on the maximum of the log likelihood and the Bayesian information criteria (BIC) of Schwarz (1978) for all estimated models. We use the auxiliary model presented in Section 2.3 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 six estimated models the AAMSTV model 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 AAMSTV$\succ$AALRRSV$\succ$EZMSTV$\succ$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 and EZMSTV models are close to the AAMSTV model, but the remainder are more than 60 orders of magnitude distant. These findings suggest that: (i) time-varying volatility in consumption growth 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 the AAMSTV, AALRRSV, EZMSTV, and EZLRRSV models, all of which feature time-varying volatility, and (ii) the AAMSTV and AALRRSV models with ambiguity, learning, and time-varying volatility are preferred to the long-run risk EZLRRSV model in the statistical model comparison. Although the model of Ju and Miao (2012), AAMS, receives less

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

statistical support than other models do, it can match key financial moments, as shown in the next section.

Rankings of estimated models, either based on log likelihood values or BIC, imply that statistical support for models is highly contingent on the following factors:

1. Ambiguity aversion: two models, AAMSTV and AALRRSV, feature both ambiguity aversion and time-varying volatility. Both occupy the top two spots regardless of the ranking criteria used.

2. Time-varying volatility: in both rankings, but especially in the ranking based on log likelihood values, models featuring time-varying volatility rank higher than those with time-invariant volatility. From a statistical standpoint, the AAMS model does poorly based on both criteria.

3. Markov-switching mean growth: in comparison with models featuring long-run risk, models with Markov-switching mean consumption growth are preferred. AAMSTV is preferred to AALRRSV by both criteria, EZMSTV is preferred to EZLRRSV by both criteria, and EZMS is preferred to EZLRRSV based on the BIC.

Taken together, we can conclude that the gains in fit stem from ambiguity aversion and Markov-switching dynamics, predominantly for the volatility process but also present for the conditional mean.

Formally, as discussed in Subsection 2.5 in Aldrich and Gallant (2011), if one compares two structural models, (p_1, ζ_1) and (p_2, ζ_2), and find that the fit of (p_2, ζ_2) is preferred, then, to gain insight as to why the data imply a preference, one can examine the posterior modes of moments of the two fits using t-statistics of the form $t = (m_1 - m_2)/\sigma_2$ where m_1 and m_2 are the posterior modes of a moment of interest for the two fits and σ_2 is the standard deviation of the posterior distribution of the moment under (p_2, ζ_2). Table 9 summarizes the t-statistics for moments of the preferred AAMSTV against the EZMSTV model and those for the comparison of the AAMSTV against the AAMS model. We denote the AAMSTV model by "2." The moments of interest are volatilities of consumption growth and dividend growth, and means and volatilities of risk-free rate and equity returns.

We first consider the EZMSTV against the AAMSTV model comparison, focusing on t-statistics greater than 2. Both models allow time-varying volatility, but the EZMSTV model lacks ambiguity aversion. The evidence suggests that ambiguity aversion generates a more credible SDF, as evidenced by the much better fit to the risk-free rate for the AAMSTV model. Because the two models imply similar variations in the dividend growth rate, we can infer for the EZMSTV model that a poor fit to the SDF is the cause of the poor match to moments of equity returns. Next, we consider the AAMS model against the AAMSTV model comparison. Both models allow ambiguity aversion, but the AAMS model lacks time-varying volatility. The removal of time-varying volatility leads to significant changes to the level of the risk-free rate and the volatility of dividend growth. Thus, we conclude that the better fit of the AAMSTV model to asset pricing moments and the dividend growth volatility delivers the gains in empirical performance rather than a better fit to consumption growth moments.

3 Asset Pricing Implications

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

3.2 Impulse-response functions

We perform impulse-response 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 8 to parameterize models and compute impulse-responses functions. Results for the AAMS, AAMSTV, AALRRSV and EZLRRSV models are plotted in Figures 9 and 10. For the long-run risk EZLRRSV and AALRRSV models, we assume that the shock to the long-run risk component occurs in the third period. For the AAMS and AAMSTV models featuring Markov-switching consumption growth, we assume that the economy stays in the good regime for a long time in the absence of innovation shocks. In the third period, the growth rate of consumption switches to the mean growth rate of the bad regime. The growth rate of consumption then follows a Markov-switching model after the regime shift. We simulate consumption growth rates from the two Markov-switching models respectively, taking into account persistence of regimes, and obtain simulated beliefs accordingly. We compute responses of financial variables to changes in simulated states and plot mean responses across simulations in Figure 9.

Figure 9 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 9 displays qualitatively similar impulse-responses of beliefs and financial variables for the AAMSTV model.[18] The notable discrepancies in the magnitude of responses between the AAMS model and the AAMSTV model are largely due to the inclusion of time-varying volatility in the AAMSTV model and different parameter estimates, as discussed in Section 2.6.

Figure 10 displays the responses of key variables in the AALRRSV and EZLRRSV models 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 an 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

[18] The impulse-response function plots for the EZMS and EZMSTV models are similar and thus omitted here for the sake of brevity.

VRP falls at first and rises afterwards, due to the response of the conditional equity volatility. Figure 10 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 with the results for the AAMS and AAMSTV models.

3.3 Financial moments

We investigate the ability of all estimated models in replicating unconditional moments of key 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 models. We examine how well our estimated models can match moments of asset returns, given that our estimation 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 moments of asset returns, we view this outcome as confirmation that the model characterizes the underlying data-generating process of returns 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, Gallant, and Tauchen (2007), Aldrich and Gallant (2011), and Jeong, Kim, and Park (2015).

Table 10 presents unconditional moments of asset returns simulated from all asset pricing models considered in this paper. For each model, we run a chain of 300,000 Monte Carlo draws for each case. We compute the reported modes from these 300,000 draws. We compute posterior means, medians, and credible intervals with a stride of 25.[19] We report mean, median, standard deviation, and 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 the 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 10 shows that among all models, the AAMS model can best match moments of returns.

[19] A stride of 25 translates into 12,000 draws: 300,000/25 = 12,000.

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, Mukerji, Sheppard, and Tallon (2018). 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 comovement of the SDF and the return volatility when the economy shifts to a bad state. The comovement 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 10 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 AAMS model. 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 for both 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.

However, the relative success of the AAMSTV model is due to the presence of ambiguity aversion, not because of incorporating time-varying consumption volatility. It is evident in Table 10 that the EZMSTV model (ambiguity-neutral case) demonstrates a significantly poorer performance than both AAMS and AAMSTV. In particular, it fails to produce a high enough market price of risk or generate a high and volatile VRP.

In the long-run risk setting, the equity premium and market price of risk implied by the AAL-RRSV 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 the AAMS and AAMSTV models, which 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 AAMSTV model is an extension of AAMS and a statistically preferred model as suggested by the model comparison, we also investigate conditional financial moments in AAMSTV. For a model of our interest, we compute conditional moments given each of the 12,000 parameter draws from the MCMC estimates of structural parameters of the model. We then obtain summary statistics for the 12,000 sets of the conditional moments computed above, including the 5th and 95th percentiles. These values of conditional moments and summary statistics characterize the posterior of the conditional moments for the model. Figure 11 shows simulated conditional equity premium, return volatility, market price of risk, and the VRP plotted against the state variable π_t in the AAMS 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's (2012) use a long sample for their calibration. Figure 12 plots simulated conditional moments for the AAMSTV model, 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 because 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.

4 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. We employ the semi-nonparametric method to estimate the auxiliary model and the GSM Bayesian method to obtain posterior estimates of structural parameters. Our structural estimation with macrofinance data provides statistical support for models with smooth ambiguity preferences. Based on our estimation results, the quantitative effects of smooth ambiguity on asset returns are significant, in both the Markov-switching and long-run risk environments.

Our main findings are: (i) the distinction between risk aversion and ambiguity aversion is remarkable in models with smooth ambiguity, and the statistical support for smooth ambiguity is robust to specifications of consumption and dividend processes; the posterior distributions of the risk-aversion parameter are centered on values between 0.8 and 5, while the median and mean of the posterior distributions of the ambiguity aversion parameter range between 6.9 and 23.5 across models, (ii) the median and mean of the EIS parameter are greater than 1 in estimated models with and without ambiguity, and this result lends support to preferences for early resolution of uncertainty, (iii) 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, and (iv) models with ambiguity, learning, and time-varying volatility are preferred to the long-run risk model according to likelihood values and the BIC; the gains in fitting are mainly driven by better performance in explaining the dynamics of asset prices.

References

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

Ai, H., and R. Bansal. 2018. Risk preferences and the macroeconomic announcement premium. *Econometrica* 86:1383–430.

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

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

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

Bansal, R., R. Gallant, and G. Tauchen. 2007. Rational pessimism and exuberance. *Review of Economic Studies* 74:1005–33.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2014. Volatility, the macroeconomy, and asset prices. *Journal of Finance* 69:2471–511.

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

———. 2016. Risks for the long run: Estimation with time aggregation. *Journal of Monetary Economics* 82:52–69.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–509.

Bianchi, F., C. Ilut, and M. Schneider. 2018. Uncertainty shocks, asset supply and pricing over the business cycle. *Review of Economic Studies* 85:810–54.

Bidder, R., and I. Dew-Becker. 2016. Long-run risk is the worst-case scenario. *American Economic Review* 106:2494–527.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31:307–27.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. Expected stock returns and variance risk premia. *Review of Financial Studies* 22:4463–92.

Branger, N., C. Schlag, and J. Thimme. 2016. Does ambiguity about volatility matter empirically? *Working Paper, University of Muenster and Goethe University* .

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

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

Chen, A. Y., R. Wasyk, and F. Winkler. 2018. A likelihood-based comparison of macro asset pricing models. *FEDS Working Paper, Federal Reserve Board* .

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

Chen, Z., and L. Epstein. 2002. Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70:1403–43.

Cogley, T., and T. J. Sargent. 2008. The market price of risk and the equity premium: A legacy of the Great Depression? *Journal of Monetary Economics* 55:454–76.

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon. 2018. Ambiguity and the historical equity premium. *Quantitative Economics* 9:945–93.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer. 2016. Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review* 106:664–98.

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* 68:1843–89.

Engle, R. F., and K. F. Kroner. 1995. Multivariate Simulteneous Generalized ARCH. *Econometric Theory* 11:122–50.

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

Epstein, L. G., and S. Ji. 2013. Ambiguous volatility and asset pricing in continuous time. *The Review of Financial Studies* 26:1740–86.

Epstein, L. G., and M. Schneider. 2007. Learning under ambiguity. *Review of Economic Studies* 74:1275–303.

Epstein, L. G., and T. Wang. 1994. Intertemporal asset pricing under Knightian uncertainty. *Econometrica* 62:283–322.

Epstein, L. G., and S. E. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57:937–69.

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

———. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25:23–49.

Fletcher, R. 1987. *Practical methods of optimization.* 2 ed. New York, NY: John Wiley and Sons.

Gallant, A. R., and J. R. Long. 1997. Estimating stochastic differential equations efficiently by minimum chi-squared. *Biometrica* 84:125–41.

Gallant, A. R., and R. E. McCulloch. 2009. On the determination of general scientific models with application to asset pricing. *Journal of the American Statistical Association* 104:117–31.

Gallant, A. R., and D. W. Nychka. 1987. Seminonparametric maximum likelihood estimation. *Econometrica* 55:363–90.

Gallant, A. R., and G. Tauchen. 1996. Which moments to match? *Econometric Theory* 12:657–81.

———. 1998. Reprojecting partially observed systems with application to interest rate diffusions. *Journal of the American Statistical Association* 93:10–24.

———. 2010. Simulated score methods and indirect inference for continuous-time models. in *Handbook of Financial Econometrics*, Amsterdam: Elsevier/North-Holland, 427–78.

———. 2014. *SNP: A program for nonparametric time series analysis – user's guide.* Pennsylvania State University and Duke University.

Gamerman, D., and H. F. Lopes. 2006. *Markov chain monte carlo: Stochastic simulation for bayesian inference.* 86, 2 ed. New York, NY: Chapman & Hall.

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

Halevy, Y. 2007. Ellsberg revisited: An experimental study. *Econometrica* 75:503–36.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–54.

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

Hansen, L. P., and T. J. Sargent. 2010. Fragile beliefs and the price of uncertainty. *Quantitative Economics* 1:129–62.

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

Ilut, C., and M. Schneider. 2014. Ambiguous business cycles. *American Economic Review* 104:2368–99.

Jahan-Parvar, M. R., and H. Liu. 2014. Ambiguity aversion and asset prices in production economies. *Review of Financial Studies* 27:3060–97.

Jarque, C. M., and A. K. Bera. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 6:255–9.

Jeong, D., H. Kim, and J. Y. Park. 2015. Does ambiguity matter? Estimating asset pricing models with a multiple-priors recursive utility. *Journal of Financial Economics* 115:361–82.

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

Ju, N., and J. Miao. 2012. Ambiguity, learning, and asset returns. *Econometrica* 80:559–91.

Klibanoff, P., M. Marinacci, and S. Mukerji. 2005. A smooth model of decision making under ambiguity. *Econometrica* 73:1849–92.

———. 2009. Recursive smooth ambiguity preferences. *Journal of Economic Theory* 144:930–76.

Leippold, M., F. Trojani, and P. Vanini. 2008. Learning and asset prices under ambiguous information. *Review of Financial Studies* 21:2565–97.

Lettau, M., S. C. Ludvigson, and J. A. Wachter. 2008. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies* 21:1653–87.

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

Liu, H., and Y. Zhang. 2015. Financial uncertainty with ambiguity and learning. *Working Paper* .

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* 45:121–45.

McConnell, M. M., and G. Perez-Quiros. 2000. Output fluctuations in the United States: What has changed since the early 1980s? *American Economic Review* 90:1464–76.

Mehra, R., and E. C. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15:145–61.

Miao, J., B. Wei, and H. Zhou. Forthcoming. Ambiguity aversion and variance premium. *Quarterly Journal of Finance* .

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

Pohl, W., K. Schmedders, and O. Wilms. 2018. Higher-order effects in asset pricing models with long-run risks. *Journal of Finance* 73:1061–111.

Schorfheide, F., D. Song, and A. Yaron. 2018. Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach. *Econometrica* 86:617–54.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–4.

Schwert, G. W. 1989. Why does stock market volatility change over time? *Journal of Finance* 44:1115–53.

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

Thimme, J., and C. Völkert. 2015. Ambiguity in the cross-section of expected returns: An empirical assessment. *Journal of Business & Economic Statistics* 33:418–29.

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

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 the 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	$\{\beta, \gamma, \psi, \eta, \mu_h, \mu_l, p_{hh}, p_{ll}, \sigma, \lambda, \sigma_d\}$
AAMSTV	$(\pi_t,\ s_t^{\sigma})$	$\{\beta, \gamma, \psi, \eta, \mu_h, \mu_l, p_{hh}^{\mu}, p_{ll}^{\mu}, \sigma_h, \sigma_l, p_{hh}^{\sigma}, p_{ll}^{\sigma}, \lambda, \sigma_d\}$
AALRRSV	$(\hat{x}_t,\ \nu_t,\ \sigma_t)$	$\{\beta, \gamma, \psi, \eta, \mu_c, \rho_x, \varphi_x, \lambda, \varphi_d, \mu_s, \rho_s, \sigma_w\}$
EZMS	π_t	$\{\beta, \gamma, \psi, \mu_h, \mu_l, p_{hh}, p_{ll}, \sigma, \lambda, \sigma_d\}$
EZMSTV	$(\pi_t,\ s_t^{\sigma})$	$\{\beta, \gamma, \psi, \mu_h, \mu_l, p_{hh}^{\mu}, p_{ll}^{\mu}, \sigma_h, \sigma_l, p_{hh}^{\sigma}, p_{ll}^{\sigma}, \lambda, \sigma_d\}$
EZLRRSV	(x_t, σ_t^2)	$\{\beta, \gamma, \psi, \mu_c, \rho_x, \varphi_x, \lambda, \varphi_d, \varphi_c, \mu_s, \rho_s, \sigma_w\}$

This table summarizes relevant state variables and structural parameters for each asset pricing model described in Section 1.

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						831.42		
Log likelihood						-391.96		
MCMC repetitions		10,000				12,000		

This table presents prior and posterior marginal means, medians, and the 5th and 95th 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 AAMSTV 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^{μ}_{ll}	0.508	0.531	0.281	0.719	0.668	0.686	0.504	0.746
p^{μ}_{hh}	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^{σ}_{ll}	0.849	0.859	0.734	0.953	0.986	0.990	0.948	0.996
p^{σ}_{hh}	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, the 5th and 95th percentiles of model parameters for the AAMSTV 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					759.78			
Log likelihood					-353.98			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, the 5th and 95th 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					813.68			
Log likelihood					-385.25			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, the 5th and 95th 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 EZMSTV Model

	Prior				Posterior			
Parameter	Mean	Median	5%	95%	Mean	Median	5%	95%
β	0.991	0.991	0.991	0.991	0.972	0.970	0.964	0.983
γ	7.995	8.250	6.750	8.250	6.153	6.219	3.844	8.469
ψ	0.975	0.938	0.938	1.063	1.673	1.664	1.086	2.273
p_{ll}^{μ}	0.595	0.594	0.594	0.594	0.888	0.894	0.780	0.960
p_{hh}^{μ}	0.812	0.813	0.813	0.813	0.986	0.985	0.976	0.996
μ_l	-0.035	-0.035	-0.035	-0.035	-0.026	-0.025	-0.036	-0.017
μ_h	0.037	0.037	0.037	0.037	0.025	0.026	0.016	0.034
p_{ll}^{σ}	0.856	0.859	0.766	0.984	0.977	0.979	0.938	0.998
p_{hh}^{σ}	0.837	0.828	0.734	0.953	0.921	0.929	0.856	0.974
σ_l	0.006	0.005	0.005	0.009	0.024	0.024	0.016	0.028
σ_h	0.006	0.006	0.006	0.006	0.044	0.048	0.021	0.058
λ	3.271	3.250	3.250	3.750	2.270	2.242	1.602	2.836
σ_d	0.102	0.102	0.086	0.117	0.126	0.122	0.099	0.179
BIC					768.55			
Log likelihood					-356.21			
MCMC repetitions	10,000				12,000			

This table presents prior and posterior marginal means, medians, the 5th and 95th percentiles of model parameters for the EZMSTV 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
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						815.94		
Log likelihood						-382.06		
MCMC repetitions		10,000				12,000		

This table presents prior and posterior marginal means, medians, the 5th and 95th 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 9
Diagnostic results for AAMSTV

	$\sigma(\Delta c)$	$\sigma(\Delta d)$	$\mathbb{E}(r_t^f)$	$\sigma(r_t^f)$	$\mathbb{E}(r_t)$	$\sigma(r_t)$
AAMSTV vs. EZMSTV	0.85	-0.48	5.03	-1.05	-2.27	-2.36
AAMSTV vs. AAMS	-1.00	-2.32	2.69	-1.29	-0.21	-2.71

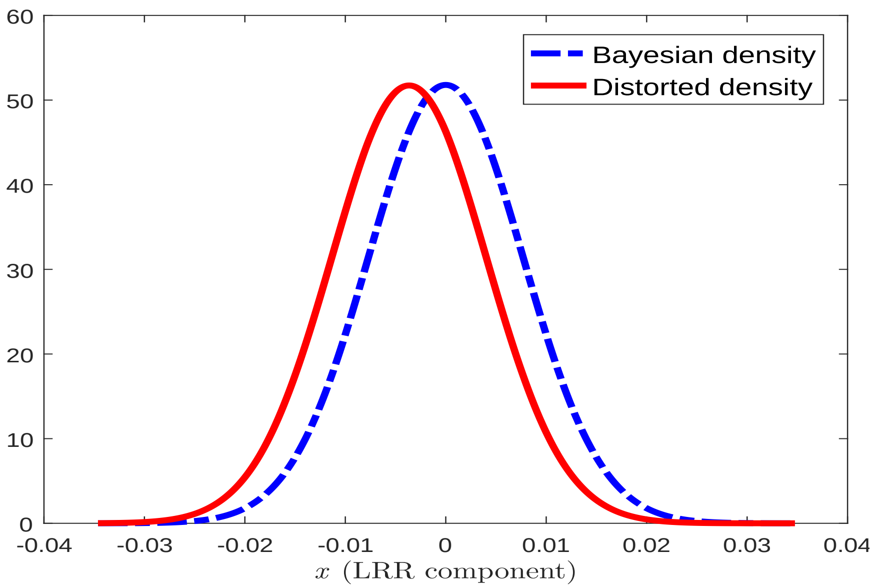
This table summarizes diagnostics results for the AAMSTV, AAMS, and EZMSTV models. The table presents t-statistics for moments of the preferred model AAMSTV vs. AAMS and t-statistics for moments of the preferred model AAMSTV against EZMSTV. The t-statistic is defined by $t = (m_1 - m_2)/\sigma_2$, where m_1 and m_2 are the posterior modes of a moment of interest for models 1 and 2, respectively, and σ_2 is the standard deviation of the posterior distribution of the moment under model 2. We denote the AAMSTV model by “2.”

Table 10
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
St. dev.	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
				AAMSTV			
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
St. dev.	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
St. dev.	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
St. dev.	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
				EZMSTV			
Mean	2.413	1.276	4.579	15.951	2.643	2.362	0.699
Median	2.484	1.260	4.712	15.828	2.156	1.944	0.689
St. dev.	0.648	0.170	1.618	1.893	1.511	1.236	0.211
95%	3.330	1.561	7.056	19.216	6.024	4.941	1.103
5%	1.072	1.016	2.017	12.883	0.993	0.900	0.378
				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
St. dev.	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

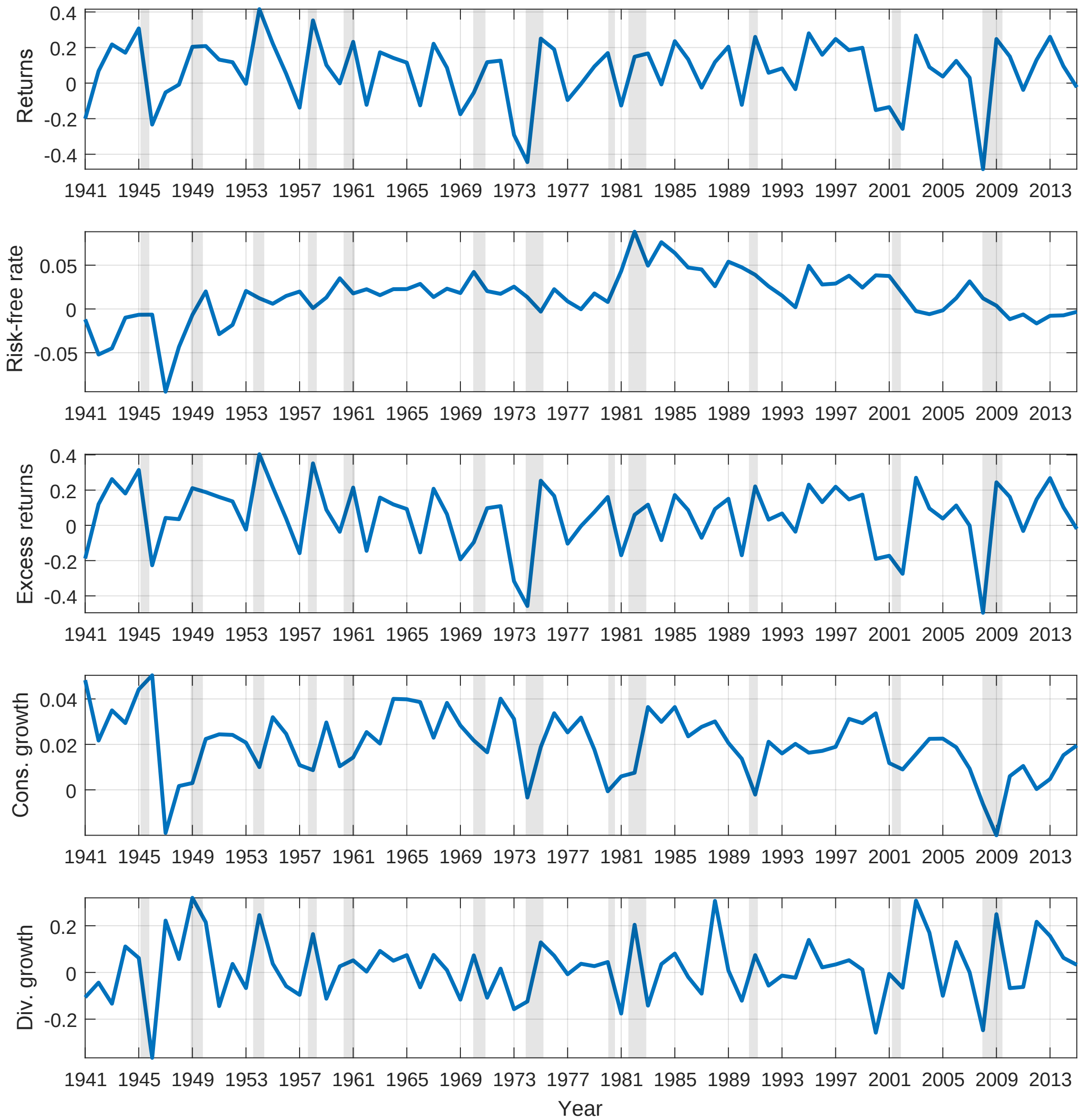
This table presents unconditional financial moments generated from the estimated models. These quantities are computed from simulated paths based on 12,000 Bayesian MCMC estimates of the structural parameters. $\mathbb{E}(r_t^f)$ and $\mathbb{E}(r_t - r_t^f)$ are mean risk-free rate and mean equity premium, 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



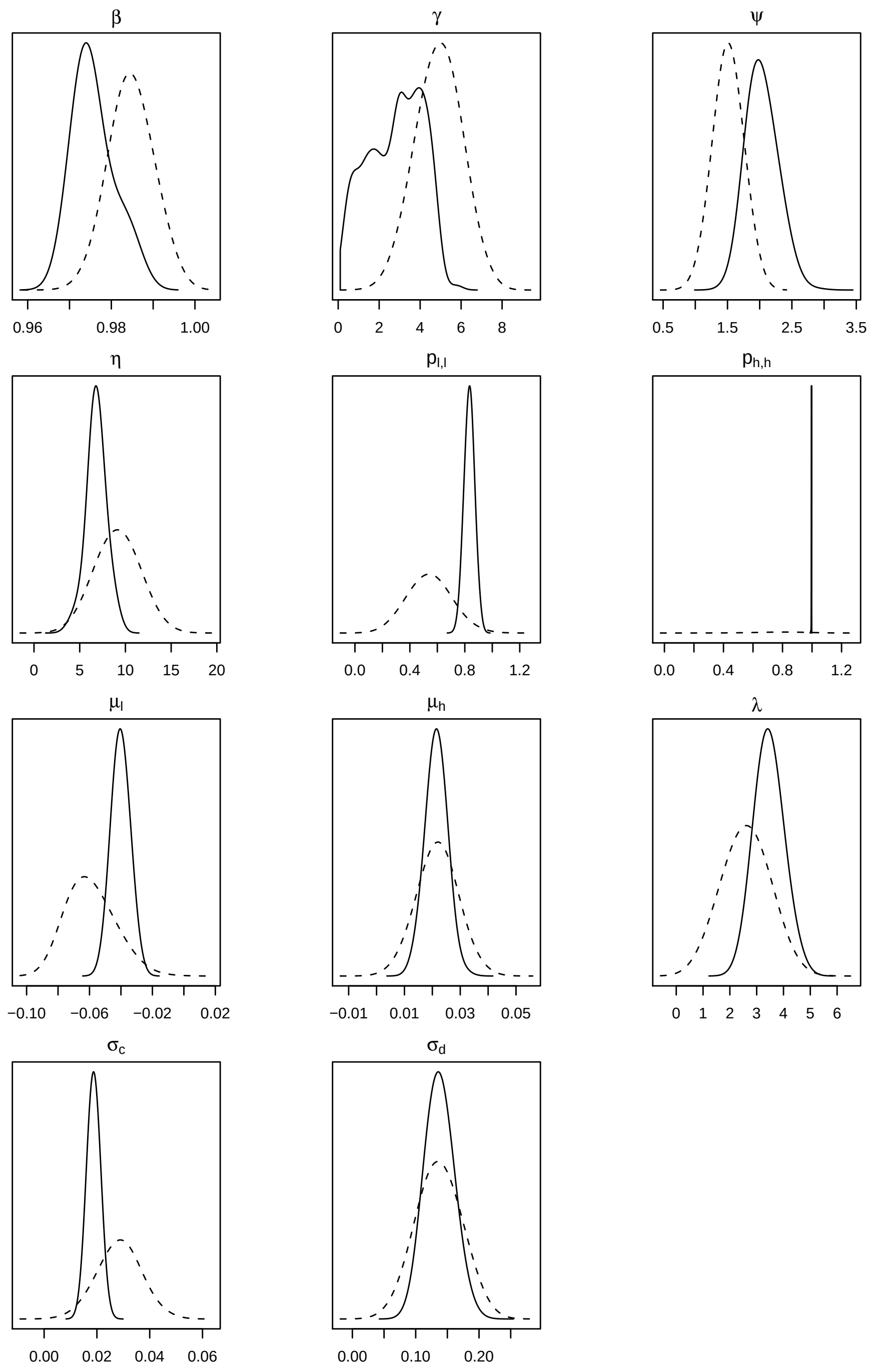
This figure plots Bayesian density and distorted density of the long-run risk component x for the AALRRSV model. The Bayesian density is $x_t \sim N(\hat{x}_t, \nu_t)$, and the distorted density is $\tilde{f}(x_t|\hat{x}_t, \nu_t, t)$. The distorted density is generated from solving the model. The state 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



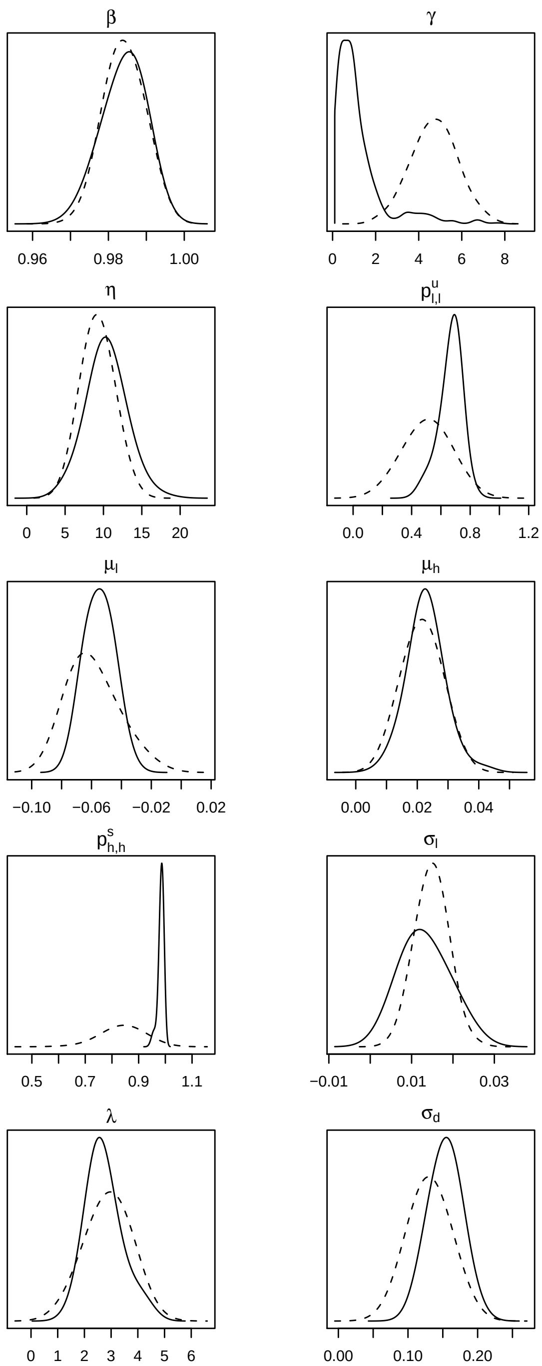
The figure shows CRSP value-weighted index returns, one-year Treasury bill rates, excess returns, per-capita log consumption growth, and log dividend growth rates for the 1941–2015 period. All series plotted are at an annual frequency and in real terms. Shaded areas represent NBER recessions.

Figure 3
Prior and posterior densities of parameters of the AAMS model



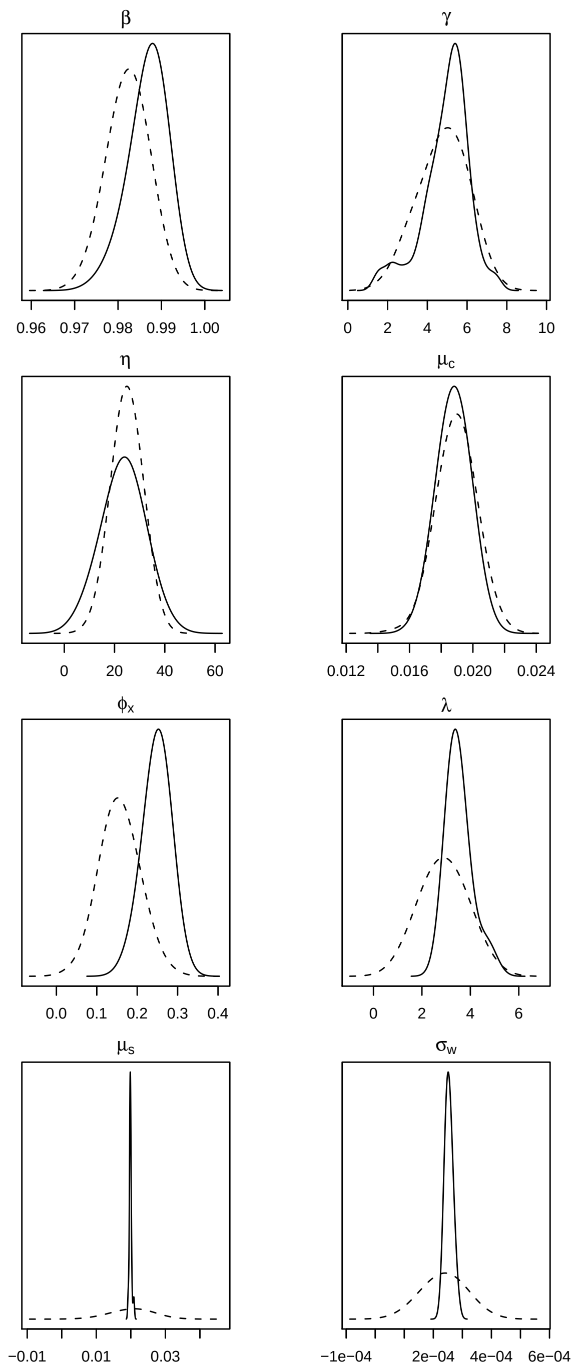
This figure plots prior and posterior densities of parameters in the AAMS model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 4
Prior and posterior densities of parameters of the AAMSTV model



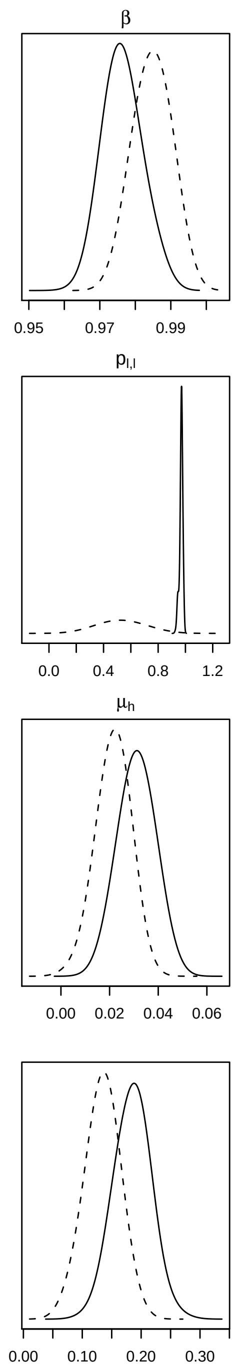
This figure plots prior and posterior densities of parameters in the AAMSTV model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 5
Prior and posterior densities of parameters of the AALRRSV model



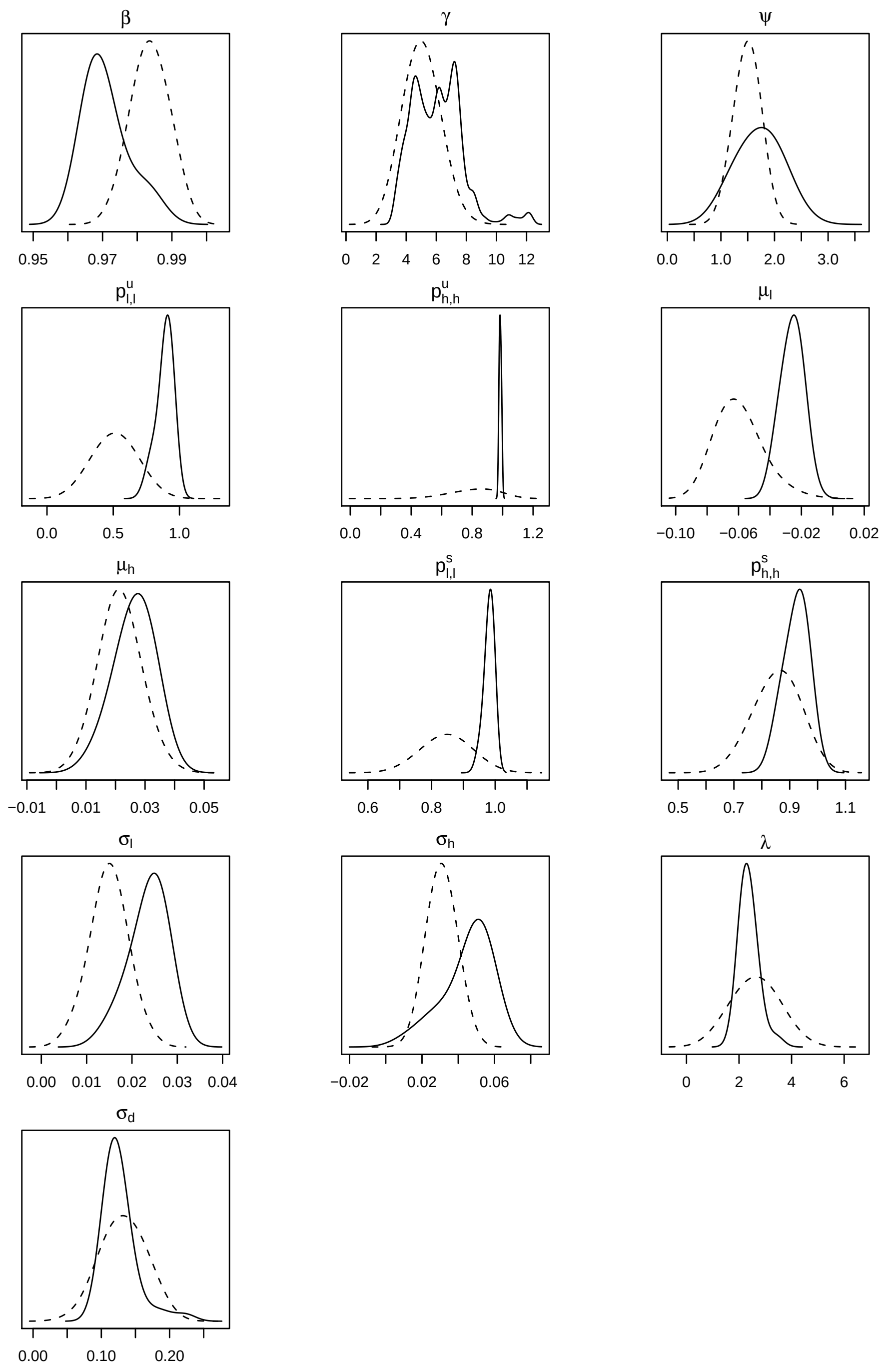
This figure plots prior and posterior densities of parameters in the AALRRSV model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 6
Prior and posterior densities of parameters of the EZMS model



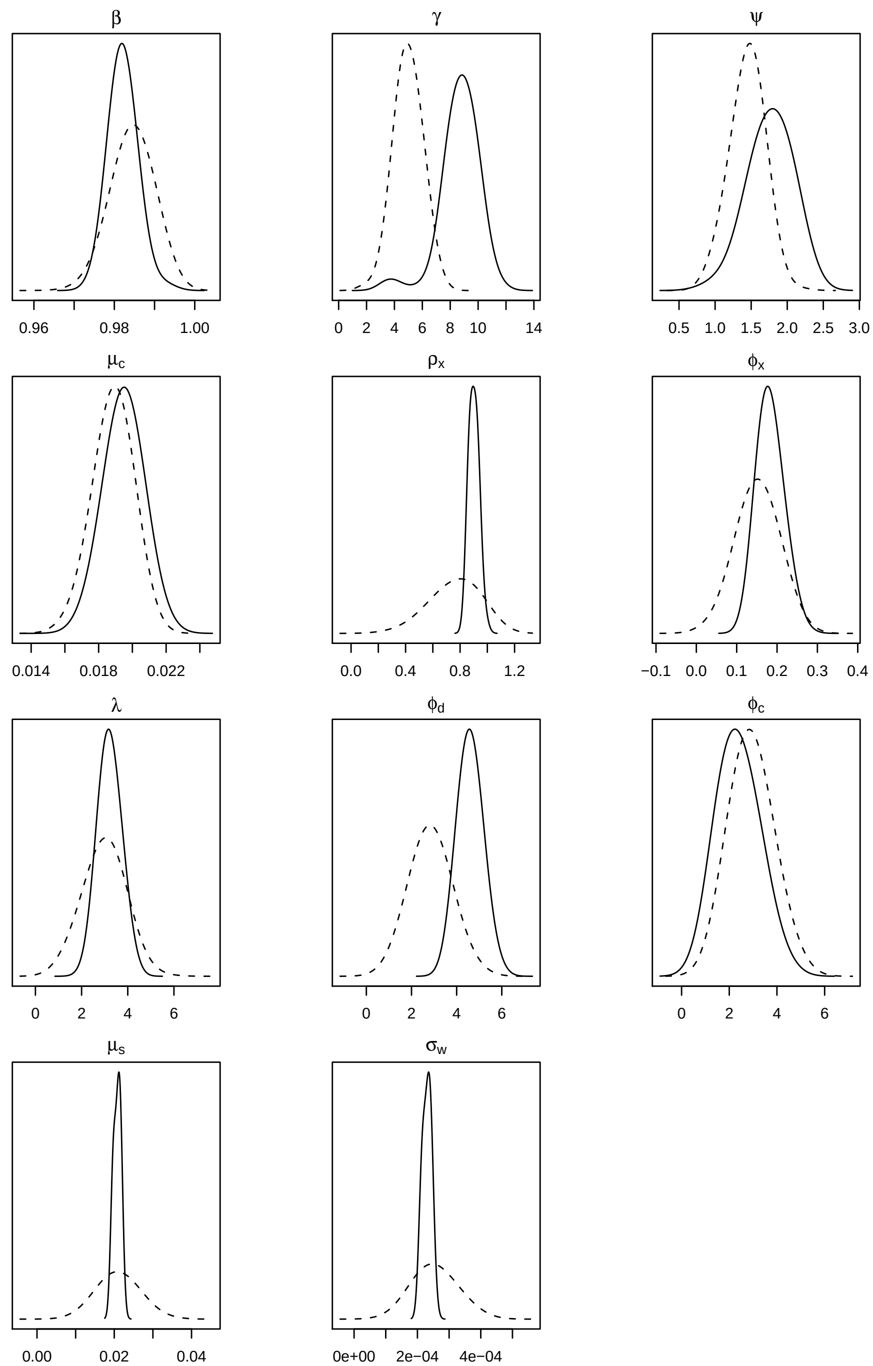
This figure plots prior and posterior densities of parameters in the EZMS model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 7
Prior and posterior densities of parameters of the EZMSTV model



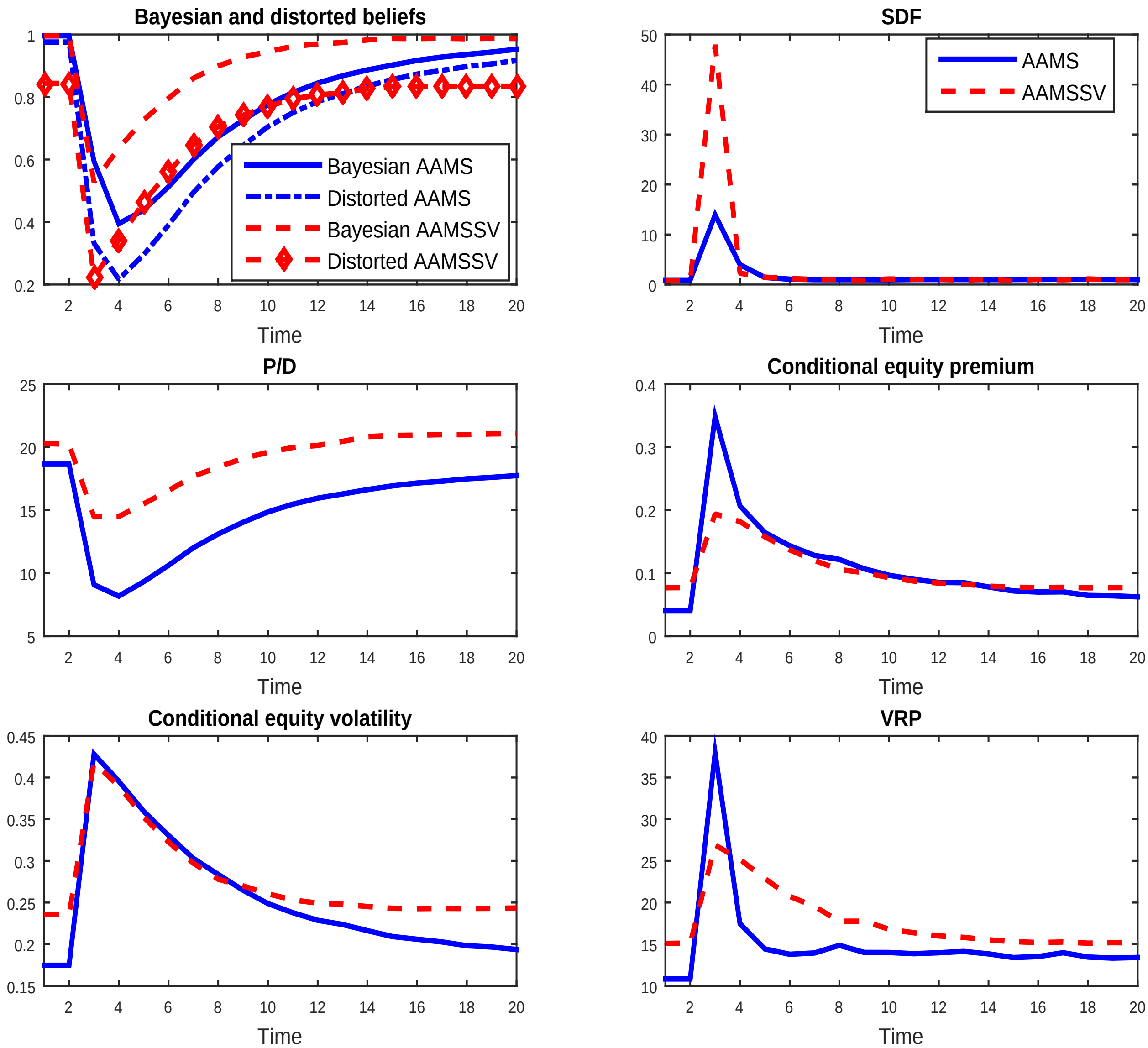
This figure plots prior and posterior densities of parameters in the EZMSTV model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 8
Prior and posterior densities of parameters of the EZLRRSV model



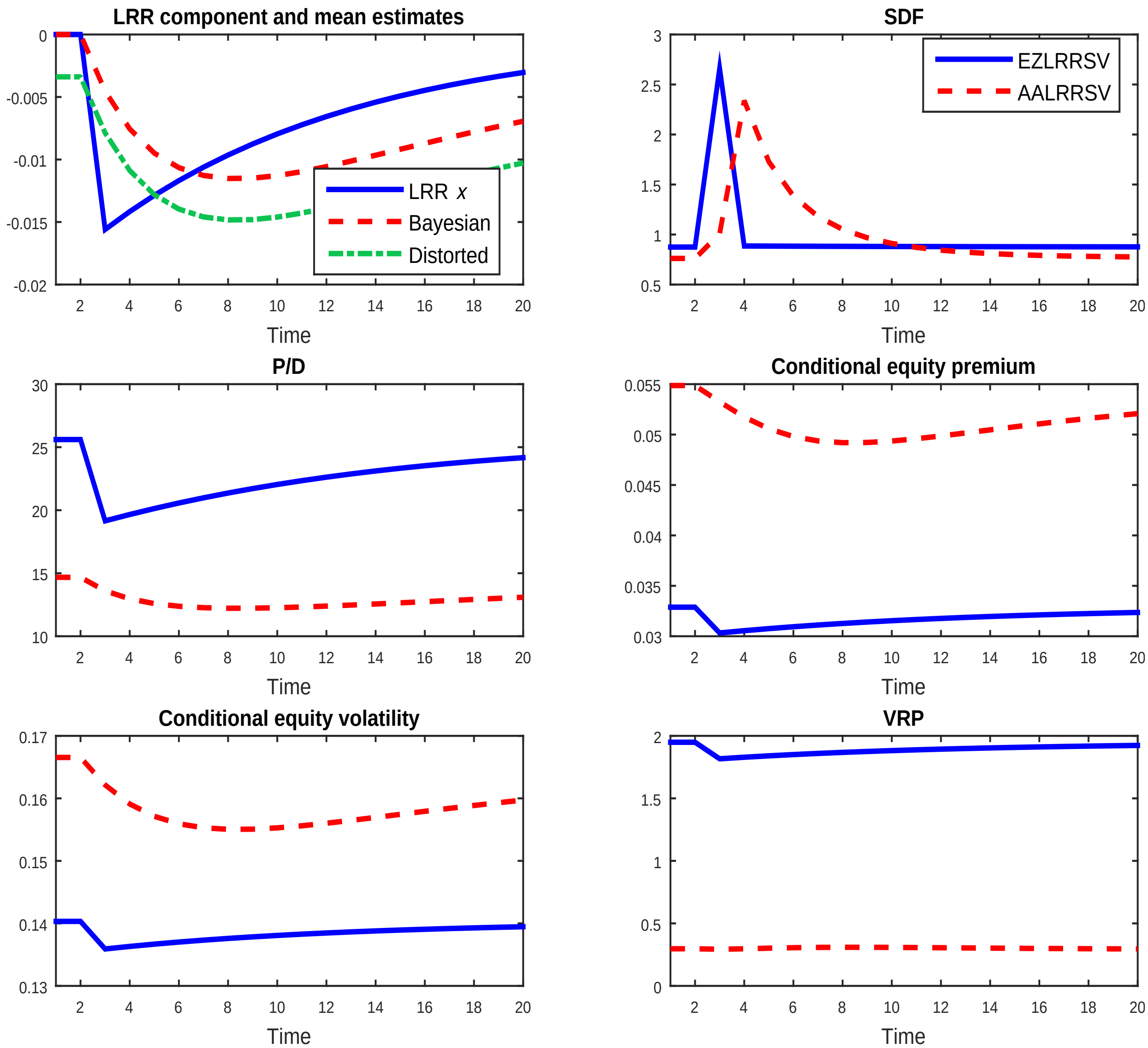
This figure plots prior and posterior densities of parameters in the EZLRRSV model. The solid lines depict posterior densities, and the dotted lines depict prior densities. The results are based on the U.S. annual data for 1941–2015.

Figure 9
Impulse-response functions: AAMS and AAMSTV



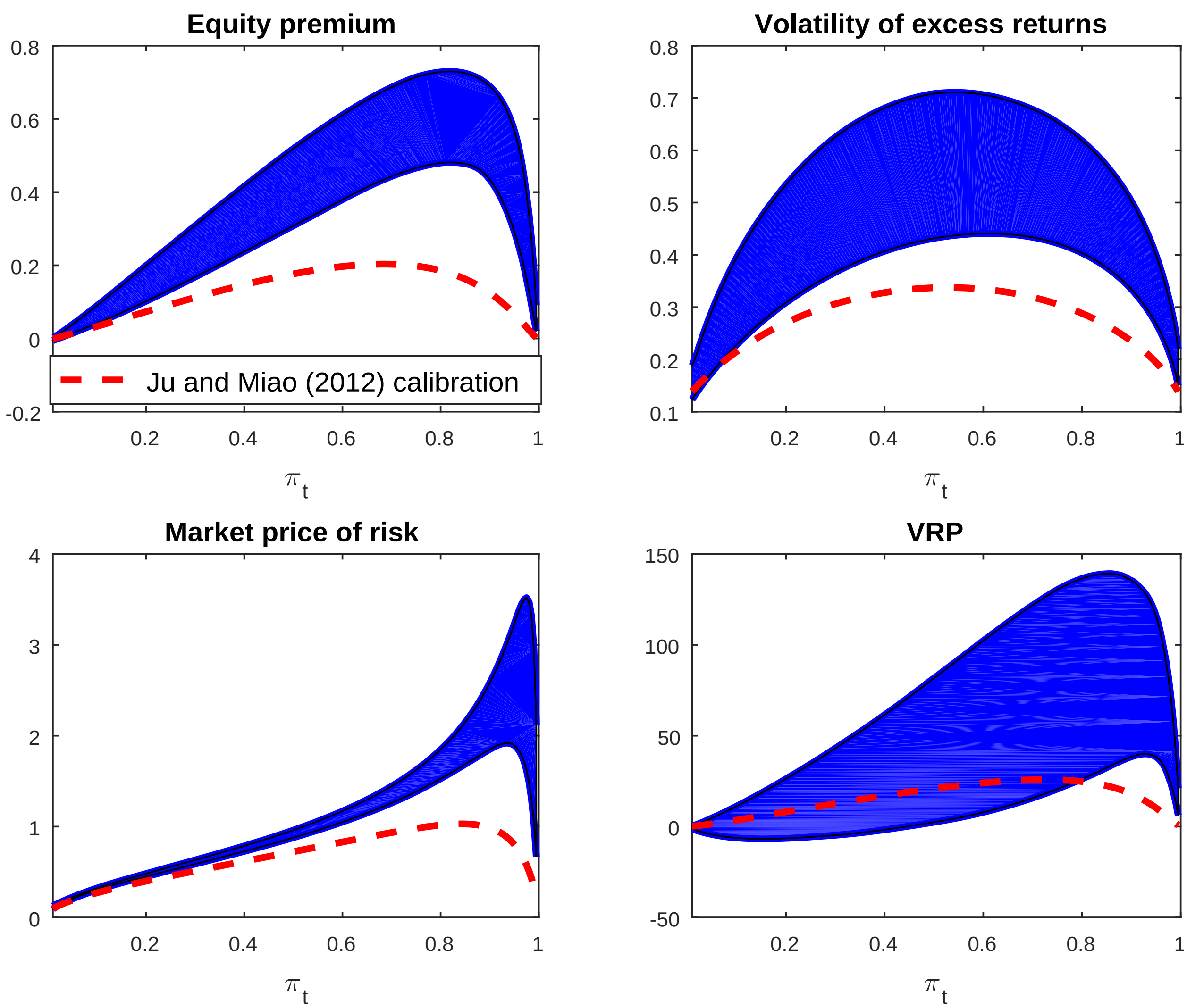
This figure plots the mean impulse-response functions for the AAMS and AAMSTV models when the mean consumption growth state shifts from μ_h to μ_l in the third period. Before the realization of the shock, mean consumption growth is assumed to stay in state μ_h without the impact of innovation shocks. The growth rate of consumption follows the respective Markov-switching models after the regime shift. We compute the mean impulse-response across 5,000 simulated paths of consumption growth. The results plotted are for model parameters set at posterior means of Bayesian MCMC estimates.

Figure 10
Impulse-response functions: AALRRSV and EZLRRSV



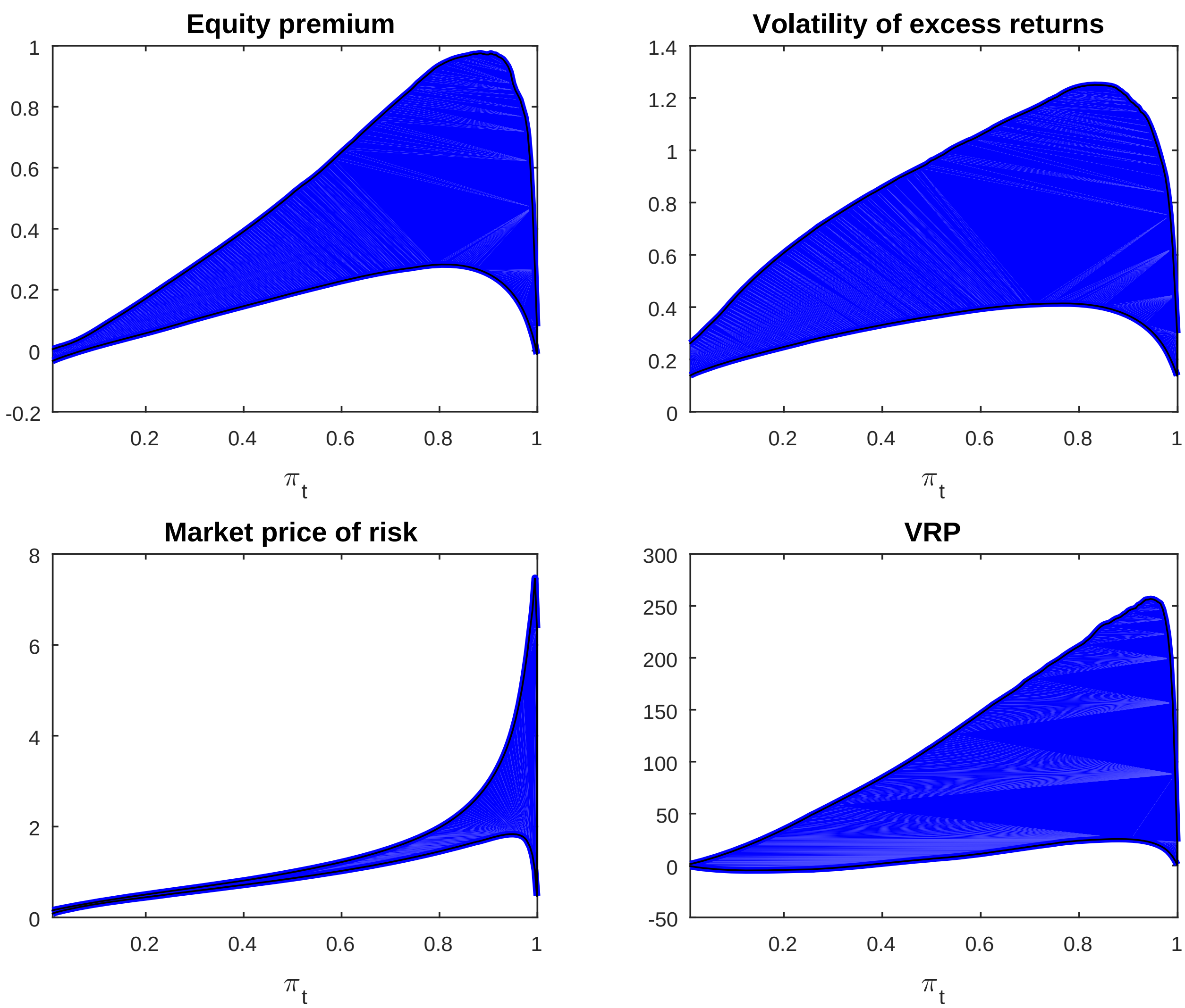
This figure plots the mean impulse-response functions for models AALRRSV and EZLRRSV when a shock of size $-4\varphi_x\mu_s$ to x_t occurs in the third period. Before the realization of the shock, the AALRRSV economy is assumed to stay in state $(\hat{x}_t, \nu_t, \sigma_t)$ for which $\Delta c_t = \mu_c,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

Figure 11
AAMS model: Conditional financial moments



This figure plots conditional financial moments ranging from the 5th to 95th 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 12
AAMSTV model: Conditional financial moments



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