FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

Ambiguity, Long-Run Risks, and Asset Prices

Bin Wei

Working Paper 2021-21
September 2021

Abstract: I generalize the long-run risks (LRR) model of Bansal and Yaron (2004) by incorporating recursive smooth ambiguity aversion preferences from Klibanoff et al. (2005, 2009) and time-varying ambiguity. Relative to the Bansal-Yaron model, the generalized LRR model is as tractable but more flexible due to its separation of ambiguity aversion from both risk aversion and the intertemporal elasticity of substitution. This three-way separation allows the model to further account for the variance premium puzzle besides the puzzles of the equity premium, the risk-free rate, and the return predictability. Specifically, the model matches reasonably well key asset-pricing moments with risk aversion under 5. Model calibration shows that the ambiguity aversion channel accounts for 77 percent of the variance premium and 40 percent of the equity premium.

JEL classification: G12, G13, D81, E44

Key words: smooth ambiguity aversion, long-run risks, equity premium puzzle, risk-free rate puzzle, variance premium puzzle, return predictability

https://doi.org/10.29338/wp2021-21

Please address questions regarding content to Bin Wei, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309, bin.wei@atl.frb.org.

1 Introduction

The equity premium puzzle (Mehra and Prescott, 1985) and the risk-free rate puzzle (Weil, 1989) are two major puzzles in the asset pricing literature. In standard representative-agent asset pricing models with the power utility function, risk aversion and the intertemporal elasticity of substitution (IES) are reciprocally related. This reciprocal relation poses a challenge to jointly explain these puzzles. On the one hand, the large equity premium in the data requires the representative agent to be implausibly risk averse. On the other hand, such strong risk aversion—needed to generate the large equity premium—generates an implausibly high risk-free rate, which is at odds with the low rate observed in the data.

The long-run risks (LRR) model proposed in Bansal and Yaron (2004) constitutes an important contribution in the asset pricing literature for its ability to jointly explain the above two puzzles, as well as the return predictability puzzle that dividend yields predict the stock market return. The Epstein-Zin preferences used in the LRR model sever the tight link between risk aversion and the IES. As a result, both risk aversion and the IES can be made large enough that the standard LRR model can generate the large equity premium and the low risk-free rate simultaneously. Furthermore, the preference for early resolution of uncertainty and the aversion to the long-run consumption risk help to further amplify the equity premium, and drive both the price-dividend ratio and the stock market return, thereby giving rise to the return predictability of the price-dividend ratio.

The standard LRR model is, however, silent about the *variance premium puzzle*: (i) the variance premium—defined as the difference between the expected stock market variances under the risk-neutral and objective measures—is too large and volatile in the data to be explained by consumption-based asset pricing models; and (ii) the variance premium predicts the stock market return.

In this paper we generalize the standard LRR model in Bansal and Yaron (2004) by incorporating *recursive smooth ambiguity aversion preferences* from Klibanoff et al. (2005, 2009) to further account for the variance premium puzzle in addition to the aforementioned puzzles. The recursive smooth ambiguity aversion preferences allow for a three-way separation between ambiguity aversion, risk aversion, and the IES. The Epstein-Zin preferences are embedded as a special case when ambiguity aversion is equal to risk aversion. As the Epstein-Zin preferences contribute to the success of the standard LRR model by separating between risk aversion and the IES, the further separation of ambiguity aversion achieved by the recursive smooth ambiguity aversion preferences enhances the flexibility of the generalized LRR model. The generalized LRR model can explain all of the aforementioned asset pricing puzzles (including the variance premium puzzle) with risk aversion less than 5.

In our generalized LRR model, the representative agent is averse to *ambiguity* about economic volatility. Consistent with the notation of "Knightian uncertainty" (Knight, 1921; Keynes, 1936), we refer to *ambiguity* as the situation where the decision maker is uncertain about probability measures due to, for example, very imprecise information, and refer to *risk* as the situation where there exists a probability measure for certain to guide choice. Specifically, we assume that there exists a set of possible conditional probability distributions regarding economic volatility and the agent is ambiguous about which conditional distribution will occur. Roughly speaking, the size of the set of possible distributions measures the degree of ambiguity. Our notation of ambiguity can also be interpreted as model uncertainty in the robustness theory developed by Hansen (2007) and Hansen and Sargent (2001, 2007, 2008), which is about decision-making that is robust to possible model misspecification.

The generalized LRR model is not only flexible but also tractable. Specifically, the equity premium, the risk-free rate, and the variance premium all admit closed-form expressions under the log-linearized model. The closed-form expressions are similar to those in Bansal and Yaron (2004), but differ by an important additional term that captures the ambiguity aversion channel. This additional term in the expression of the equity or variance premium measures ambiguity-induced compensation, which increases with either ambiguity or ambiguity aversion. We thus refer to this term as the "*ambiguity premium*." Quantitatively, we show that the standard LRR model with an ambiguity-neutral agent has difficulty matching the mean and volatility of the variance premium in the data. Introducing smooth ambiguity preferences results in an eightfold increase in the *level* of the model-implied variance premium to match the data, and also enables the model to reproduce about 25% of the *volatility* of the variance premium in the data. Furthermore, we find that the ambiguity premium component accounts for 77% of the variance premium, and for 40% of the equity premium. Our findings suggest that a large portion of the variance premium is attributable to the ambiguity aversion channel, and that ambiguity aversion has a much larger effect on the variance premium than on the equity premium.

The ambiguity aversion channel amplifies the variance premium by distorting the ambiguity-averse agent's beliefs. As detailed later, the variance premium in the generalized LRR model can be written as a weighted average of the model-specific variance premiums that would prevail under a specific probability distribution (i.e., model). The corresponding weight is the *distorted* probability of the specific scenario; that is, the real probability distribution distorted or tilted toward bad states. The distortion is caused by the agent's pessimistic beliefs. The more ambiguity averse the agent is or the more ambiguity there is, the more tilted toward bad states the weights are. As a result, if the agent is more ambiguity averse or the environment becomes more ambiguous, more compensation is demanded, resulting in a higher variance premium. A similar intuition applies to other asset pricing moments. As such, the ambiguity aversion channel arising from the three-way separation provides a new amplification mechanism that helps the generalized LRR model to better match the asset pricing moments in the data.

In the standard LRR model, the equity premium is sizable and time-varying, but the variance premium is small and constant. This outcome implies that the variance premium has zero volatility and no power of predicting the stock market return. Through the ambiguity aversion channel, the generalized LRR model generates a time-varying variance premium and the return predictability as more ambiguity drives up both the variance premium and the stock market return. The generalized LRR model is able to match about 25% of the volatility of the variance premium observed in the data. Our findings suggest that both time variation and the return predictability of the variance premium are both closely tied to time variation in ambiguity, rather than the underlying economic volatility.

Lastly, we make inferences about the latent long-run risks and ambiguity via particle filtering based on Schorfheide et al. (2018). Specifically, we derive a nonlinear state-space system that relates the unobserved state to the observables based on the annual data of consumption and dividend growth rates, the market return, the risk-free rate, the price-dividend ratio, and the variance premium. The state-space system is linear conditional on the latent volatility and ambiguity states, and it can be used to back out the latent state variables via particle filtering (see Herbst and Schorfheide, 2015). Using the filtered state variables, we find that the generalized LRR model tracks the observed data series reasonably well, particularly the evolution of the variance premium over time. The state-space approach also allows us to conduct variance decomposition to analyze the relative contributions of the long-run risks, economic volatility, and ambiguity to the volatilities of the stock market return, the risk-free rate, the price-dividend ratio, and the variance premium. The variance decomposition results suggest that almost all time variation in the variance premium is driven by time variation in ambiguity. Moreover, the variabilities of the stock market return and the risk-free rate are, in almost equal parts, attributable to the variations in economic volatility and ambiguity.

This paper contributes to the literature on the long-run risks asset pricing models. Following the seminal work by Bansal and Yaron (2004), there have been important subsequent work to explain asset pricing puzzles for the term structure (Piazzesi and Schneider, 2007), for bond risk premiums (Bansal and Shaliastovich, 2013), for credit spreads (Chen, 2010), for option prices ((Drechsler and Yaron, 2011; Eraker and Shaliastovich, 2008), for cross-sectional stock returns (Bansal et al., 2005; Hansen et al., 2008; Bansal et al., 2009), for the wealth-consumption ratio (Ai, 2010), and for exchange rate (Colacito and Croce, 2011).

Building upon the literature on ambiguity, robustness, and asset pricing,[1] this paper introduces smooth ambiguity preferences into the long-run risks model in an attempt to solve the variance premium puzzle. The present paper is thus closely related to Branger et al. (2016) and Gallant et al. (2018). These papers also introduce smooth ambiguity preferences into the long-run risks model. Branger et al. (2016) assumes that the conditional volatility of consumption growth is unobservable and constructs the distribution of the unobservable volatility based on analysts' GDP forecasts in the Survey of Professional Forecasters. Gallant et al. (2018) assumes that the long-run risks component is unobservable and infers the distribution of the unobservable long-run risks component via Kalman filtering. Complementing these studies, the present paper takes a dramatically different approach to incorporate smooth ambiguity preferences. Both of the above papers assume that the agent is ambiguous about the *value* of a specific variable. Our approach instead models the agent as facing ambiguity about the *distribution* of a certain variable. Our notion of ambiguity is based on Knightian uncertainty; that is, ambiguity refers to the situation where no known *probabilities* are available.

Our paper also contributes to the literature on the variance premium (e.g., Bakshi et al., 2015; Bollerslev et al., 2009; Carr and Wu, 2009; Todorov, 2009). In order to explain the variance premium puzzle, Drechsler and Yaron (2011) introduce jumps into the standard LRR model and show that the extended model is able to generate many of the quantitative features of the variance premium. Rather than generalizing cash flow processes in the standard LRR model, these authors note that one could generalize the preferences further: "*[a] possible direction for generating interesting transient dynamics like the ones documented here is by generalizing preferences to include features of ambiguity aversion and a desire for robustness.*" A few prior studies provide ambiguity-based explanations for the variance premium puzzle. Drechsler (2013) extends the Bansal-Yaron model in a continuous-time robust control framework to further explain anomalies in the option market, namely the large magnitude of the variance premium and the "volatility skew." To account for these anomalies and properties of equity returns and the risk-free rate, Drechsler (2013) develops a very flexible framework in which ambiguity operates through multiple channels. So it is very difficult to know which channels are important to explain a certain anomaly (e.g., the variance premium). In contrast, our model is parsimonious, aiming to primarily explain the variance

[1]See Epstein and Wang (1994), Chen and Epstein (2002), Cao et al. (2005), Garlappi et al. (2007), Epstein and Schneider (2008), Leippold et al. (2008), Routledge and Zin (2009), Drechsler (2013), Collard et al. (2018), and Shi (2019) for asset-pricing applications of multiple-priors models, and Ju and Miao (2012), Jahan-Parvar and Liu (2014), Ai and Bansal (2018), Miao et al. (2019) for asset-pricing applications of smooth ambiguity models. See Hansen and Sargent (1999), Hansen and Sargent (2001), Anderson et al. (2003), Uppal and Wang (2003), Maenhout (2004), Liu et al. (2005), and Cagetti et al. (2015) for models of robustness and applications.

premium puzzle through the channel of ambiguity in economic volatility. Furthermore, our model reproduces the large variance premium without introducing jumps, whereas shutting off the jumps channel in Drechsler (2013) would reduce the variance premium by an order of magnitude. Lastly, the multiple-priors approach used in Drechsler (2013) is embedded as a special case of smooth ambiguity preferences in this paper in which the ambiguity aversion coefficient is infinitely large. As a result, relative to Drechsler (2013), we can separately examine the role of ambiguity aversion in generating the dynamics of the variance premium, and estimate the magnitude of the ambiguity aversion coefficient based on particular filtering. Miao et al. (2019) show that the regime-switching model in Ju and Miao (2012) can generate a sizable variance premium as in the data. However, due to lack of tractability, their model does not admit closed form expressions for the variance premium and can only be numerically solved. In contrast, our generalized model is not only tractable, but can also distinguish between risk and ambiguity which are inseparable in their paper in which the agent's belief, the only state variable, drives both.

The rest of the paper is organized as follows. Section 2 presents the generalized LRR model. We then discuss its asset pricing implications using analytical log-linearization approximations in Section 3. In Section 4, we describe the data sample, our calibration methodology, and we discuss our central findings. Section 5 concludes. We delegate proofs and details in the log-linearization analysis to Appendix A and particle-filter-based estimation to Appendix B.

2 A Generalized Long-Run Risks Model

We generalize the standard long-run risks model of Bansal and Yaron (2004) by introducing recursive smooth ambiguity aversion preferences. Our generalized LRR model departs from the Bansal-Yaron model in terms of modeling of ambiguity and ambiguity aversion.

As in the standard LRR model, the dynamics for consumption and dividend growth rates, $\Delta c_{t+1} \equiv \log(C_{t+1}/C_t)$ and $\Delta d_{t+1} \equiv \log(D_{t+1}/D_t)$, respectively, are given by

$$\Delta c_{t+1} = \mu_c + x_t + \sigma_t \epsilon_{c,t+1}, \tag{1}$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \varphi_d \sigma_t (\rho_d \epsilon_{c,t+1} + \sqrt{1-\rho_d^2}\epsilon_{d,t+1}), \tag{2}$$

$$x_{t+1} = \rho_x x_t + \varphi_x \sigma_t \epsilon_{x,t+1}, \tag{3}$$

where $\epsilon_{c,t+1}$, $\epsilon_{d,t+1}$, and $\epsilon_{x,t+1}$ have independent standard normal distributions, μ_c and μ_d represent their respective unconditional average growth rates, and σ_t represents fluctuating economic volatility. The small persistent component x_t captures the long-run consumption risks in the economy. Dividend growth is exposed to the long-run risks, but also to short-run consumption risks captured by the coefficient of correlation ρ_d.

Ambiguity. As one of key departures from Bansal and Yaron (2004), we introduce ambiguity into the model such that the agent is ambiguous about economic volatility σ_t^2. Specifically, the economic volatility follows the following process

$$\sigma_{t+1}^2 = \sigma^2 + \rho_\sigma \left(\sigma_t^2 - \sigma^2\right) + \varphi_\sigma \epsilon_{\sigma,t+1}, \tag{4}$$

where the shock $\epsilon_{\sigma,t+1}$ has a conditional normal distribution with its mean $\widetilde{z}_t$ being stochastic and normally distributed

$$\epsilon_{\sigma,t+1} \sim \mathcal{N}\left(\widetilde{z}_t, 1\right), \tag{5}$$

$$\widetilde{z}_t \sim \mathcal{N}\left(0, \tau_t^2\right). \tag{6}$$

The degree of ambiguity is determined by τ_t^2 and evolves as follows:

$$\tau_{t+1}^2 = \tau^2 + \rho_\tau \left(\tau_t^2 - \tau^2\right) + \varphi_\tau \sigma_t \epsilon_{\tau,t+1}, \tag{7}$$

where the shock $\epsilon_{\tau,t+1}$ has a standard normal distribution. Absent ambiguity (i.e., $\tau = 0$), the generalized LRR model reduces to the Bansal-Yaron model.

Our notation of ambiguity builds upon "Knightian uncertainty" (Knight, 1921; Keynes, 1936) in the sense that the agent is uncertain about the conditional probability distributions of σ_t^2. It can also be interpreted as model uncertainty in robustness theory.[2] From the perspective of the agent, there are an infinite number of possible "models" for the process of σ_t^2 because each random draw $\widetilde{z}_t = z$ leads to a "model" or a conditional distribution of $\mathcal{N}\left(\sigma^2 + \rho_\sigma \left(\sigma_t^2 - \sigma^2\right) + \varphi_\sigma z, \varphi_\sigma^2\right)$ regarding the economic volatility next period.

Depending on whether ambiguity is time-invariant or not, there are two cases to consider. We refer to the *homoskedastic*-ambiguity case with constant ambiguity (i.e., $\tau_t \equiv \tau$ for any time t) as Case I of the generalized LRR model, or simply as "gLRR1." We will show shortly that with a sufficiently large coefficient of ambiguity aversion, this case can generate a large enough variance premium to match the magnitude in the data. However, the variance premium in this case is constant. This motivates us to consider the *heteroskedastic*-ambiguity case with fluctuating ambiguity. We refer to this case as Case II of the generalized LRR model, or simply as "gLRR2." Note that in both cases the equity premium is time-varying, driven partly by fluctuating economic volatility. Moreover, economic volatility also feeds back positively into ambiguity because the diffusion term in Eq. (7) is proportional to σ_t.

[2] For this reason, we use ambiguity and model uncertainty interchangeably in the paper.

In times of greater volatility, the agent faces more ambiguity as the dispersion among all possible models increases.

The timing in the model is as follows. At the beginning of period t when the agent makes consumption-investment decisions, he observes the history of consumption and dividends up to the current period $s^t \equiv \{C_0, D_0, C_1, D_1, ..., C_t, D_t\}$. However, he does not observe $\widetilde{z}_t$ and thus faces ambiguity as described above. Ambiguity is only resolved following the realization of $\widetilde{z}_t$ after decisions have been made. Then in the end of period t, all shocks (e.g., $\epsilon_{x,t+1}$) occur and then consumption and dividends are realized. The same sequence of events repeats in all future periods.

Before we turn to the recursive smooth ambiguity aversion preferences, it is convenient to express the dynamics of the system in vector form:

$$Y_{t+1} = \mu + FY_t + G_t \epsilon_{t+1}, \tag{8}$$

where

$$Y_t \equiv \begin{pmatrix} \Delta c_t \\ \Delta d_t \\ x_t \\ \sigma_t^2 \\ \tau_t^2 \end{pmatrix}, \mu \equiv \begin{pmatrix} \mu_c \\ \mu_d \\ 0 \\ (1-\rho_\sigma)\,\sigma^2 \\ (1-\rho_\tau)\,\tau^2 \end{pmatrix}, \epsilon_{t+1} \equiv \begin{pmatrix} \epsilon_{c,t+1} \\ \epsilon_{d,t+1} \\ \epsilon_{x,t+1} \\ \epsilon_{\sigma,t+1} \\ \epsilon_{\tau,t+1} \end{pmatrix},$$

and

$$F \equiv \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \phi & 0 & 0 \\ 0 & 0 & \rho_x & 0 & 0 \\ 0 & 0 & 0 & \rho_\sigma & 0 \\ 0 & 0 & 0 & 0 & \rho_\tau \end{pmatrix}, G_t \equiv \begin{pmatrix} \sigma_t & 0 & 0 & 0 & 0 \\ \rho_d \varphi_d \sigma_t & \sqrt{1-\rho_d^2}\varphi_d \sigma_t & 0 & 0 & 0 \\ 0 & 0 & \varphi_x \sigma_t & 0 & 0 \\ 0 & 0 & 0 & \varphi_\sigma & 0 \\ 0 & 0 & 0 & 0 & \varphi_\tau \sigma_t \end{pmatrix}.$$

Due to ambiguity, the agent has a different view on the conditional expectation of Y_{t+1} under a different dynamics: given a particular realization of $\widetilde{z}_t = z_t$,

$$\mathbb{E}_{\pi_{z,t}}[Y_{t+1}] = (\mu + FY_t) + \mu_t,$$

where we use $\mathbb{E}_{\pi_{z,t}}$ to denote the conditional expectation operator under the condition distribution fixing a particular realization of $\widetilde{z}_t$, and

$$\mu_t \equiv (0, 0, 0, \varphi_\sigma z_t, 0)'. \tag{9}$$

μ_t reflects the effect of ambiguity on the conditional mean of Y_{t+1} perceived by the agent. Note that because ambiguity only affects the volatility process σ_t, all elements of μ_t are zero except the second-to-last element. Recall that from Eq. (6), μ_t follows a normal distribution $\mu_t \sim \mathcal{N}(0, G_t \Omega_t G_t')$, where

$$\Omega_t \equiv diag\left(\left[0, 0, 0, \tau_t^2, 0\right]\right).$$

Although we focus on the above specification for the sake of simplicity, ambiguity could also be introduced with a more general structure. For example, we could allow for ambiguity in the process of the long-run risks (e.g., $\tau_{x,t}^2$) and other processes. We leave these extenions of the model for future research.

Smooth Ambiguity Aversion. As another key departure from Bansal and Yaron (2004), the representative agent has the recursive smooth ambiguity aversion preferences, proposed by Hayashi and Miao (2011) and Ju and Miao (2012) who generalize the model of Klibanoff et al. (2009).

Let $V_t(C)$ denote the continuation utility at date t. Following Ju and Miao (2012), assume that $V_t(C)$ satisfies the following recursive equation:

$$V_t(C) = \left[(1-\beta) C_t^{1-\rho} + \beta \left\{\mathcal{R}_t\left(V_{t+1}(C)\right)\right\}^{1-\rho}\right]^{\frac{1}{1-\rho}}, \tag{10}$$

$$\mathcal{R}_t\left(V_{t+1}(C)\right) = \left\{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1-\eta}{1-\gamma}}\right\}^{\frac{1}{1-\eta}}, \tag{11}$$

where $\mathcal{R}_t\left(V_{t+1}(C)\right)$ is an uncertainty aggregator that maps an s^{t+1}-measurable random variable $V_{t+1}(C)$ to an s^t-measurable random variable, and $\mathbb{E}_{\pi_{z,t}}$ denotes the conditional expectation operator that is applied to ϵ_{t+1} under the distribution $\mathcal{N}(z_t, \mathcal{I})$ given a particular realization z_t, and $\mathbb{E}_{\mu_t}$ denotes the conditional expectation operator that is applied to μ_t under the distribution $\mathcal{N}(0, G_t \Omega_t G_t')$. In addition, $\beta \in (0,1)$ represents the subjective discount factor, $1/\rho > 0$ represents the IES, $\gamma > 0$ represents the degree of risk aversion, and $\eta \geq \gamma$ represents the degree of ambiguity aversion.

The major advantage offered by the recursive smooth ambiguity aversion preferences is the three-way separation among ambiguity aversion, risk aversion, and the IES. The coefficient of ambiguity aversion η is tied to the curvature of the uncertainty aggregator $\mathcal{R}_t\left(V_{t+1}(C)\right)$ in Eq. (11). When $\eta = \gamma$, the smooth ambiguity preferences reduce to the Epstein-Zin preferences. In this case, the agent is ambiguity-neutral and computes the certainty equivalent value as $\left(\mathbb{E}_{\mu_t} \circ \mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1}{1-\gamma}}$.

When $\eta > \gamma$, the agent is ambiguity averse. The intuition is the following. At time t, each possible realization of $\widetilde{z}_t$ leads to a different conditional distribution (or "model") of Y_{t+1}, resulting in the certainty equivalent of expected continuation value, $\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1}{1-\gamma}}$.

Uncertainty resulting from $\widetilde{z}_t$ leads to ambiguity about the above certainty equivalent of expected continuous value which changes with the realization of $\widetilde{z}_t$. Aversion to such ambiguity implies that the aggregate certainty equivalent value $\mathcal{R}_t\left(V_{t+1}\left(C\right)\right)$ evaluated across all possible "models" is lower than it would be in an ambiguity-neutral case; that is,

$$\mathcal{R}_t\left(V_{t+1}\left(C\right)\right) < \left(\mathbb{E}_{\mu_t} \circ \mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}\left(C\right)\right]\right)^{\frac{1}{1-\gamma}},$$

which holds if and only if $\eta > \gamma$. Note that when $\eta > \gamma$, the compound conditional distributions for ϵ_{t+1} and μ_t cannot be reduced to a single distribution in (11). This irreducibility of compound distributions captures ambiguity aversion.

For recursive smooth ambiguity aversion preferences, Ju and Miao (2012) show that the pricing kernel, conditional on a particular model with conditional mean $\mu_t = \sigma_t G z_t$, is given by

$$M_{z_t,t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\rho-\gamma}\left(\frac{\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}. \tag{12}$$

The last term on the right-hand side of (12) arises from ambiguity aversion. When the agent is ambiguity neutral (i.e., $\eta = \gamma$), this term vanishes and the pricing kernel reduces to the same one as studied in Bansal and Yaron (2004). In the general case with ambiguity aversion (i.e., $\eta > \gamma$), this adjustment term exists, capturing the feature that an ambiguity averse agent puts a higher weight on the states in which his continuation value is lower. To prepare for the log-linearization analysis, we derive an equivalent expression for the pricing kernel below:

$$M_{z_t,t+1} = \underbrace{\beta^{\theta}\left(\frac{C_{t+1}}{C_t}\right)^{-\rho\theta} R_{c,t+1}^{-(1-\theta)}}_{\equiv M_{t+1}^{EZ}}\underbrace{\left(\beta^{\theta}\mathbb{E}_{\pi_{z,t}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\rho\theta} R_{c,t+1}^{\theta}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}_{\equiv M_{z_t,t}^{A}}, \tag{13}$$

where $R_{c,t+1}$ denotes the return on the consumption claim and $\theta \equiv \frac{1-\gamma}{1-\rho}$.

3 Asset Pricing Implications

We now provide a log-linearization analysis of the generalized long-run risks model with smooth ambiguity preferences. The log-linearized model is so tractable that we are able to derive closed-form expressions for the risk-free rate, the equity premium, and the variance premium. The closed-form expressions are very useful to bring out important asset pricing implications of the model.

Log-linearization. Following Campbell and Shiller (1988), we first log-linearize $R_{c,t+1}$ as follows

$$r_{c,t+1} = \log R_{c,t+1} = \kappa_0 + \kappa_1 v_{t+1} - v_t + \Delta c_{t+1}, \tag{14}$$

where $v_t \equiv \log(\mathcal{V}_t)$ denotes the logarithm of the wealth-to-consumption ratio $\mathcal{V}_t$. The coefficients $\kappa_1 = \frac{\exp(\overline{v})}{\exp(\overline{v})+1}$ and $\kappa_0 = \log(\exp(\overline{v}) + 1) - \kappa_1 \overline{v}$ are determined based on the long-run mean of the wealth-to-consumption rate $\overline{v}$.

The rest of this section will discuss the log-linearization results and the model's asset pricing implications. The discussion will focus on the return on the consumption claim for the sake of simplicity. The same argument also applies to the equity return on the dividend claim. We derive the results for the equity return in the appendix and use them for numerical analysis in the next section.

The model economy is characterized by the state variables x_t, σ_t^2, and τ_t^2. We conjecture that the logarithm of the wealth-to-consumption ratio v_t is affine in the state variables:

$$v_t = A_0 + A_x x_t + A_\sigma \sigma_t^2 + A_\tau \tau_t^2, \tag{15}$$

where the log-linearization coefficients A_0, A_x, A_σ, A_τ are determined via the Euler condition (see Appendix A.1 for the derivation):

$$A_0 = \frac{\log\beta + \kappa_0 + (1-\rho)\,\mu_c + \kappa_1 A_\sigma (1-\rho_\sigma)\,\sigma^2 + \kappa_1 A_\tau (1-\rho_\tau)\,\tau^2 + \frac{1}{2}\theta\,(\kappa_1 A_\sigma \varphi_\sigma)^2}{1-\kappa_1}, \tag{16a}$$

$$A_x = \frac{1-\rho}{1-\kappa_1\rho_x}, \tag{16b}$$

$$A_\sigma = \frac{1-\gamma}{1-\rho}\frac{(1-\rho)^2 + (\kappa_1 A_x \varphi_x)^2 + (\kappa_1 A_\tau \varphi_\tau)^2}{2(1-\kappa_1\rho_\sigma)}, \tag{16c}$$

$$A_\tau = \frac{1-\eta}{1-\rho}\frac{(\kappa_1 A_\sigma \varphi_\sigma)^2}{2(1-\kappa_1\rho_\tau)}. \tag{16d}$$

The expressions for A_x and A_σ are almost the same as in Bansal and Yaron (2004). When the IES is greater than one (i.e., $\rho < 1$), $A_x > 0$ implies that the wealth-to-consumption ratio increases in response to higher expected growth, because the intertemporal substitution effect dominates the wealth effect such that the agent buys more assets. If, in addition, the agent prefers early resolution of uncertainty (i.e., $\gamma > 1/\rho > 1$), then $A_\sigma < 0$, which implies that economic volatility has an adverse effect on the wealth-to-consumption ratio.

The coefficient of ambiguity A_τ depends directly on the ambiguity aversion parameter η. When the IES is greater than one (i.e., $\rho < 1$), the coefficient A_τ is negative if ambiguity aversion is large enough (i.e., $\eta > 1$), and becomes more negative as the agent becomes more ambiguity averse. That is, ambiguity aversion causes the wealth-to-consumption ratio to response more negatively to an increase in uncertainty, because a more ambiguity-averse agent puts more weight on bad models. In the homoskedastic-ambiguity case, an ambiguity-neutral agent (i.e., $\eta = \gamma$) simply treats ambiguity as an additional source of economic volatility. From the viewpoint of the ambiguity-neutral agent, the magnitude of economic volatility is effectively increased to $\varphi_\sigma^2 (1 + \tau^2)$.[3]

In the rest of the paper, we focus on the following empirically relevant parameter restriction:

$$\eta > \gamma > 1 > \rho. \tag{17}$$

Under the above restriction, the wealth-to-consumption ratio decreases in response to decreased expected growth and increased volatility or ambiguity.

Asset Return and Volatility. It is more convenient to work with the vector form: $v_t = A_0 + A' Y_t$, where $A = (A_c, A_d, A_x, A_\sigma, A_\tau)'$. As shown in our derivation, $A_c = A_d = 0$ since neither the consumption nor dividend growth rate affect on v_t. Based on the log-linearization result in Eq. (14), we can derive the innovation to the return on the consumption claim

$$r_{c,t+1} - \mathbb{E}_t \left[r_{c,t+1} \right] = \left(e_c + \kappa_1 A \right)' \left(Y_{t+1} - \mathbb{E}_t \left[Y_{t+1} \right] \right) \equiv B_r' G_t \epsilon_{t+1}, \tag{18}$$

where $e_c = (1, 0, 0, 0, 0)'$ denotes the selector vector that selects Δc_{t+1} from Y_{t+1} and $B_r \equiv e_c + \kappa_1 A \equiv (1, 0, B_{r,x}, B_{r,\sigma}, B_{r,\tau})'$ denotes the return innovation's sensitivities to various shocks. It follows that the conditional variance of $r_{c,t+1}$ is given by

$$\begin{aligned}
\Sigma_t &\equiv Var_t \left[r_{c,t+1} \right] = \mathbb{E}_t \left(B_r' G_t \epsilon_{t+1} \epsilon_{t+1}' G_t' B_r \right) = B_r' G_t \left(\mathcal{I} + \Omega_t \right) G_t' B_r \qquad (19) \\
&= \left[1 + \left(\kappa_1 A_x \varphi_x \right)^2 + \left(\kappa_1 A_\tau \varphi_\tau \right)^2 \right] \sigma_t^2 + \left(\kappa_1 A_\sigma \varphi_\sigma \right)^2 \left(1 + \tau_t^2 \right) \\
&\equiv B_{\Sigma,0} + B_{\Sigma,\sigma} \sigma_t^2 + B_{\Sigma,\tau} \tau_t^2 \equiv B_{\Sigma,0} + B_\Sigma' Y_t,
\end{aligned}$$

where $B_\Sigma \equiv (0, 0, 0, B_{\Sigma,\sigma}, B_{\Sigma,\tau})'$ denotes sensitivities with respect to various state variables. Note that the conditional variance Σ_t is driven by σ_t^2 and τ_t^2, with the sensitivities given by $B_{\Sigma,\sigma}$ and $B_{\Sigma,\tau}$. Therefore, the innovation to the conditional variance is given by

$$\begin{aligned}
\Sigma_{t+1} - \mathbb{E}_t \left[\Sigma_{t+1} \right] &= B_\Sigma' \left(Y_{t+1} - \mathbb{E}_t \left[Y_{t+1} \right] \right) = B_\Sigma' G_t \epsilon_{t+1} \qquad (20) \\
&= B_{\Sigma,\sigma} \varphi_\sigma \epsilon_{\sigma,t+1} + B_{\Sigma,\tau} \varphi_\tau \sigma_t \epsilon_{\tau,t+1}.
\end{aligned}$$

[3] In this case with constant ambiguity and ambiguity neutrality, the coefficient A_0 is replaced by $A_0 + A_\tau \tau^2$, given by $A_0 = \frac{\ln \beta + \kappa_0 + (1-\rho)\mu_c + \kappa_1 A_\sigma (1-\rho_\sigma)\sigma^2 + \frac{1}{2}\theta(\kappa_1 A_\sigma \varphi_\sigma)^2 \left(1+\tau^2\right)}{1-\kappa_1}$. That is, the parameter φ_σ^2 in the standard LRR model is now replaced by $\varphi_\sigma^2 \left(1 + \tau^2\right)$. Equivalently, this result implies that from the viewpoint of the ambiguity-neutral agent, The amount of model uncertainty τ^2 thus adds to the volatility of the economic uncertainty.

Pricing Kernel. We now log-linearize the pricing kernel in Eq. (13), which has two components: $M_{z_t,t+1} = M^{EZ}_{t+1} M^{A}_{z_t,t}$. The first component M^{EZ}_{t+1} is the pricing kernel under the Epstein-Zin preferences. As shown in Bansal and Yaron (2004), the innovation to the logarithm of this component (i.e., $m^{EZ}_{t+1} = \ln M^{EZ}_{t+1}$) is given by

$$m^{EZ}_{t+1} - \mathbb{E}_t\left[m^{EZ}_{t+1}\right] = -\Lambda'\left(Y_{t+1} - \mathbb{E}_t\left[Y_{t+1}\right]\right) = -\Lambda' G_t \epsilon_{t+1},$$

where $\Lambda = \gamma e_c - (\theta - 1)\kappa_1 A \equiv (\gamma, 0, \Lambda_x, \Lambda_\sigma, \Lambda_\tau)$ represents the prices of risk for the shocks ϵ_{t+1} for an *ambiguity-neutral* agent. It is worthwhile to point out that the price of the expected growth (long-run) risk, Λ_x, is given by $-(\theta - 1)\kappa_1 A_x = -(\theta - 1)\frac{\kappa_1(1-\rho)}{1-\kappa_1\rho_x}$, which increases with ρ_x if the agent prefers early resolution of uncertainty (i.e., $\gamma > \rho$). This is an important reason that the aversion to persistent (yet small) long-run consumption risk in the standard LRR model can generate a large equity risk premium.

The second component $M^{A}_{z_t,t}$ is an additional adjustment term in the pricing kernel, resulting from ambiguity aversion. It is straightforward to show that (see Appendix A.3)

$$\begin{aligned} m^{A}_{z_t,t} &= \log M^{A}_{z_t,t} = -\frac{\eta-\gamma}{1-\rho}\left(\log\beta + \kappa_0 + (\kappa_1 - 1)A_0 - A'Y_t + \Gamma'\left(\mu + FY_t + \mu_t\right) + \frac{1}{2}\theta\Gamma' G_t G_t'\Gamma\right) \\ &= \frac{\eta-\gamma}{1-\rho}\left[\frac{1}{2}\frac{1-\eta}{1-\rho}\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 \tau_t^2 - \left(\kappa_1 A_\sigma \varphi_\sigma\right) z_t\right]. \end{aligned}$$

When $\eta = \gamma$, the adjustment term $M^{A}_{z_t,t}$ becomes one and the pricing kernel reduces to the one under the Epstein-Zin preferences. Under the parameter restriction in Eq. (17), the coefficient of τ_t^2 in the above equation is negative, suggesting that higher ambiguity adversely impacts asset valuation in an ambiguity-averse world.

Risk Premiums. We now turn to the determination of the risk premiums, which are shown to have log-linearization approximations in closed form. Proposition 1 below reports the results.

Proposition 1. *Under the log-linearized generalized LRR model with smooth ambiguity preferences, the risk-free rate is given by*

$$\begin{aligned} r_{f,t} &= -\log\beta + \rho\left(\mu_c + x_t\right) - \frac{1}{2}\gamma^2\sigma_t^2 + \frac{1}{2}(1-\theta)^2(1-\rho)^2\sigma_t^2 \\ &\quad -\frac{1}{2}(1-\theta)\Gamma' G_t\left(\mathcal{I} + \Omega_t\right)G_t'\Gamma - \frac{1}{2}\frac{\eta-\gamma}{1-\rho}\Gamma' G_t \Omega_t G_t'\Gamma, \end{aligned} \tag{21}$$

where $\Gamma = (1-\rho)e_c + \kappa_1 A$. Furthermore, the equity and variance premiums for the return *on the consumption claim are given by*

$$
\begin{aligned}
EP_t &\equiv \log \mathbb{E}_t\left[R_{c,t+1}\right] - r_{f,t} \\
&= B_r' G_t \left(\mathcal{I} + \Omega_t\right) G_t' \Lambda + \frac{\eta - \gamma}{1 - \rho} B_r' G_t \Omega_t G_t' \Gamma,
\end{aligned}
\tag{22}
$$

and

$$
\begin{aligned}
VP_t &\equiv \mathbb{E}_t^Q\left[\Sigma_{t+1}\right] - \mathbb{E}_t\left[\Sigma_{t+1}\right] \\
&= -\left[B_\Sigma' G_t \left(\mathcal{I} + \Omega_t\right) G_t' \Lambda + \frac{\eta - \gamma}{1 - \rho} B_\Sigma' G_t \Omega_t G_t' \Gamma\right].
\end{aligned}
\tag{23}
$$

The results for the return on the dividend claim are very similar.

Proof. See Appendice A.3, A.5, and A.6. □

First, let us consider Case I of the generalized LRR model (i.e., model "gLRR1") where ambiguity is time invariant (i.e., τ_t^2 is equal to constant τ^2). If we hold ambiguity constant, fluctuation in economic volatility σ_t^2 alone are insufficient to generate the time-varying variance premium, although the equity premium will vary in time. In fact, we can see from Eq. (23) that the variance premium in this case is time invariant:

$$
VP_t|_{\tau_t^2 \equiv \tau^2} = -\frac{\gamma - \rho}{1 - \rho} B_{\Sigma,\sigma} \kappa_1 A_\sigma \varphi_\sigma^2 - \frac{\eta - \rho}{1 - \rho} B_{\Sigma,\sigma} \kappa_1 A_\sigma \varphi_\sigma^2 \tau^2.
$$

Note that consistent with the data, the model-implied variance premium is positive under the parameter restriction in Eq. (17). Furthermore, the magnitude of the variance premium increases with the ambiguity aversion coefficient η, the degree of ambiguity τ^2, the volatility of economic volatility φ_σ^2, as well as the sensitivity of the wealth-consumption ratio with respect to economic volatility A_σ. This result of a time-invariant variance premium in the case with constant ambiguity implies that the fluctuation in the variance premium observed in the data should be closely related to fluctuation in ambiguity.

Next, we turn to Case II of the generalized LRR model (i.e., model "gLRR2") where we introduce fluctuating ambiguity τ_t^2 in order to generate a time-varying variance premium. We start with the special ambiguity-neutral case with ambiguity neutrality (i.e., $\eta = \gamma$). In this case, based on Proposition 1, the risk-free rate, the equity premium, and the variance premium have the following expressions:

$$
\begin{aligned}
r_{f,t}|_{\eta=\gamma} &= -\log\beta + \rho(\mu_c + x_t) - \frac{1}{2}\gamma^2\sigma_t^2 + \frac{1}{2}(1-\theta)^2(1-\rho)^2\sigma_t^2 \\
&\quad -\frac{1}{2}(1-\theta)\Gamma' G_t(\mathcal{I}+\Omega_t)G_t'\Gamma,
\end{aligned}
$$

and

$$
\begin{aligned}
EP_t|_{\eta=\gamma} &= B_r' G_t(\mathcal{I}+\Omega_t)G_t'\Lambda \\
&= \left[\gamma + B_{r,x}\Lambda_x\varphi_x^2 + B_{r,\tau}\Lambda_\tau\varphi_\tau^2\right]\sigma_t^2 + B_{r,\sigma}\Lambda_\sigma\varphi_\sigma^2\left(1+\tau_t^2\right),
\end{aligned}
$$

and

$$
\begin{aligned}
VP_t|_{\eta=\gamma} &= -B_\Sigma' G_t(\mathcal{I}+\Omega_t)G_t'\Lambda \\
&= -B_{\Sigma,\tau}\Lambda_\tau\varphi_\tau^2\sigma_t^2 - B_{\Sigma,\sigma}\Lambda_\sigma\varphi_\sigma^2\left(1+\tau_t^2\right).
\end{aligned}
$$

The expressions of $r_{f,t}|_{\eta=\gamma}$ and $EP_t|_{\eta=\gamma}$ are almost identical to those in Bansal and Yaron (2004), except for the fact that the ambiguity-neutral agent perceives ambiguity simply as an additional source of economic volatility and thus considers both indistinguishable. From the expression of $VP_t|_{\eta=\gamma}$, the model under the Epstein-Zin preferences can still generate a positive variance premium, because $A_x > 0$ and $A_\sigma, A_\tau < 0$ under the parameter restriction in Eq. (17). However, a sizable variance premium will not obtain if neither volatility-of-volatility φ_σ nor ambiguity τ_t^2 is large enough. In fact, based on the estimated parameters in Bansal et al. (2016), the standard LRR model can only account for less than 10 percent of the magnitude of the variance premium in the data.

Amplification via Ambiguity Aversion. Introducing the recursive smooth ambiguity preferences allows for a three-way separation among ambiguity aversion, risk aversion, and the IES. Being able to separate out ambiguity aversion provides us with an additional degree of freedom with which to match the data, especially the variance premium.

Introducing ambiguity aversion can further lower the risk-free rate, and amplify the equity and variance premiums to get them closer to the data. Specifically, when $\eta > \gamma$, the representative agent is ambiguity averse and has a higher ambiguity-induced demand for savings, which tends to lower the risk-free rate:

$$
r_{f,t} - r_{f,t}|_{\eta=\gamma} = -\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\Gamma' G_t\Omega_t G_t'\Gamma = -\frac{1}{2}\frac{\eta-\gamma}{1-\rho}(\kappa_1 A_\sigma\varphi_\sigma)^2\tau_t^2,
$$

which is negative under the parameter restriction in Eq. (17). Furthermore, the agent demands additional compensation, which is the last term in Eq. (22) or Eq. (23), given by

$$EP_t - EP_t|_{\eta=\gamma} = \frac{\eta-\gamma}{1-\rho} B_r' G_t \Omega_t G_t' \Gamma = \frac{\eta-\gamma}{1-\rho} \left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 \tau_t^2,$$

and

$$\begin{aligned} VP_t - VP_t|_{\eta=\gamma} &= -\frac{\eta-\gamma}{1-\rho} B_\Sigma' G_t \Omega_t G_t' \Gamma = -\frac{\eta-\gamma}{1-\rho} B_{\Sigma,\sigma} \left(\kappa_1 A_\sigma\right) \varphi_\sigma \tau_t^2 \\ &= -\frac{\eta-\gamma}{1-\rho} \kappa_1 A_\sigma \left(1 + \left(\kappa_1 A_x \varphi_x\right)^2 + \left(\kappa_1 A_\tau \varphi_\tau\right)^2\right) \varphi_\sigma^2 \tau_t^2. \end{aligned}$$

We refer to the additional compensation demanded by an ambiguity-averse agent relative to an ambiguity-neutral agent as an "*ambiguity premium*." Under the parameter restriction in Eq. (17), the ambiguity premium is positive, and, importantly, increases with the coefficient of ambiguity aversion η. Later we will calibrate the model to quantitatively examine how much of the equity or variance premium can be accounted for by the ambiguity premium.

Why is the Ambiguity Premium Positive in the Model? This is because the ambiguity-averse agent's beliefs are distorted more toward bad states. To understand this intuition, let us take the variance premium as an example. From the proof of Proposition 1, the variance premium can be rewritten as

$$\begin{aligned} VP_t &= \mathbb{E}_t^Q \left[\Sigma_{t+1}\right] - \mathbb{E}_t \left[\Sigma_{t+1}\right] \\ &= \mathbb{E}_{\mu_t} \left[\underbrace{\frac{\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ}\right] M_{z_t,t}^A}{\mathbb{E}_{\mu_t} \left(\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ}\right] M_{z_t,t}^A\right)}}_{\text{ambiguity-induced distortion } d(\mu_t)} \times \underbrace{\frac{\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ} \left(\Sigma_{t+1} - \mathbb{E}_t \left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ}\right]}}_{\text{VP under a specific model } VP_t^{EZ}(\mu_t)} \right] \\ &\equiv \mathbb{E}_{\mu_t} \left[d\left(\mu_t\right) VP_t^{EZ}\left(\mu_t\right)\right]. \end{aligned} \tag{24}$$

This expression for the variance premium has a very intuitive interpretation: it is roughly a weighted average of the variance premiums across all possible models. More precisely, it is the average of the weighted (or distorted) variance premiums within each model, taken across all possible models under the objective probability measure. The variance premium within a specific model (i.e., a fixed μ_t), $VP_t^{EZ}\left(\mu_t\right)$, can be shown as

$$VP_t^{EZ}\left(\mu_t\right) \equiv \frac{\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ} \left(\Sigma_{t+1} - \mathbb{E}_t \left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_{\pi_{z,t}} \left[M_{t+1}^{EZ}\right]} = -B_\Sigma' G_t G_t' \Lambda + B_\Sigma' \mu_t.$$

Due to ambiguity around μ_t, the ambiguity-averse agent weights the model-specific variance premium differently, using a weighting scheme different from the physical probability. The weighting scheme tilts more toward bad states. Specifically, the weight or "distortion" due to ambiguity aversion can be derived as:

$$d\left(\mu_t\right) \equiv \frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^{A}}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^{A}\right)} = \frac{\exp\left(-\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)' \mu_t\right)}{\mathbb{E}_{\mu_t}\left[\exp\left(-\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)' \mu_t\right)\right]}.$$

The presence of the distortion factor is equivalent to a change of the probability measure, and the variance premium demanded by the ambiguity-averse agent is the conditional expectation under the new probability measure. In fact,

$$VP_t = \mathbb{E}_{\mu_t}\left(d\left(\mu_t\right) VP_t^{EZ}\left(\mu_t\right)\right) = \widehat{\mathbb{E}}_{\mu_t}\left(VP_t^{EZ}\left(\mu_t\right)\right) \tag{25}$$

where the expectation operator $\widehat{\mathbb{E}}_{\mu_t}$ operates under the new probability measure under which μ_t follows the distribution of $N\left(-\Omega_t\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right), \Omega_t\right)$, instead of the distribution of $N\left(0, \Omega_t\right)$ under the original physical measure. In other words, because of ambiguity aversion ($\eta > \gamma$), the agent tilts the mean of μ_t by $-\frac{\eta-\gamma}{1-\rho}\Omega_t\Gamma = -\frac{\eta-\gamma}{1-\rho}\left(0,0,0,\kappa_1 A_\sigma \tau_t^2\right)'$. Under the parameter restriction in Eq. (17), $A_\sigma < 0$ and the agent facing ambiguity believes that economic volatility is likely to increase (i.e., $-\frac{\eta-\gamma}{1-\rho}\kappa_1 A_\sigma \tau_t^2 > 0$). As a result, the agent demands a higher variance premium (along with a higher equity premium) relative to an ambiguity-neutral agent.

4 Main Findings

In this section, we demonstrate the generalized LRR model's flexibility in terms of its ability to simultaneously match key moments of asset prices, particularly those of the variance premium. This demonstration uses a version of the model calibrated at an annual frequency using annual data between 1931 and 2009.

4.1 Data

Annual consumption data are obtained from the Bureau of Economic Analysis's NIPA Table 7.1. The real quarterly consumption growth rate is the real per-capita personal consumption expenditure on nondurable goods and services over the quarter divided by the per-capita personal consumption expenditure on nondurable goods and services over the previous year.

Inflation data are obtained from the Bureau of Labor Statistics, particularly the monthly seasonal-adjusted consumer price index (CPI) for all urban consumers. The annual inflation rate is then constructed as the log growth rate of the CPI in the final month of the current year over the final month in the previous year.

Annual stock market data are obtained from the CRSP, particularly the returns on the value-weighted index including distributions (VWRETD) and the value-weighted index excluding distributions (VWRETX). The indices consist of all stocks on the NYSE, AMEX, and NASDAQ, from which we construct the annual stock price index level and annual dividend. The annual dividend is the sum of the dividends paid over the course of the year. We then calculate the log year-over-year growth rate in dividends. The annual stock market return is the sum of the stock market index at the end of the year and the annual dividend, divided by the stock market index at the beginning of the year. We then obtain the annual real stock market return by deflating the nominal return by the growth rate of the CPI. The annual price-dividend ratio is constructed as the stock market index at the end of the year divided by the sum of the dividends paid over the previous twelve months.

We obtain the ex ante real risk-free rate following Beeler and Campbell (2012). Specifically, we obtain the three-month yield from the CRSP Fama Risk Free Rates data and then calculate the ex post real risk-free rate by subtracting the log inflation rate from the nominal log yield for the same three-month period. The ex ante real risk-free rate is then obtained as the fitted value from the regression of the ex post real risk-free rate on the three-month nominal yield and the realized growth rate of the CPI over the previous twelve months.

Historical variance premium data are available only after 1990 when we are able to compute implied volatility from options data following the mothod of Drechsler (2013). The variance premium is formally defined as the difference between the risk-neutral expectation $\mathbb{E}_t^{\mathbb{Q}}(\cdot)$ and the objective expectation $\mathbb{E}_t(\cdot)$ of the return variance Σ_{t+1}; that is, $VP_t \equiv \mathbb{E}_t^{\mathbb{Q}}(\Sigma_{t+1}) - \mathbb{E}_t(\Sigma_{t+1})$, where $\mathbb{Q}$ represents the risk-neutral measure. The availability of the Chicago Board Options Exchange (CBOE) VIX index makes it straightforward to measure the risk-neutral expectation of stock market return variances.[4] We compute $\mathbb{E}_t^{\mathbb{Q}}(\Sigma_{t+1})$ by squaring the CBOE VIX index and then dividing it by 12 to get a monthly quantity. We estimate $\mathbb{E}_t(\Sigma_{t+1})$ as the conditional forecast of the realized variance in the following month. Following Drechsler (2013), we measure the realized variance of the returns on the S&P 500 index by summing up the squared five-minute log returns on the S&P 500 futures and on the S&P 500 index over the whole month. We obtain the high-frequency data used to construct these realized variance measures from TICKDATA. Next we use a simple time-series model

[4] The CBOE VIX index is based on the highly liquid S&P 500 index options along with the "model-free" approach explicitly tailored to replicate the risk-neutral variance of a fixed 30-day maturity. See, e.g., Carr and Wu (2009) for the definition of model-free implied variance.

to obtain one-step-ahead forecasts of the realized variance in the following month. Specifically, we regress the futures realized variance on the value of the squared VIX and on a lagged cash realized variance. The estimates of the variance premium are truncated to zero whenever negative values appear since theoretically the physical measure of variance should be less than the risk-neutral measure.

4.2 Calibration

Bansal and Yaron (2004) assume a monthly decision interval of the agent and calibrate the standard long-run risks model at the monthly frequency to match the time-aggregated annual data. Bansal et al. (2016) estimate the standard model using a GMM framework that explicitly accounts for time-aggregation and treats the decision interval of the agent as an additional parameter to estimate. Their estimate of the decision frequency corresponds to a decision interval of about 33 days, or roughly one month. Furthermore, they show that if the decision interval is fixed to one month, the parameter estimates and pricing implications are very similar to those of the unrestricted model. At monthly frequency, the model fits the data reasonably well.

For the above reasons, in this paper we use as the benchmark parameter specification the estimated parameter values under the assumption of a monthly decision interval from Bansal et al. (2016) (see Column "Monthly" in Table VI in their paper). Panel A of Table 1 reports the parameter values in the benchmark specification.

To illustrate the ability of the generalized LRR model to account for moments of the variance premium along with the other key asset pricing moments, we recalibrate the parameters of risk aversion, the IES, and ambiguity aversion (i.e., γ, ρ, and η) to match the levels of the equity premium, the risk-free rate, and the variance premium. Note that we do not require the agent be ambiguity averse in calibration. [5] Furthermore, in Case II of the generalized LRR model (i.e., model "gLRR2"), we can further calibrate the parameter φ_τ to match the volatility of the variance premium. As we will show in the following subsection, the calibrated generalized LRR model does a reasonably good job of fitting all these moments with risk aversion below 5.

4.3 Main Findings

We first recapitulate the success of the standard LRR model in matching the moments of the equity premium and the risk-free rate before highlighting a few challenges. Later in this

[5] See Makarov (2019) on ambiguity-seeking behavior in models with smooth ambiguity preferences.

section, we will show how these challenges can be addressed in the generalized LRR model.

Asset Pricing Moments under the Standard LRR Model. We consider a broad set of asset pricing moments, including the mean and volatility of the variance premium, the market return, the risk-free rate, and the price-dividend ratio.

In Panel B of Table 1 we display the asset pricing moments in the data and those implied by the standard LRR model based on the benchmark parameter specification. The moments are very similar to those in Bansal et al. (2016) (see Columns "Sample" and "Model" under "LRR Model" in Table III of their paper), suggesting that the data constructed in this paper closely matches theirs and the benchmark parameter specification is close enough to the specification based on the estimated decision interval in their paper (estimated to be about 33 days, close to a month). Overall, the standard LRR model under the benchmark parameter specification does a reasonably good job at matching most of the asset pricing moments. As argued in Bansal and Yaron (2004), this quantitative success largely comes from the ability of the Epstein-Zin preferences to separate the IES from risk aversion. As a result of this two-way separation, the IES is not tied to the reciprocal of risk aversion and thus can be possibly larger than 1. Therefore, when both risk aversion and the IES take relatively large values, the LRR model can generate a relatively large equity premium and a low risk-free rate. The existence of persistent long-run risks can further improve the fit of the model.

Despite its overall good performance, the standard LRR model under the benchmark parameter specification matches about 75% of the equity premium and nearly triples the risk-free rate seen in the data. To facilitate model comparisons, we recalibrate the parameters of risk aversion and the IES only to *exactly* match both the equity premium and the risk-free rate, keeping all the other parameter values unchanged. The recalibration results are reported in Column "LRR" in Table 2. Intuitively, one can see that both parameters need to be increased to match the two key moments. Specifically, the recalibrated value of risk aversion increases from 7.13 to 8.548, and the IES increases from 2.08 to 2.97 (i.e., its reciprocal, the parameter ρ, decreases from 0.481 to 0.336). Under the recalibrated parameter values, the model-implied moments remain similar and match the moments in the data reasonably well except that the price-dividend ratio has a somewhat lower mean.

It is important to point out that under either the benchmark parameter specification or the recalibrated one, the standard LRR model implies a very small variance premium compared to the data. It accounts for only about 10%-15% of the average variance premium in the data. Furthermore, although the standard LRR model is able to generate the time-varying equity premium, the variance premium remains constant. As a result, its volatility implied by the model is zero, and thus has no power of predicting the stock market return.

Next we turn to the generalized LRR model and discuss its asset pricing implications. We will highlight how the model's flexibility helps to improve the model's fit to the data and in tackling the above challenges.

Asset Pricing Moments under the Generalized LRR Model. Much like the two-way separation between risk aversion and the IES behind the success of the standard LRR model, the three-way separation in the generalized LRR model that further separates ambiguity aversion from risk aversion and the IES further improves the model's quantitative performance. In particular, ambiguity aversion as the additional degree of freedom is very helpful for the model to match the average level of the variance premium. The last two columns of Table 2 (i.e., Columns "gLRR1" and "gLRR2") display various asset pricing moments implied by Case I and Case II of the generalized LRR model (with constant and time-varying ambiguity), respectively.

Consider Case I of the generalized LRR model first. The parameter τ governs the degree of ambiguity in the model. It increases the volatility of the underlying economic risks because $Var\left(\sigma_t^2\right) = \varphi_\sigma^2\left(1+\tau^2\right)/\left(1-\rho_\sigma^2\right)$. We calibrate the parameter τ to be $\sqrt{2.5^2-1} = 2.29$; that is, compared to the ambiguity-free environment in the standard LRR model, the volatility of economic volatility σ_t^2 increases by 150%. The parameters of γ, ρ, and η are recalibrated to match the equity premium, the risk-free rate, and the variance premium simultaneously. The recalibration results are reported in Column "gLRR1". The parameter of risk aversion is recalibrated to be 4.70, which is substantially lower than the value of 8.55 in the standard LRR model. The IES is recalibrated to a lower value of 2.08 as well. The parameter of ambiguity aversion is calibrated to be around 32.5, which is broadly consistent with estimates of ambiguity aversion in the literature. For example, the value of ambiguity aversion is calibrated to be 8.86 in Ju and Miao (2012) and 10.38 in Miao et al. (2019). Gallant et al. (2018) estimate the same asset pricing model in these two papers using Bayesian methods and estimate ambiguity aversion at around 30 for their annual data sample.

Most importantly, the calibrated generalized LRR model with constant ambiguity is able to match the variance premium. In fact, introducing smooth ambiguity preferences to the standard LRR model increases the variance premium from 1.107 to 8.285 as in the data (see Columns "LRR" and "gLRR1" of Table 2), an almost eightfold increase. As discussed in Section 3, the model can generate a sizable variance premium because the ambiguity averse agent has a *distorted* belief that tilts more toward bad states and thus demands a higher variance premium.

However, the variance premium is constant in Case I of the generalized LRR model and thus it cannot account for the volatile variance premium in the data. The generalized LRR model with time-varying ambiguity (i.e., Case II) is able to tackle these challenges. The last Column "gLRR2" in Table 2 reports the calibration results. Relative to Case I, there are two additional parameters to calibrate: ρ_τ and φ_τ. These parameters determine the persistence and volatility of the time-varying ambiguity. We set ρ_τ to be 0.985, implying a very persistent ambiguity process. The parameter φ_τ is then calibrated to make the model match the volatility of the variance premium in the data as closely as possible. For this purpose, we calibrate φ_τ to be 25 and the model-implied volatility of the variance premium is 1.498, or almost 25% of the observed volatility of the variance premium in the data. Similarly as before, the parameters γ, ρ, and η are calibrated to match the equity premium, the risk-free rate, and the variance premium. The calibrated values are similar to those in Case I of the generalized LRR model (see Column "gLRR1").

Ambiguity Aversion vs. Risk Aversion. As we have shown so far, the aversion to both long-run risks and ambiguity have distinct contributions to the equity premium and the variance premium. Their contributions are referred to as the *risk-premium* and *ambiguity-premium* components of a given asset pricing moment. To quantify the relative contributions of the risk and ambiguity channels, we calculate the equity premium and the variance premium implied by Case II of the generalized LRR model (i.e., model "gLRR2") when the agent is either ambiguity averse or ambiguity neutral. These results are reported in Table 3.

As shown in Table 3, Column "Ambiguity-Averse" reports the model-implied equity and variance premiums using the calibrated parameter values in Column "gLRR2" in Table 2 for an ambiguity-averse agent with $\eta = 30.807$ and $\gamma = 4.576$. By construction, the calibrated model is able to generate exactly the same levels of the equity and variance premiums in the data, or 7.7% and 8.285, respectively. To isolate the contribution of the ambiguity channel, we set ambiguity aversion and risk aversion to be equal and compute the model-implied equity and variance premiums for an ambiguity-neutral agent (i.e., $\eta = \gamma = 4.576$). As shown in Column "Ambiguity-Neutral" in Table 3, the equity and variance premiums drop to 4.6% and 1.921, respectively, in this ambiguity-neutral case. The difference in the equity and variance premiums between these two cases represents the *ambiguity-premium component*. The results suggest that the ambiguity-premium component accounts for about 77% of the variance premium, but 40.3% of the equity premium. Our findings indicate that a large portion of the variance premium is attributable to the ambiguity aversion channel, and that ambiguity aversion has a much larger effect on the variance premium than on the equity premium.

Ambiguity vs. Risk. We are able to make inferences about long-run risks and ambiguity via particle filtering. As discussed above, we assume a monthly decision interval and use annual data in calibration following Bansal et al. (2016). We then compute model-implied unconditional moments that account for time aggregation (see Appendix B.1) for the log annual consumption and dividend growth rates, the log annual stock market return and risk-free rate, the annual price-dividend ratio, and the annual variance premium. In Appendix B.2, we derive a state-space system for the observables and the unobserved state assuming measurement errors. The measurement equations in the state-space system follow the measurement error model in Schorfheide et al. (2018) with measurement errors fixed at 1% of the sample variance of various asset pricing moments. Because of the presence of stochastic volatility and ambiguity, the state-space system is nonlinear and we thus use a particle filter to back out the latent variables following Schorfheide et al. (2018) (see Appendix B.3).

Figure 1 plots time series of the filtered latent variables with solid lines for the generalized LRR model ("gLRR2") and dashed lines for the standard LRR model. Panel A of Figure 1 depicts the estimate of the expected growth component x_t. As argued in Schorfheide et al. (2018), this expected growth component is extracted largely based on cash-flow data on consumption and dividend. In fact, our estimate of this long-run risk component is similar to theirs, except for the 1930s. For example, x_t tends to fall in recessions after mid 1950s. At the same time, our estimate of x_t is also influenced to some extent by asset prices, particularly the risk-free rate. Schorfheide et al. (2018) introduce exogenous preference shocks in order for the model-implied risk-free rate to track closely the risk-free rate in the data. They show that with no added preference shocks, the model's fit to the risk-free rate data deteriorates substantially. In this paper, we try to stay as closely as possible with the Bansal-Yaron model for the purpose of illustrating the novel ambiguity aversion channel, and we avoid incorporating preference shocks for this reason. The relatively small measurement errors imply that asset prices influence our estimates of the latent variables to a larger extent than Schorfheide et al. (2018). This difference explains the divergence between our estimate of x_t and theirs during the period in 1930s.

The filtered volatility and ambiguity processes are plotted in Panels B and C of Figure 1. Both processes exhibit substantial fluctuations over the sample period and tend to increase during recessions. The volatility process most closely resembles the volatility process of the expected consumption growth x_t in Schorfheide et al. (2018) who allow for different volatility processes for the innovations to consumption growth, dividend growth, and the expected consumption growth component. Due to ambiguity, the filtered economic volatility under the generalized LRR model ("gLRR2") is somewhat more volatile than that under the standard LRR model.

Under our generalized LRR model, we are able to infer about the ambiguity process τ_t

from cash-flow and asset price data. From Panel C of Figure 1, we can see that in the post-1990 period, the estimate of ambiguity τ_t is extracted largely based on the variance premium data when it is available. It is still interesting to see that we can make inferences about ambiguity even before 1990 using other asset prices. As shown in Panel C of Figure 1, both economic volatility (σ_t) and ambiguity (τ_t) spiked during the Great Depression and the Second World War.

Figure 2 plots filtered time series of the variance premium between 1931 and 2009 with solid lines for the generalized LRR model and dashed lines for the standard LRR model. It shows that the generalized LRR model exactly matches the average value of the variance premium in the data (i.e., 8.285) and generates its dynamics reasonably well. For example, the variance premium in the data jumps to around 20 in 1998 (the Russian default crisis) and 2009 (the financial crisis). The model-implied variance premium also spikes in these years. In contrast, the variance premium implied in the standard LRR model is constant and very small, around 1.1. By the variance decomposition described shortly below, we show that the dynamics of the model-implied variance premium are driven almost entirely by variation in ambiguity. This result is consistent with Drechsler (2013) who demonstrates with a robust control framework that variation in ambiguity generates variance premium fluctuations.

Not only can the generalized LRR model reproduce the dynamics of the variance premium in the data, it also does a reasonably good job of tracking closely other key asset prices. In Figure 3 we plot filtered time series of the stock market return, the risk-free rate, and the price-dividend ratio between 1931 and 2009 with solid lines for the generalized LRR model and dashed lines for the standard LRR model. As shown in Figure 3, both models do a reasonably good job of tracking the data series closely. The model-implied risk-free rate is smoother than the data, which is consistent with Schorfheide et al. (2018) and can be addressed by following their approach of introducing preference shocks.

Figure 4 plots the filtered expected consumption growth x_t together with the consumption growth rate (Panel A) and the dividend growth rate (Panel B). We can see that the expected consumption growth x_t is estimated to be small and persistent, and, more importantly, it is an important driver of both consumption and dividend growth rates. Based on our particle-filter estimation results, we find that the expected consumption growth has significant predictive power for both consumption and dividend growth rates after the mid-1930s.

The state-space approach allows us to conduct variance decomposition of several key asset prices. Table 4 reports the contribution of the fluctuations in growth prospects, x_t, economic volatility, σ_t^2, and ambiguity, τ_t^2, to the volatilities of the stock market return, the risk-free rate, the price-dividend ratio, and the variance premium for both the standard LRR model (see Column "LRR") and the generalized LRR model (see Column "gLRR2").

As shown in Table 4, almost all time variation in the variance premium is driven by time variation in ambiguity in the generalized LRR model. At the same time, fluctuations in ambiguity contribute little to the variance of the price-dividend ratio. The variabilities of the stock market return and the risk-free rate are, in almost equal parts, attributable to variation in economic volatility and ambiguity.

In summary, our results demonstrate that *ambiguity aversion* is important in determining the *level* of the variance premium, while fluctuating *ambiguity* is important in driving the *dynamics* of the variance premium. This finding is reminiscent of the well-known result in standard asset pricing models that *risk aversion* is the key determinant of the level of the equity premium, and *risk* (or fluctuating volatility) is the key driver of variability in the equity premium.

5 Conclusion

In this paper we generalize the long-run risks model in Bansal and Yaron (2004) by incorporating the recursive smooth ambiguity aversion preferences of Klibanoff et al. (2005, 2009) and time-varying ambiguity in an innovative way. The generalized LRR model remains as tractable as the Bansal-Yaron model, and gains an additional degree of freedom due to the separation of ambiguity aversion from risk aversion and the IES. A new ambiguity aversion channel arises as the ambiguity averse agent holds a pessimistic view tilted toward worse states and demands larger compensations. The generalized LRR model is thus more flexible in matching many asset pricing moments in the data, such as, the average level and volatility of the equity premium, the risk-free rate, and the variance premium, as well as the return predictability of dividend yield and the variance premium.

The calibrated model implies that the ambiguity aversion channel plays an important role in generating sizable equity and variance premiums. This channel accounts for about 77 percent of the variance premium and only about 40 percent of the equity premium. Our variance decomposition results further suggest that almost all variation in the variance premium is driven by variation in ambiguity.

Our generalized LRR framework is very flexible such that ambiguity can be introduced in a very general way. For example, ambiguity can be introduced to the long-run risks component or the consumption growth volatility component, or both. To highlight the quantitative performance of the generalized model in matching the variance premium, we choose to focus on ambiguity in the volatility component for the sake of simplicity in this paper. The analysis can be easily extended to more than one ambiguous state variables without sacrificing tractability; we leave this work for future research.

Although the main focus of this paper is to use calibrated model to illustrate its asset pricing implications, as an interesting research direction the particle-filter-based Bayesian method can be used to estimate the complete generalized LRR model so that we can estimate key structural parameters as well as the latent variables (see Schorfheide et al., 2018). Another possible future research direction is to introduce jumps similarly as in Drechsler and Yaron (2011) and allow for ambiguity about the jump process. These possible extensions may improve the quantitative performance of the model and provide further insight about asset pricing.

References

Ai, H. (2010): "Information Quality and Long-Run Risk: Asset Pricing Implications," *Journal of Finance*, 65, 1333–1367.

Ai, H. and R. Bansal (2018): "Risk Preferences and the Macroeconomic Announcement Premium," *Econometrica*, 86, 1383–1430.

Anderson, E. W., L. P. Hansen, and T. J. Sargent (2003): "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection," *Journal of the European Economic Association*, 1, 68–123.

Bakshi, G., N. Kapadia, and D. Madan (2015): "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *Review of Financial Studies*, 16, 101–143.

Bansal, R., R. Dittmar, and D. Kiku (2009): "Cointegration and Consumption Risks in Asset Returns," *Review of Financial Studies*, 22, 1343–1375.

Bansal, R., R. F. Dittmar, and C. T. Lundblad (2005): "Consumption, Dividends, and the Cross Section of Equity Returns," *Journal of Finance*, 60, 1639–1672.

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

Bansal, R. and I. Shaliastovich (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22, 4463–4492.

Branger, N., C. Schlag, and J. Thimme (2016): "Does Ambiguity About Volatility Matter Empirically?" Working Paper.

Cagetti, M., L. P. Hansen, T. Sargent, and N. Williams (2015): "Robustness and Pricing with Uncertain Growth," *Review of Financial Studies*, 15, 363–404.

Campbell, J. Y. and R. J. Shiller (1988): "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Cao, H. H., T. Wang, and H. H. Zhang (2005): "Model Uncertainty, Limited Market Participation, and Asset Prices," *Review of Financial Studies*, 18, 1219–1251.

Carr, P. and L. Wu (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

Chen, H. (2010): "Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure," *The Journal of Finance*, 65, 2171–2212.

Chen, Z. and L. Epstein (2002): "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

Colacito, R. and M. M. Croce (2011): "Risks for the Long Run and the Real Exchange Rate," *Journal of Political Economy*, 119, 153–181.

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon (2018): "Ambiguity and the Historical Equity Premium," *Quantitative Economics*, 9, 945–993.

Drechsler, I. (2013): "Uncertainty, Time-Varying Fear, and Asset Prices," *Journal of Finance*, 68, 1843–1889.

Drechsler, I. and A. Yaron (2011): "What's Vol Got to Do with It," *Review of Financial Studies*, 24, 1–45.

Epstein, L. G. and M. Schneider (2008): "Ambiguity, Information Quality, and Asset Pricing," *Journal of Finance*, 63, 197–228.

Epstein, L. G. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, 283–322.

Eraker, B. and I. Shaliastovich (2008): "An Equilibrium Guide to Designing Affine Pricing Models," *Mathematical Finance*, 18, 519–543.

Gallant, A. R., M. R Jahan-Parvar, and H. Liu (2018): "Does Smooth Ambiguity Matter for Asset Pricing?" *Review of Financial Studies*, 32, 3617–3666.

Garlappi, L., R. Uppal, and T. Wang (2007): "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *Review of Financial Studies*, 20, 41–81.

Hansen, L. P. (2007): "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," *American Economic Review*, 97, 1–30.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116, 260–302.

Hansen, L. P. and T. J. Sargent (1999): "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 873–907.

——— (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

——— (2007): "Recursive Robust Estimation and Control without Commitment," *Journal of Economic Theory*, 136, 1–27.

——— (2008): *Robustness*, Princeton: Princeton University Press.

Hayashi, T. and J. Miao (2011): "Intertemporal substitution and recursive smooth ambiguity preferences," *Theoretical Economics*, 6, 423–472.

Herbst, E. P. and F. Schorfheide (2015): *Bayesian Estimation of DSGE Models*, Princeton and Oxford: Princeton University Press.

Jahan-Parvar, M. R. and H. Liu (2014): "Ambiguity Aversion and Asset Prices in Production Economies," *Review of Financial Studies*, 27, 3060–3097.

Ju, N. and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80, 559–591.

Keynes, J. M. (1936): *The General Theory of Employment, Interest, and Money*, London: Palgrave Macmillan.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892.

——— (2009): "Recursive smooth ambiguity preferences," *Journal of Economic Theory*, 144, 930–976.

Knight, F. (1921): *Risk, Uncertainty, and Profit*, Boston: Houghton Mifflin.

Leippold, M., F. Trojani, and P. Vanini (2008): "Learning and Asset Prices Under Ambiguous Information," *Review of Financial Studies*, 21, 2565–2597.

Liu, J., J. Pan, and T. Wang (2005): "An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks," *Review of Financial Studies*, 18, 131–164.

Maenhout, P. J. (2004): "Robust Portfolio Rules and Asset Pricing," *Review of Financial Studies*, 17, 951–983.

Makarov, D. (2019): "On Ambiguity-Seeking Behavior in Finance Models with Smooth Ambiguity," Working Paper.

Mehra, R. and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Miao, J., B. Wei, and H. Zhou (2019): "Ambiguity Aversion and the Variance Premium," *Quarterly Journal of Finance*, 09, 1–36.

Piazzesi, M. and M. Schneider (2007): "Equilibrium Yield Curves," in *NBER Macroeconomics Annual*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, Cambridge, MA: MIT Press, 389–442.

Routledge, B. R. and S. E. Zin (2009): "Model Uncertainty and Liquidity," *Review of Economic Dynamics*, 12, 543–566.

Schorfheide, F., D. Song, and A. Yaron (2018): “Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach,” *Econometrica*, 86, 617–654.

Shi, Z. (2019): “Time-varying Ambiguity, Credit Spreads, and the Levered Equity Premium,” *Journal of Financial Economics*, 134, 617–646.

Todorov, V. (2009): “Variance Risk-Premium Dynamics: The Role of Jumps,” *Review of Financial Studies*, 23, 345–383.

Uppal, R. and T. Wang (2003): “Model Misspecification and Underdiversification,” *Journal of Finance*, 58, 2465–2486.

Weil, P. (1989): “The Equity Premium Puzzle and the Risk-free Rate Puzzle,” *Journal of Monetary Economics*, 24, 401–421.

A Log-linearization

In this appendix, we provide details in our log-linearization analysis.

A.1 Return on the Consumption Claim $r_{c,t+1}$

It is more convenient to work with the vector form and express the log wealth-consumption ratio as

$$v_t = A_0 + A'Y_t,$$

where $A \equiv (A_c, A_d, A_x, A_\sigma, A_\tau)'$. As we will see shortly, $A_c = A_d = 0$ since the consumption or dividend growth rate has no effect on v_t.

Recall that from Eq. (13), the pricing kernel can also be re-expressed as

$$\begin{aligned} M_{z_t,t+1} &= \underbrace{\beta^{\frac{1-\gamma}{1-\rho}} \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{1-\rho}} R_{c,t+1}^{\frac{\rho-\gamma}{1-\rho}}}_{M_{t+1}^{EZ}} \underbrace{\left(\beta^{\frac{1-\gamma}{1-\rho}} \mathbb{E}_{\pi_{z,t}} \left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{1-\rho}} R_{c,t+1}^{\frac{1-\gamma}{1-\rho}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}_{M_{z_t,t}^{A}} \\ &= M_{t+1}^{EZ} \cdot M_{z_t,t}^{A} \end{aligned}$$

Furthermore,

$$\begin{aligned} & \mathbb{E}_{\pi_{z,t}} \left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{1-\rho}} R_{c,t+1}^{\frac{1-\gamma}{1-\rho}}\right] \\ = \; & \mathbb{E}_{\pi_{z,t}} \left[\exp\left(-\rho\theta\Delta c_{t+1} + \theta r_{c,t+1}\right)\right] \\ = \; & \mathbb{E}_{\pi_{z,t}} \left[\exp\left(\left(1-\rho\right)\theta\Delta c_{t+1} + \theta\left(\kappa_0 + \kappa_1 v_{t+1} - v_t\right)\right)\right] \\ = \; & \exp\left(\theta\left(\kappa_0 + \left(\kappa_1 - 1\right)A_0 - A'Y_t\right)\right) \mathbb{E}_{\pi_{z,t}} \left[\exp\left(\theta\Gamma' Y_{t+1}\right)\right] \\ = \; & \exp\left(\theta\left(\kappa_0 + \left(\kappa_1 - 1\right)A_0 - A'Y_t\right) + \theta\Gamma'\left(\mu + FY_t + \mu_t\right) + \frac{1}{2}\theta^2\Gamma' G_t G_t' \Gamma\right) \end{aligned}$$

where

$$\Gamma = \left(1-\rho\right)e_c + \kappa_1 A,$$

Therefore, from the Euler equation,

$$\begin{aligned} 1 &= \mathbb{E}_t\left[M_{z_t,t+1} R_{c,t+1}\right] \\ &= \mathbb{E}_{\mu_t}\left[M_{z_t,t}^{A} \mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ} R_{c,t+1}\right]\right] = \mathbb{E}_{\mu_t}\left[\left(\beta^{\frac{1-\gamma}{1-\rho}} \mathbb{E}_{\pi_{z,t}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{1-\rho}} R_{c,t+1}^{\frac{1-\gamma}{1-\rho}}\right]\right)^{\frac{1-\eta}{1-\gamma}}\right] \\ &= \mathbb{E}_{\mu_t}\left[\exp\left(\begin{array}{c} \frac{1-\eta}{1-\rho}\left(\ln\beta + \kappa_0 + \left(\kappa_1 - 1\right)A_0 - A'Y_t\right) \\ +\frac{1-\eta}{1-\rho}\Gamma'\left(\mu + FY_t + \mu_t\right) + \frac{1}{2}\frac{1-\eta}{1-\rho}\theta\Gamma' G_t G_t' \Gamma \end{array}\right)\right] \\ &= \exp\left(\begin{array}{c} \frac{1-\eta}{1-\rho}\left(\ln\beta + \kappa_0 + \left(\kappa_1 - 1\right)A_0 - A'Y_t + \Gamma'\left(\mu + FY_t\right)\right) \\ +\frac{1}{2}\frac{1-\eta}{1-\rho}\theta\Gamma' G_t G_t' \Gamma + \frac{1}{2}\left(\frac{1-\eta}{1-\rho}\right)^2 \Gamma' G_t \Omega_t G_t' \Gamma \end{array}\right). \end{aligned}$$

Matching the constant term and the coefficient in front of Y_t yields

$$\left(1-\kappa_1\right)A_0 = \ln\beta + \kappa_0 + \Gamma'\mu + \frac{1}{2}\theta\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2$$
$$A = F'\Gamma + \frac{1}{2}\theta\Gamma' G G'\Gamma e_\sigma + \frac{1}{2}\frac{1-\eta}{1-\rho}\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 e_\tau,$$

where

$$G \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \rho_d\varphi_d & \sqrt{1-\rho_d^2}\varphi_d & 0 & 0 & 0 \\ 0 & 0 & \varphi_x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varphi_\tau \end{pmatrix}.$$

Solving the above system of equations leads to the solution in Eqs. (16a)-(16d).[6]

Innovation to the return on the consumption claim is

$$r_{c,t+1} - \mathbb{E}_t\left[r_{c,t+1}\right] = \left(e_c + \kappa_1 A\right)'\left(Y_{t+1} - \mathbb{E}_t\left[Y_{t+1}\right]\right) = B_r' G_t \epsilon_{t+1}$$

where $B_r = e_c + \kappa_1 A = \left(1, 0, \kappa_1 A_x, \kappa_1 A_\sigma, \kappa_1 A_\tau\right)'$. It follows that the conditional variance of $r_{c,t+1}$ is given by

$$\begin{aligned} \Sigma_t &= Var_t\left[r_{c,t+1}\right] = \mathbb{E}_t\left(B_r' G_t \epsilon_{t+1}\epsilon_{t+1}' G_t' B_r\right) = B_r' G_t\left(\mathcal{I} + \Omega_t\right) G_t' B_r \\ &= \left[1 + \left(\kappa_1 A_x \varphi_x\right)^2 + \left(\kappa_1 A_\tau \varphi_\tau\right)^2\right]\sigma_t^2 + \left(\kappa_1 A_\sigma \varphi_\sigma\right)^2\left(1 + \tau_t^2\right) \\ &\equiv B_{\Sigma,0} + B_{\Sigma,\sigma}\sigma_t^2 + B_{\Sigma,\tau}\tau_t^2 \\ &\equiv B_{\Sigma,0} + B_\Sigma' Y_t, \end{aligned}$$

where $B_{\Sigma,0} = B_{\Sigma,\tau} = \left(\kappa_1 A_\sigma \varphi_\sigma\right)^2$, $B_{\Sigma,\sigma} = 1 + \left(\kappa_1 A_x \varphi_x\right)^2 + \left(\kappa_1 A_\tau \varphi_\tau\right)^2$ and $B_\Sigma \equiv \left(0, 0, 0, B_{\Sigma,\sigma}, B_{\Sigma,\tau}\right)$. The above derivation of the expression of Σ_t uses the following result:

$$\mathbb{E}_t\left(\epsilon_{t+1}\epsilon_{t+1}'\right) = \mathbb{E}_{\mu_t} \circ \mathbb{E}_{\pi_{z,t}}\left(\epsilon_{t+1}\epsilon_{t+1}'\right) = \mathbb{E}_{\mu_t}\left(\mathcal{I} + \varphi_\sigma^{-2}\mu_t\mu_t'\right) = \mathcal{I} + \Omega_t.$$

Note that the conditional variance is driven by σ_t^2 and τ_t^2, with the sensitivity given by $B_{\Sigma,\sigma}$ and $B_{\Sigma,\tau}$, respectively (i.e., the last two components of the vector B_Σ). The innovation to the conditional variance is

$$\begin{aligned} & \Sigma_{t+1} - \mathbb{E}_t\left[\Sigma_{t+1}\right] \\ =\ & B_\Sigma'\left(Y_{t+1} - \mathbb{E}_t\left[Y_{t+1}\right]\right) = B_\Sigma' G_t \epsilon_{t+1} \\ =\ & B_{\Sigma,\sigma}\varphi_\sigma \epsilon_{\sigma,t+1} + B_{\Sigma,\tau}\varphi_\tau \sigma_t \epsilon_{\tau,t+1}. \end{aligned}$$

[6] In the benchmark model with constant model uncertainty (i.e., $\rho_\tau = \varphi_\tau = 0$), the constant term A_0 in the benchmark model corresponds to $A_0 + A_\tau\tau^2$ in the general model, given by

$$A_0 = \frac{\ln\beta + \kappa_0 + \left(1-\rho\right)\mu_c + \kappa_1 A_\sigma\left(1-\rho_\sigma\right)\sigma^2 + \frac{1}{2}\theta\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2 + \frac{1}{2}\frac{1-\eta}{1-\rho}\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2\tau^2}{1-\kappa_1}.$$

A.2 Intertemporal Marginal Rate of Substitution (IMRS) and the Risk-Free Rate

The logarithms of both components of the pricing kernel $M_{z_t,t+1} = M_{t+1}^{EZ} M_{z_t,t}^{A}$ can be re-written as

$$
\begin{aligned}
m_{t+1}^{EZ} &= \ln M_{t+1}^{EZ} = \theta \ln \beta - \rho\theta\Delta c_{t+1} + (\theta - 1)\, r_{c,t+1} \\
&= \theta \ln \beta - \rho\theta\Delta c_{t+1} + (\theta - 1)\left[\kappa_0 + \kappa_1 v_{t+1} - v_t + \Delta c_{t+1}\right] \\
&= \theta \ln \beta - \gamma\Delta c_{t+1} + (\theta - 1)\left[\kappa_0 + \kappa_1 \left(A_0 + A'Y_{t+1}\right) - \left(A_0 + A'Y_t\right)\right] \\
&= \theta \ln \beta + (\theta - 1)\left(\kappa_0 + (\kappa_1 - 1)\, A_0\right) - (\theta - 1)\, A'Y_t - \Lambda' Y_{t+1}
\end{aligned}
$$

and

$$
\begin{aligned}
m_{z_t,t}^{A} &= \ln M_{z_t,t}^{A} = -\frac{\eta - \gamma}{1 - \gamma}\left(\theta \ln \beta + \ln\left(\mathbb{E}_{\pi_{z,t}}\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\rho(1-\gamma)}{1-\rho}} R_{c,t+1}^{\frac{1-\gamma}{1-\rho}}\right]\right)\right) \\
&= -\frac{\eta - \gamma}{1 - \gamma}\left(\theta \ln \beta + \ln\left(\mathbb{E}_{\pi_{z,t}}\left[\exp\left((1-\rho)\,\theta\Delta c_{t+1} + \theta\left[\kappa_0 + \kappa_1 v_{t+1} - v_t\right]\right)\right]\right)\right) \\
&= -\frac{\eta - \gamma}{1 - \gamma}\left(\theta \ln \beta + \theta\left[\kappa_0 + (\kappa_1 - 1)\, A_0 - A'Y_t\right] + \ln\left(\mathbb{E}_{\pi_{z,t}}\left[\theta\Gamma' Y_{t+1}\right]\right)\right) \\
&= -\frac{\eta - \gamma}{1 - \gamma}\left(\theta \ln \beta + \theta\left[\kappa_0 + (\kappa_1 - 1)\, A_0 - A'Y_t\right] + \theta\Gamma'\left(\mu + FY_t + \mu_t\right) + \frac{1}{2}\theta^2\Gamma' G_t G_t' \Gamma\right) \\
&= -\frac{\eta - \gamma}{1 - \rho}\left(\ln \beta + \kappa_0 + (\kappa_1 - 1)\, A_0 - A'Y_t + \Gamma'\left(\mu + FY_t + \mu_t\right) + \frac{1}{2}\theta\Gamma' G_t G_t' \Gamma\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\Lambda &= \gamma e_c - (\theta - 1)\,\kappa_1 A \\
\Gamma &= (1 - \rho)\, e_c + \kappa_1 A
\end{aligned}
$$

Substituting the expressions of A_0 and A into the expression of $m_{z_t,t}^{A}$, we have

$$
\begin{aligned}
m_{z_t,t}^{A} &= -\frac{\eta - \gamma}{1 - \rho}\left(-A'Y_t + \Gamma'\left(FY_t + \mu_t\right) + \frac{1}{2}\theta\Gamma' G G' \Gamma \sigma_t^2\right) \\
&= \frac{\eta - \gamma}{1 - \rho}\left[\frac{1}{2}\frac{1 - \eta}{1 - \rho}\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 \tau_t^2 - \left(\kappa_1 A_\sigma \varphi_\sigma\right) z_t\right].
\end{aligned}
$$

Next, we determine the risk-free rate. We first derive the conditional mean of the pricing kernel

$$
\begin{aligned}
& \mathbb{E}_t\left[M_{z_t,t+1}\right] \\
= \; & \mathbb{E}_{\mu_t}\left[M^A_{z_t,t}\mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}\right]\right] \\
= \; & \mathbb{E}_{\mu_t}\left[\exp\begin{pmatrix} \theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0\right)-(\theta-1)A'Y_t \\ -\Lambda'\left(\mu+FY_t+\mu_t\right)+\frac{1}{2}\Lambda'G_tG_t'\Lambda \\ -\frac{\eta-\gamma}{1-\rho}\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t+\Gamma'\left(\mu+FY_t+\mu_t\right)+\frac{1}{2}\theta\Gamma'G_tG_t'\Gamma\right) \end{pmatrix}\right] \\
= \; & \exp\begin{pmatrix} \theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0\right)-(\theta-1)A'Y_t \\ -\Lambda'\left(\mu+FY_t\right)+\frac{1}{2}\Lambda'G_tG_t'\Lambda \\ -\frac{\eta-\gamma}{1-\rho}\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t+\Gamma'\left(\mu+FY_t\right)+\frac{1}{2}\theta\Gamma'G_tG_t'\Gamma\right) \end{pmatrix} \\
& \mathbb{E}_{\mu_t}\left[\exp\left(-\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'\mu_t\right)\right] \\
= \; & \exp\begin{pmatrix} \theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0\right)-(\theta-1)A'Y_t \\ -\Lambda'\left(\mu+FY_t\right)+\frac{1}{2}\Lambda'G_tG_t'\Lambda \\ -\frac{\eta-\gamma}{1-\rho}\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t+\Gamma'\left(\mu+FY_t\right)+\frac{1}{2}\theta\Gamma'G_tG_t'\Gamma\right) \\ +\frac{1}{2}\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right) \end{pmatrix}
\end{aligned}
$$

where in deriving the second equality, we used the fact that

$$
\begin{aligned}
& \mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}\right] \\
= \; & \mathbb{E}_{\pi_{z,t}}\left[\exp\left(\theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0\right)-(\theta-1)A'Y_t-\Lambda'Y_{t+1}\right)\right] \\
= \; & \exp\begin{pmatrix} \theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0\right)-(\theta-1)A'Y_t \\ -\Lambda'\left(\mu+FY_t+\mu_t\right)+\frac{1}{2}\Lambda'G_tG_t'\Lambda \end{pmatrix}
\end{aligned}
$$

Therefore, the risk-free rate is

$$
\begin{aligned}
r_{f,t} \; = \; & -\ln\mathbb{E}_t\left[M_{z_t,t+1}\right] \\
= \; & -\begin{pmatrix} \theta\ln\beta+(\theta-1)\left(\kappa_0+(\kappa_1-1)A_0-A'Y_t\right) \\ -\Lambda'\left(\mu+FY_t\right)+\frac{1}{2}\Lambda'G_tG_t'\Lambda \\ -\frac{\eta-\gamma}{1-\rho}\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t+\Gamma'\left(\mu+FY_t\right)+\frac{1}{2}\theta\Gamma'G_tG_t'\Gamma\right) \\ +\frac{1}{2}\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right) \end{pmatrix} \\
= \; & -\left(\ln\beta+\left(\theta-1-\frac{\eta-\gamma}{1-\rho}\right)\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t\right)\right) \\
& +\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'\left(\mu+FY_t\right)-\frac{1}{2}\Lambda'G_tG_t'\Lambda+\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\theta\Gamma'G_tG_t'\Gamma \\
& -\frac{1}{2}\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)
\end{aligned}
$$

Substituting the expression of A_0 into the above equation yields

$$\begin{aligned} r_{f,t} &= -\ln\beta+\rho\mu_c+\left(\theta-1-\frac{\eta-\gamma}{1-\rho}\right)\frac{1}{2}\theta\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2+\left(\theta-1-\frac{\eta-\gamma}{1-\rho}\right)A'Y_t \\ &\quad +\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)' FY_t-\frac{1}{2}\Lambda' G_t G_t'\Lambda+\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\theta\Gamma' G_t G_t'\Gamma-\frac{1}{2}\left(\frac{\eta-\rho}{1-\rho}\right)^2\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2\tau_t^2 \\ &= -\ln\beta+\rho\left(\mu_c+x_t\right)+\frac{1}{2}\theta\left(\theta-1\right)\Gamma' G_t G_t'\Gamma-\frac{1}{2}\Lambda' G_t G_t'\Lambda-\frac{1}{2}\frac{\eta-\rho}{1-\rho}\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2\tau_t^2 \\ &= -\ln\beta+\rho\left(\mu_c+x_t\right)-\frac{1}{2}\gamma^2\sigma_t^2+\frac{1}{2}\left(1-\theta\right)^2\left(1-\rho\right)^2\sigma_t^2-\frac{1}{2}\left(1-\theta\right)\Gamma' G_t G_t'\Gamma-\frac{1}{2}\frac{\eta-\rho}{1-\rho}\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2\tau_t^2. \end{aligned}$$

Therefore, we can show that the above expression of $r_{f,t}$ can be simplied as the one in Eq. (21) in Proposition 1. After further simplification, we can express $r_{f,t}$ as a linear function of the latent variables:

$$\begin{aligned} r_{f,t} &= \left[-\ln\beta-\frac{1}{2}\left(1-\theta\right)\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2+\rho\mu_c\right]+\rho x_t \\ &\quad -\frac{1}{2}\left(\gamma^2-\left(\gamma-\rho\right)^2+\left(1-\theta\right)\left[\left(1-\rho\right)^2+\left(\kappa_1 A_x\varphi_x\right)^2+\left(\kappa_1 A_\tau\varphi_\tau\right)^2\right]\right)\sigma_t^2 \\ &\quad -\frac{1}{2}\frac{\eta-\rho}{1-\rho}\left(\kappa_1 A_\sigma\varphi_\sigma\right)^2\tau_t^2 \\ &\equiv \mu_{f,0}+\mu_{f,1}x_t+\mu_{f,2}\sigma_t^2+\mu_{f,3}\tau_t^2. \end{aligned} \tag{26}$$

A.3 The Return on the Dividend Claim

The logarithm of the price-dividend ratio can be expressed in the vector form: $v_{m,t}=A_{0,m}+A_m'Y_t$. Then the return on the dividend claim is

$$\begin{aligned} r_{m,t+1} &= \kappa_{0,m}+\kappa_{1,m}v_{m,t+1}-v_{m,t}+\Delta d_{t+1} \\ &= \kappa_{0,m}+\left(\kappa_{1,m}-1\right)A_{0,m}-A_m'Y_t+B_{r,m}'Y_{t+1} \end{aligned}$$

where $B_{r,m}=e_d+\kappa_{1,m}A_m=\left(0,1,\kappa_{1,m}A_{x,m},\kappa_{1,m}A_{\sigma,m},\kappa_{1,m}A_{\tau,m}\right)'$. We can further simplify it as follows:

$$\begin{aligned} r_{m,t+1} &= \kappa_{0,m}+\left(\kappa_{1,m}-1\right)A_{0,m}+\mu_d+\left(\kappa_{1,m}A_{x,m}\right)x_{t+1}+\left(\phi-A_{x,m}\right)x_t \\ &\quad +\varphi_d\rho_d\sigma_t\epsilon_{c,t+1}+\sqrt{1-\rho_d^2}\varphi_d\sigma_t\epsilon_{d,t+1} \\ &\quad +\left(\kappa_{1,m}A_{\sigma,m}\right)\sigma_{t+1}^2-A_{\sigma,m}\sigma_t^2+\left(\kappa_{1,m}A_{\tau,m}\right)\tau_{t+1}^2-A_{\tau,m}\tau_t^2 \\ &\equiv \mu_{r,0}+\mu_{r,1}x_{t+1}+\mu_{r,2}x_t+\mu_{r,3}\sigma_t\epsilon_{c,t+1}+\mu_{r,4}\varphi_d\sigma_t\epsilon_{d,t+1} \\ &\quad +\mu_{r,5}\sigma_{t+1}^2+\mu_{r,6}\sigma_t^2+\mu_{r,7}\tau_{t+1}^2+\mu_{r,8}\tau_t^2. \end{aligned} \tag{27}$$

From the Euler equation, we have

$$
\begin{aligned}
1 &= \mathbb{E}_t\left[M_{z_t,t+1}R_{m,t+1}\right] = \mathbb{E}_{\mu_t}\left[M^{A}_{z_t,t}\mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}R_{m,t+1}\right]\right] \\
&= \mathbb{E}_{\mu_t}\left[\exp\begin{pmatrix} -\frac{\eta-\gamma}{1-\rho}\left(\ln\beta+\kappa_0+(\kappa_1-1)A_0-A'Y_t+\Gamma'(\mu+FY_t+\mu_t)+\frac{1}{2}\theta\Gamma'G_tG_t'\Gamma\right) \\ +\theta\ln\beta+(\theta-1)(\kappa_0+(\kappa_1-1)A_0)+\kappa_{0,m}+(\kappa_{1,m}-1)A_{0,m} \\ -(A_m+(\theta-1)A)'Y_t+(B_{r,m}-\Lambda)'(\mu+FY_t+\mu_t) \\ +\frac{1}{2}(B_{r,m}-\Lambda)'G_tG_t'(B_{r,m}-\Lambda) \end{pmatrix}\right] \\
&= \exp\begin{pmatrix} \frac{1-\eta}{1-\rho}\ln\beta-\frac{\eta-\rho}{1-\rho}(\kappa_0+(\kappa_1-1)A_0)+\kappa_{0,m}+(\kappa_{1,m}-1)A_{0,m} \\ +\left(B_{r,m}-\Lambda-\frac{\eta-\gamma}{1-\rho}\Gamma\right)'(\mu+FY_t)-\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\theta\Gamma'G_tG_t'\Gamma \\ -\left(A_m-\frac{\eta-\rho}{1-\rho}A\right)'Y_t+\frac{1}{2}(B_{r,m}-\Lambda)'G_tG_t'(B_{r,m}-\Lambda) \\ +\frac{1}{2}\left(B_{r,m}-\Lambda-\frac{\eta-\gamma}{1-\rho}\Gamma\right)'G_t\Omega_tG_t'\left(B_{r,m}-\Lambda-\frac{\eta-\gamma}{1-\rho}\Gamma\right) \end{pmatrix}
\end{aligned}
$$

where in deriving the third equality we used the fact

$$
\begin{aligned}
&\mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}R_{m,t+1}\right] \\
&= \mathbb{E}_{\pi_{z,t}}\left[\exp\begin{pmatrix} \theta\ln\beta+(\theta-1)(\kappa_0+(\kappa_1-1)A_0)-(\theta-1)A'Y_t-\Lambda'Y_{t+1} \\ +\kappa_{0,m}+(\kappa_{1,m}-1)A_{0,m}-A_m'Y_t+B_{r,m}'Y_{t+1} \end{pmatrix}\right] \\
&= \exp\begin{pmatrix} \theta\ln\beta+(\theta-1)(\kappa_0+(\kappa_1-1)A_0)+\kappa_{0,m}+(\kappa_{1,m}-1)A_{0,m} \\ -(A_m+(\theta-1)A)'Y_t+(B_{r,m}-\Lambda)'(\mu+FY_t+\mu_t) \\ +\frac{1}{2}(B_{r,m}-\Lambda)'G_tG_t'(B_{r,m}-\Lambda) \end{pmatrix}
\end{aligned}
$$

Matching the constant term and the coefficients in front of Y_t yields

$$
\begin{aligned}
0 &= \frac{1-\eta}{1-\rho}\ln\beta-\frac{\eta-\rho}{1-\rho}(\kappa_0+(\kappa_1-1)A_0)+\kappa_{0,m}+(\kappa_{1,m}-1)A_{0,m} \\
&\quad+\left(B_{r,m}-\Lambda-\frac{\eta-\gamma}{1-\rho}\Gamma\right)'\mu-\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\theta(\kappa_1A_\sigma\varphi_\sigma)^2 \\
&\quad+\frac{1}{2}(\kappa_{1,m}A_{\sigma,m}-(1-\theta)\kappa_1A_\sigma)^2\varphi_\sigma^2
\end{aligned}
$$

and

$$
\begin{aligned}
0 &= \left(B_{r,m}-\Lambda-\frac{\eta-\gamma}{1-\rho}\Gamma\right)'F-\frac{1}{2}\frac{\eta-\gamma}{1-\rho}\theta\Gamma'GG'\Gamma e_\sigma \\
&\quad-\left(A_m-\frac{\eta-\rho}{1-\rho}A\right)'+\frac{1}{2}(B_{r,m}-\Lambda)'GG'(B_{r,m}-\Lambda)e_\sigma \\
&\quad+\frac{1}{2}\left(\kappa_{1,m}A_{\sigma,m}-\left(1-\theta+\frac{\eta-\gamma}{1-\rho}\right)\kappa_1A_\sigma\right)^2\varphi_\sigma^2e_\tau.
\end{aligned}
$$

The solution is given by[7]

$$
\begin{aligned}
A_{0,m} &= \frac{\left[\begin{array}{c} \ln\beta + \kappa_{0,m} - \rho\mu_c + \mu_d + \kappa_{1,m}A_{\sigma,m}(1-\rho_\sigma)\sigma^2 \\ +\kappa_{1,m}A_{\tau,m}(1-\rho_\tau)\tau^2 - \frac{1}{2}\frac{\rho-\gamma}{1-\rho}\theta(\kappa_1 A_\sigma \varphi_\sigma)^2 \\ +\frac{1}{2}(\kappa_{1,m}A_{\sigma,m} - (1-\theta)\kappa_1 A_\sigma)^2\varphi_\sigma^2 \end{array}\right]}{1-\kappa_{1,m}}, \\
A_{c,m} &= A_{d,m} = 0, \\
A_{x,m} &= \frac{\phi - \rho}{1-\kappa_{1,m}\rho_x}, \\
A_{\sigma,m} &= \frac{\gamma^2 - 2\gamma\rho_d\varphi_d + \varphi_d^2 + H_x + H_\tau + \frac{\gamma-\rho}{1-\rho}\theta\left((1-\rho)^2 + (\kappa_1 A_x\varphi_x)^2 + (\kappa_1 A_\tau \varphi_\tau)^2\right)}{2(1-\kappa_{1,m}\rho_\sigma)}, \\
A_{\tau,m} &= \frac{\frac{\eta-\rho}{1-\rho}\frac{1-\eta}{1-\rho}(\kappa_1 A_\sigma\varphi_\sigma)^2 + \left(\kappa_{1,m}A_{\sigma,m} - \frac{\eta-\rho}{1-\rho}\kappa_1 A_\sigma\right)^2\varphi_\sigma^2}{2(1-\kappa_{1,m}\rho_\tau)},
\end{aligned}
$$

where

$$
\begin{aligned}
H_x &= (\kappa_{1,m}A_{x,m} + (\theta-1)\kappa_1 A_x)^2\varphi_x^2 \\
H_\tau &= (\kappa_{1,m}A_{\tau,m} + (\theta-1)\kappa_1 A_\tau)^2\varphi_\tau^2
\end{aligned}
$$

A.4 The Equity Premium

We first derive the risk premium (referred to as the equity premium as well) for the return on the consumption claim. The derivation for the equity premium for the return on the dividend claim is very similar. On the one hand, from the Euler equation, we have

$$
\begin{aligned}
1 &= \mathbb{E}_t\left[M_{z_t,t+1}R_{c,t+1}\right] \\
&= \mathbb{E}_{\mu_t}\left[M^A_{z_t,t}\mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}R_{c,t+1}\right]\right] \\
&= \mathbb{E}_{\mu_t}\left[M^A_{z_t,t}\exp\left(\begin{array}{c} \mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right) + \mathbb{E}_{\pi_{z,t}}(r_{c,t+1}) + \frac{1}{2}Var_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right) \\ +\frac{1}{2}Var_{\pi_{z,t}}(r_{c,t+1}) + Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1}, r_{c,t+1}\right) \end{array}\right)\right] \\
&= \exp\left(\frac{1}{2}Var_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right) + \frac{1}{2}Var_{\pi_{z,t}}(r_{c,t+1}) + Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1}, r_{c,t+1}\right)\right) \\
&\quad \mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t} + \mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right) + \mathbb{E}_{\pi_{z,t}}(r_{c,t+1})\right)\right]
\end{aligned}
$$

[7] In the benchmark model with constant model uncertainty (i.e., $\rho_\tau = \varphi_\tau = 0$), the constant term $A_{0,m}$ in the benchmark model corresponds to $A_{0,m} + A_{\tau,m}\tau^2$ in the general model, given by

$$
A_{0,m} = \frac{\left[\begin{array}{c} \ln\beta + \kappa_{0,m} - \rho\mu_c + \mu_d + \kappa_{1,m}A_{\sigma,m}(1-\rho_\sigma)\sigma^2 \\ +\frac{1}{2}\frac{\eta-\rho}{1-\rho}\frac{1-\eta}{1-\rho}(\kappa_1 A_\sigma\varphi_\sigma)^2\tau^2 + \frac{1}{2}\left(\kappa_{1,m}A_{\sigma,m} - \frac{\eta-\rho}{1-\rho}\kappa_1 A_\sigma\right)^2\varphi_\sigma^2\tau^2 \\ -\frac{1}{2}\frac{\rho-\gamma}{1-\rho}\theta(\kappa_1 A_\sigma\varphi_\sigma)^2 + \frac{1}{2}(\kappa_{1,m}A_{\sigma,m} - (1-\theta)\kappa_1 A_\sigma)^2\varphi_\sigma^2 \end{array}\right]}{1-\kappa_{1,m}}.
$$

where in deriving the last equality, we used the fact that the conditional variances and covariance for a fixed μ_t are independent of μ_t and thus can be taken out of the conditional expectation operator $\mathbb{E}_{\mu_t}$. On the other hand, by a similar argument, we have

$$\begin{aligned}
\mathbb{E}_t\left[R_{c,t+1}\right] &= \mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)+\frac{1}{2}Var_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right] \\
&= \exp\left(\frac{1}{2}Var_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right] \\
\mathbb{E}_t\left[M_{z_t,t+1}\right] &= \mathbb{E}_{\mu_t}\left[M^A_{z_t,t}\mathbb{E}_{\pi_{z,t}}\left[M^{EZ}_{t+1}\right]\right] = \mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\frac{1}{2}Var_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\right] \\
&= \exp\left(\frac{1}{2}Var_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\right]
\end{aligned}$$

Based on the above results, we have

$$\begin{aligned}
1 &= \exp\left(\frac{1}{2}Var_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\frac{1}{2}Var_{\pi_{z,t}}\left(r_{c,t+1}\right)+Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1},r_{c,t+1}\right)\right) \\
&\quad \mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right] \\
&= \mathbb{E}_t\left[R_{c,t+1}\right]\mathbb{E}_t\left[M_{z_t,t+1}\right]\exp\left(Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1},r_{c,t+1}\right)\right) \\
&\quad \frac{\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]}{\mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\right]}
\end{aligned}$$

Taking logs of both sides, we have

$$\begin{aligned}
&\ln\mathbb{E}_t\left[R_{c,t+1}\right]-r_{f,t} \\
=\ & -Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1},r_{c,t+1}\right)-\ln\frac{\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]}{\mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\right]}
\end{aligned}$$

Note that

$$Cov_{\pi_{z,t}}\left(m^{EZ}_{t+1},r_{c,t+1}\right)=Cov_{\pi_{z,t}}\left(-\Lambda'Y_{t+1},B_r'Y_{t+1}\right)=-B_r'G_tG_t'\Lambda$$

and

$$\begin{aligned}
&\frac{\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)+\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]}{\mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]\mathbb{E}_{\mu_t}\left[\exp\left(m^A_{z_t,t}+\mathbb{E}_{\pi_{z,t}}\left(m^{EZ}_{t+1}\right)\right)\right]} \\
=\ & \frac{\mathbb{E}_{\mu_t}\left[\exp\left(-\frac{\eta-\gamma}{1-\rho}\Gamma'\mu_t-\Lambda'\mu_t+B_r'\mu_t\right)\right]}{\mathbb{E}_{\mu_t}\left[\exp\left(B_r'\mu_t\right)\right]\mathbb{E}_{\mu_t}\left[\exp\left(-\frac{\eta-\gamma}{1-\rho}\Gamma'\mu_t-\Lambda'\mu_t\right)\right]} \\
=\ & \exp\begin{pmatrix}\frac{1}{2}\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda-B_r\right)'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda-B_r\right) \\ -\frac{1}{2}B_r'G_t\Omega_tG_t'B_r-\frac{1}{2}\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)\end{pmatrix} \\
=\ & \exp\left(-B_r'G_t\Omega_tG_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \ln \mathbb{E}_t\left[R_{c,t+1}\right]-r_{f,t} \\
= & -Cov_{\pi_{z,t}}\left(m_{t+1}^{EZ}, r_{c,t+1}\right)-\ln \frac{\mathbb{E}_{\mu_t}\left[\exp\left(m_{z_t,t}^{A}+\mathbb{E}_{\pi_{z,t}}\left(m_{t+1}^{EZ}\right)+\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]}{\mathbb{E}_{\mu_t}\left[\exp\left(\mathbb{E}_{\pi_{z,t}}\left(r_{c,t+1}\right)\right)\right]\mathbb{E}_{\mu_t}\left[\exp\left(m_{z_t,t}^{A}+\mathbb{E}_{\pi_{z,t}}\left(m_{t+1}^{EZ}\right)\right)\right]} \\
= & B_r' G_t G_t' \Lambda + B_r' G_t \Omega_t G_t' \left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right) \\
= & B_r' G_t\left(\mathcal{I}+\Omega_t\right) G_t' \Lambda+\frac{\eta-\gamma}{1-\rho} B_r' G_t \Omega_t G_t' \Gamma \\
= & \gamma\sigma_t^2+(1-\theta)\left(B_{r,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2\left(1+\tau_t^2\right)+\left[B_{r,x}\kappa_1 A_x \varphi_x^2+B_{r,\tau}\kappa_1 A_\tau \varphi_\tau^2\right]\sigma_t^2\right) \\
& +\frac{\eta-\gamma}{1-\rho} B_{r,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2 \tau_t^2 \\
= & \gamma\sigma_t^2+(1-\theta)\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2\left(1+\tau_t^2\right)+(1-\theta)\left[\left(\kappa_1 A_x \varphi_x\right)^2+\left(\kappa_1 A_\tau \varphi_\tau\right)^2\right]\sigma_t^2 \\
& +\frac{\eta-\gamma}{1-\rho}\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 \tau_t^2
\end{aligned}$$

A.5 The Variance Premium

We first derive the variance premium for the return on the consumption claim. The derivation for the variance premium for the return on the dividend claim is very similar.

First, the following result is needed in the derivation of the variance premiums. Suppose $\epsilon \sim N(\mu, \Sigma)$ follows a multi-variate normal distribution and a and b are known constant vectors, then it is well known that

$$\mathbb{E}\left[\left(b'\epsilon\right)\exp\left(a'\epsilon\right)\right]=b'\left(\mu+\Sigma a\right)\mathbb{E}\left[\exp\left(a'\epsilon\right)\right]. \tag{28}$$

The variance premium for the return on the consumption claim is defined as

$$\begin{aligned}
VP_t & = \mathbb{E}_t^Q\left[\Sigma_{t+1}\right]-\mathbb{E}_t\left[\Sigma_{t+1}\right] \\
& = \frac{\mathbb{E}_t\left[M_{z_t,t+1}\left(\Sigma_{t+1}-\mathbb{E}_t\left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_t\left[M_{z_t,t+1}\right]} \\
& = \frac{\mathbb{E}_{\mu_t}\left(M_{z_t,t}^{A}\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\left(\Sigma_{t+1}-\mathbb{E}_t\left[\Sigma_{t+1}\right]\right)\right]\right)}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]M_{z_t,t}^{A}\right)} \\
& = \mathbb{E}_{\mu_t}\left(\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]M_{z_t,t}^{A}}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]M_{z_t,t}^{A}\right)}\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\left(\Sigma_{t+1}-\mathbb{E}_t\left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]}\right)
\end{aligned}$$

The expression of the variance premium has a very intuitive interpretation: it is roughly a weighted average of the variance premiums across all possible models. Importantly, the weight (i.e., the term $\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]M_{z_t,t}^{A}}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]M_{z_t,t}^{A}\right)}$) is tilted in a way to reflect ambiguity aversion. Based on the well-known finding in Eq. (28), we can show that for a given model with fixed

μ_t, the variance premium under this particular model is

$$\begin{aligned} & \frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\left(\Sigma_{t+1}-\mathbb{E}_t\left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right]} \\ = & \frac{\mathbb{E}_{\pi_{z,t}}\left[\exp\left(-\Lambda' G_t \epsilon_{t+1}\right)\left(\Sigma_{t+1}-\mathbb{E}_t\left[\Sigma_{t+1}\right]\right)\right]}{\mathbb{E}_{\pi_{z,t}}\left[\exp\left(-\Lambda' G_t \epsilon_{t+1}\right)\right]} \\ = & B_\Sigma' G_t\left(-G_t'\Lambda + z_t\right) = -B_\Sigma' G_t G_t' \Lambda + B_\Sigma' \mu_t \end{aligned}$$

and also the corresponding weight function is given by

$$\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A\right)} = \frac{\exp\left(-\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'\mu_t\right)}{\mathbb{E}_{\mu_t}\left[\exp\left(-\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right)'\mu_t\right)\right]}$$

Therefore, the variance premium that takes into account model uncertainty is given by

$$\begin{aligned} VP_t &= \mathbb{E}_{\mu_t}\left(\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A\right)}\left(-B_\Sigma' G_t G_t'\Lambda + B_\Sigma'\mu_t\right)\right) \\ &= \mathbb{E}_{\mu_t}\left(\frac{\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A}{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[M_{t+1}^{EZ}\right] M_{z_t,t}^A\right)} B_\Sigma'\mu_t\right) - B_\Sigma' G_t G_t'\Lambda \\ &= -B_\Sigma' G_t \Omega_t G_t'\left(\frac{\eta-\gamma}{1-\rho}\Gamma+\Lambda\right) - B_\Sigma' G_t G_t'\Lambda \\ &= -(1-\theta)\left[B_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2\left(1+\tau_t^2\right) + B_{\Sigma,\tau}\kappa_1 A_\tau \varphi_\tau^2 \sigma_t^2\right] - \frac{\eta-\gamma}{1-\rho} B_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2 \tau_t^2 \\ &= -(1-\theta)\left(1+\left(\kappa_1 A_x \varphi_x\right)^2 + \left(\kappa_1 A_\tau \varphi_\tau\right)^2\right)\kappa_1 A_\sigma \varphi_\sigma^2 - \frac{\gamma-\rho}{1-\rho}\left(\kappa_1 A_\sigma \varphi_\sigma\right)^2 \kappa_1 A_\tau \varphi_\tau^2 \sigma_t^2 \\ &\quad -\frac{\eta-\rho}{1-\rho}\left(1+\left(\kappa_1 A_x \varphi_x\right)^2 + \left(\kappa_1 A_\tau \varphi_\tau\right)^2\right)\kappa_1 A_\sigma \varphi_\sigma^2 \tau_t^2. \end{aligned}$$

A.6 Predictability

In this subsection, we derive the expressions for the slope coefficients in the predictive regressions of next period's stock market return on the price-dividend ration or the variance premium. The expressions are similar for the predictive regressions of stock market excess returns and are thus omitted here.

Recall that the (log) price-dividend ration is given by $v_{m,t} = A_{0,m} + A_{x,m}x_t + A_{\sigma,m}\sigma_t^2 + A_{\tau,m}\tau_t^2$. Furthermore,

$$\mathbb{E}_t\left(r_{m,t+1}\right) \equiv B_{m,0} + B_{m,x}x_t + B_{m,\sigma}\sigma_t^2 + B_{m,\tau}\tau_t^2,$$

where

$$
\begin{aligned}
B_{m,0} &\equiv \kappa_{0,m} + (\kappa_{1,m} - 1)\left(A_{0,m} + A_{\sigma,m}\sigma^2 + A_{\tau,m}\tau^2\right) + \mu_d \\
&\quad + (1 - \rho_\sigma \kappa_{1,m}) A_{\sigma,m}\sigma^2 + (1 - \rho_\tau \kappa_{1,m}) A_{\tau,m}\tau^2 \\
B_{m,x} &\equiv \rho, \\
B_{m,\sigma} &\equiv -A_{\sigma,m}(1 - \kappa_{1,m}\rho_\sigma), \\
B_{m,\tau} &\equiv -A_{\tau,m}(1 - \kappa_{1,m}\rho_\tau).
\end{aligned}
$$

Similar to the expression in the end of Section B.5, the conditional variance premium of the equity return is given by:

$$
\begin{aligned}
VP_t^m &= -(1-\theta)\left[B^m_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2 + B^m_{\Sigma,\tau}\kappa_1 A_\tau \varphi_\tau^2 \sigma_t^2\right] - \frac{\eta - \rho}{1-\rho} B^m_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2 \tau_t^2 \\
&\equiv B_{VP,0} + B_{VP,\sigma}\sigma_t^2 + B_{VP,\tau}\tau_t^2, \qquad (29)
\end{aligned}
$$

where

$$
\begin{aligned}
B_{VP,0} &= -(1-\theta) B^m_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2, \\
B_{VP,\sigma} &= -(1-\theta) B^m_{\Sigma,\tau}\kappa_1 A_\tau \varphi_\tau^2, \\
B_{VP,\tau} &= -\frac{\eta-\rho}{1-\rho} B^m_{\Sigma,\sigma}\kappa_1 A_\sigma \varphi_\sigma^2,
\end{aligned}
$$

and

$$
\begin{aligned}
B^m_{\Sigma,0} &= (\kappa_{1,m} A_{\sigma,m}\varphi_\sigma)^2, \\
B^m_{\Sigma,\sigma} &= \varphi_d^2 + (\kappa_{1,m} A_{x,m}\varphi_x)^2 + (\kappa_{1,m} A_{\tau,m}\varphi_\tau)^2, \\
B^m_{\Sigma,\tau} &= (\kappa_{1,m} A_{\sigma,m}\varphi_\sigma)^2.
\end{aligned}
$$

B Time Aggregation and Particle-Filter Estimation

B.1 Time Aggregation

We assume a monthly decision interval and use annual data in calibration following Bansal et al. (2016). Let j denote the time index for data sampling frequency and therefore j increments annually. Let calendar time t denote the time index for the agent's monthly decision interval and therefore t increments monthly. We assume that the annual data in year j is observed at the last month of the year, that is, the month of $12j$.

First, derive log annual consumption growth rate that accounts for time aggregation. Following Bansal et al. (2016) and Schorfheide et al. (2018), we define annual consumption in year j as the sum of monthly consumption, denoted by $C_{12(j-1)+m}$, $m = 1, \cdots, 12$ for each month m within the year j. Therefore, the log annual consumption growth in year j, denoted by $\Delta c_{(j)}$, can be well approximated as

$$
\begin{aligned}
\Delta c_{(j)} &= \log \frac{\sum_{m=1}^{12} C_{12(j-1)+m}}{\sum_{m=1}^{12} C_{12(j-2)+m}} \\
&\approx \sum_{k=2}^{24} \frac{12 - |12 - (k-1)|}{12} \Delta c_{t-24+k}.
\end{aligned}
$$

Similarly,

$$\Delta d_{(j)} \approx \sum\nolimits_{k=2}^{24} \frac{12-|12-(k-1)|}{12}\Delta d_{t-24+k}.$$

Next, we derive the annual time-aggregated market return, the risk-free rate, and the variance premium. Next, recall that $r_{m,t+1}$, $r_{f,t}$, VP_t^m in the time interval between month t and month $t+1$ are given in Eqs. (27), (26), and (29). Aggregating over the span of the year j, we have

$$\begin{aligned}
r_{f,(j)} &= \sum\nolimits_{k=1}^{M}\left[\mu_{f,0}+\mu_{f,1}x_{t-k-1}+\mu_{f,2}\sigma_{c,t-k}^2+\mu_{f,3}\tau_{t-k}^2\right], \\
r_{m,(j)} &= \sum\nolimits_{k=0}^{M-1}\left[\begin{array}{c}\mu_{r,0}+\mu_{r,1}x_{t-k}+\mu_{r,2}x_{t-k-1}+\mu_{r,3}\sigma_{t-k-1}\epsilon_{c,t-k}+\mu_{r,4}\sigma_{d,t-k-1}\epsilon_{d,t-k} \\ +\mu_{r,5}\sigma_{t-k}^2+\mu_{r,6}\sigma_{t-k-1}^2+\mu_{r,7}\tau_{t-k}^2+\mu_{r,8}\tau_{t-k-1}^2\end{array}\right], \\
VP_{(j)}^m &= E_t\left[\sum\nolimits_{k=0}^{M-1}VP_{t+k}^m\right]=E_t\left[\sum\nolimits_{k=0}^{M-1}\left(B_{VP,0}+B_{VP,\sigma}\sigma_{t+k}^2+B_{VP,\tau}\tau_{t+k}^2\right)\right] \\
&= hB_{VP,0}+B_{VP,\sigma}\left[h\sigma^2+\left(\sigma_t^2-\sigma^2\right)\frac{1-\rho_\sigma^M}{1-\rho_\sigma}\right]+B_{VP,\tau}\left[M\tau^2+\left(\tau_t^2-\tau^2\right)\frac{1-\rho_\tau^M}{1-\rho_\tau}\right] \\
&\equiv B_{VP,0}^{(a)}+B_{VP,\sigma}^{(a)}\sigma_t^2+B_{VP,\tau}^{(a)}\tau_t^2.
\end{aligned}$$

Lastly, we derive the annual time-aggregated price-dividend ratio.

$$\begin{aligned}
pd_{(j)} &= \log\frac{P_t}{D_t+D_{t-1}+D_{t-2}+\cdots+D_{t-11}} \\
&\approx \log\frac{P_t}{D_t}-\log M+\log D_t-\frac{1}{12}\left(\log D_t+\cdots+\log D_{t-11}\right) \\
&= v_{m,t}-\log M+\sum\nolimits_{k=1}^{M}\frac{M-k}{M}\Delta d_{t+1-k} \\
&= \left[A_{0,m}-\log M+\frac{1}{2}\left(M-1\right)\mu_d\right]+A_{x,m}x_t+\sum\nolimits_{k=1}^{M}\frac{M-k}{M}\phi x_{t-k}+A_{\sigma,m}\sigma_t^2 \\
&\quad +A_{\tau,m}\tau_t^2+\sum\nolimits_{k=1}^{M}\frac{M-k}{M}\left[\rho_d\varphi_d\sigma_{t-k}\epsilon_{c,t+1-k}+\varphi_d\sqrt{1-\rho_d^2}\sigma_{t-k}\epsilon_{d,t+1-k}\right].
\end{aligned}$$

B.2 State Space Representation

We use the general LRR model to provide details about particle-filter-based Bayesian estimation in this appendix. The standard LRR model is a special case and thus entails a very similar implementation.

Stacking the observed annual data into the vector $y_t^o \equiv \left(\Delta c_{(j)}^o, \Delta d_{(j)}^o, r_{m,(j)}^o, r_{f,(j)}^o, v_{m,(j)}^o, VP_{(j)}^{m,o}\right)'$, where the superscript "o" denotes the corresponding *observed* variable.[8] Similarly, we stack the model-implied counterparts into the vector $y_t \equiv \left(\Delta c_{(j)}, \Delta d_{(j)}, r_{m,(j)}, r_{f,(j)}, v_{m,(j)}, VP_{(j)}^m\right)'$, where we drop the superscript "o" from a variable to refer to its model-implied counterpart. In this appendix, we derive a state-space system for the observables y_t^o and the underlying latent state.

[8]When estimating the standard LRR model, we do not use the variance premium data since the calibrated model implies a negligible constant variance premium.

First, we make the following assumptions about measurement errors:

$$\begin{aligned}
\Delta c^o_{(j)} &= \Delta c_{(j)} + \sigma^a_\epsilon \left(\epsilon^a_{(j)} - \epsilon^a_{(j-1)}\right), \\
\Delta d^o_{(j)} &= \Delta d_{(j)} + \sigma^a_{d,\epsilon} \left(\epsilon^a_{d,(j)} - \epsilon^a_{d,(j-1)}\right), \\
r^o_{f,(j)} &= r_{f,(j)} + \sigma^a_{f,\epsilon}\epsilon^a_{f,(j)}, \\
pd^o_{(j)} &= pd_{(j)} + \sigma^a_{pd,\epsilon}\epsilon^a_{pd,(j)}, \\
VP^o_{(j)} &= VP_{(j)} + \sigma^a_{vp,\epsilon}\epsilon^a_{vp,(j)},
\end{aligned}$$

where $\epsilon^a_{(j)}$, $\epsilon^a_{d,(j)}$, $\epsilon^a_{f,(j)}$, $\epsilon^a_{pd,(j)}$, $\epsilon^a_{vp,(j)}$ are i.i.d. standard normal random variables. Note that the first assumption is the same as in Schorfheide et al. (2018) where multiplicative i.i.d. measurement errors are assumed for the *level* of annual consumption. Similar assumption is made for the dividend process in this paper.

We now derive the measurement equations for each observable before we derive the state equations to complete the construction of the state-space system.

First, note that

$$\begin{aligned}
\Delta c^o_{(j)} &= \Delta c_{(j)} + \sigma^a_\epsilon \left(\epsilon^a_{(j)} - \epsilon^a_{(j-1)}\right) \\
&= \left[\begin{array}{c} \frac{1}{M}\Delta c_t + \frac{2}{M}\Delta c_{t-1} + \cdots + \frac{M-1}{M}\Delta c_{t-(M-2)} + \frac{M}{M}\Delta c_{t-(M-1)} \\ + \frac{M-1}{M}\Delta c_{t-M} + \cdots + \frac{2}{M}\Delta c_{t-(2M-3)} + \frac{1}{M}\Delta c_{t-(2M-2)} + \sigma^a_\epsilon\left(\epsilon^a_{(j)} - \epsilon^a_{(j-1)}\right) \end{array}\right].
\end{aligned}$$

Denote $\Xi = \left[\frac{1}{M} \quad \frac{2}{M} \quad \cdots \quad \frac{M-1}{M} \quad 1 \quad \frac{M-1}{M} \quad \cdots \quad \frac{2}{M} \quad \frac{1}{M}\right]$, then

$$\begin{aligned}
&\Delta c^o_{(j)} \\
&= \Xi \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \vdots \\ \Delta c_{t-(2M-2)} \end{bmatrix} + \sigma^a_\epsilon \left(\epsilon^a_{(j)} - \epsilon^a_{(j-1)}\right) \\
&= \Xi \left[\mu_c 1_{(2M-1)\times 1} + \left[0_{(2M-1)\times 1}, I_{2M-1}\right] s^{(1)}_t + I_{2M-1} s^{(2)}_t\right] + \left[1, 0_{1\times(M-1)}, -1\right] s^{(1)}_{\epsilon,t} \\
&\equiv D^{(1)} + Z^{(1,1)} s^{(1)}_t + Z^{(1,2)} s^{(2)}_t + Z^{(1,1)}_\epsilon s^{(1)}_{\epsilon,t},
\end{aligned}$$

where $1_{m\times n}$ or $0_{m\times n}$ denotes a $n \times 1$ vector of ones or zeros, respectively, I_n denotes the $n \times n$ identity matrix, and

$$\underset{(2M)}{s^{(1)}_t} = \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-(2M-1)} \end{bmatrix}, \quad \underset{(2M-1)}{s^{(2)}_t} = \begin{bmatrix} \sigma_{t-1}\epsilon_{c,t} \\ \sigma_{t-2}\epsilon_{c,t-1} \\ \vdots \\ \sigma_{t-(2M-1)}\epsilon_{c,t-(2M-2)} \end{bmatrix},$$

$$\underset{(2M-1)\times(M+1)}{Z^{(1,1)}_\epsilon} = \begin{bmatrix} 0_{(M-1)\times(M+1)} \\ \left[1, 0_{1\times(M-1)}, -1\right] \\ 0_{(M-1)\times(M+1)} \end{bmatrix}, \quad \underset{(M+1)}{s^{(1)}_{\epsilon,t}} = \begin{bmatrix} \sigma^a_\epsilon \epsilon^a_t \\ \vdots \\ \sigma^a_\epsilon \epsilon^a_{t-M} \end{bmatrix}.$$

Similarly, for the dividend process we have

$$\begin{aligned}\Delta d^{o}_{(j)} &= \Delta d_{(j)} + \sigma^{a}_{d,\epsilon}\left(\epsilon^{a}_{d,(j)} - \epsilon^{a}_{d,(j-1)}\right) \\ &\equiv D^{(2)} + Z^{(2,1)} s^{(1)}_{t} + Z^{(2,2)} s^{(2)}_{t} + Z^{(2,3)} s^{(3)}_{t} + Z^{(2,2)}_{\epsilon} s^{(2)}_{\epsilon,t},\end{aligned}$$

where $D^{(2)} = M\mu_d$, $Z^{(2,1)} = [0, \Xi]$, $Z^{(2,2)} = \rho_d \varphi_d \Xi$, $Z^{(2,3)} = \sqrt{1-\rho_d^2}\,\Xi$, $Z^{(2,2)}_{\epsilon} = \left[1, 0_{1\times(M-1)}, -1\right]$, and

$$\underset{(2M-1)}{s^{(3)}_{t}} = \begin{bmatrix} \varphi_d \sigma_{t-1}\epsilon_{d,t} \\ \varphi_d \sigma_{t-2}\epsilon_{d,t-1} \\ \vdots \\ \varphi_d \sigma_{t-(2M-1)}\epsilon_{d,t-(2M-2)} \end{bmatrix}, \quad \underset{(M+1)}{s^{(2)}_{\epsilon,t}} = \begin{bmatrix} \sigma^{a}_{d,\epsilon}\epsilon^{a}_{d,t} \\ \vdots \\ \sigma^{a}_{d,\epsilon}\epsilon^{a}_{d,t-M} \end{bmatrix}.$$

Next, note that

$$\begin{aligned}\begin{bmatrix} r^{o}_{m,(j)} \\ r^{o}_{f,(j)} \\ pd^{o}_{(j)} \\ VP^{m,o}_{(j)} \end{bmatrix} &= \begin{bmatrix} r_{m,(j)} \\ r_{f,(j)} \\ pd_{(j)} \\ VP^{m}_{(j)} \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma^{a}_{f,\epsilon}\epsilon^{a}_{f,(j)} \\ \sigma^{a}_{pd,\epsilon}\epsilon^{a}_{pd,(j)} \\ \sigma^{a}_{vp,\epsilon}\epsilon^{a}_{vp,(j)} \end{bmatrix} \\
&= \begin{bmatrix} M\mu_{r,0} \\ M\mu_{f,0} \\ \mu_{pd,0} \\ B^{(a)}_{VP,0} \end{bmatrix} + \begin{bmatrix} \left[[\mu_{r,1}, \mu_{r,1}+\mu_{r,2}, \cdots, \mu_{r,1}+\mu_{r,2}, \mu_{r,2}], 0_{1\times(M-1)}\right] \\ \left[0, \mu_{f,1} 1_{1\times M}, 0_{1\times(M-1)}\right] \\ \left[\left[A_{x,m}, \phi\frac{M-1}{M}, \cdots, \phi\frac{1}{M}, 0\right], , 0_{1\times(M-1)}\right] \\ \left[0_{1\times(2M)}\right] \end{bmatrix} s^{(1)}_{t} \\
&+ \begin{bmatrix} \left[\mu_{r,3} 1_{1\times M}, 0_{1\times(M-1)}\right] \\ 0_{1\times(2M-1)} \\ \left[\rho_d\varphi_d\left[\frac{M-1}{M}, \cdots, \frac{1}{M}, 0\right], 0_{1\times(M-1)}\right] \\ 0_{1\times(2M-1)} \end{bmatrix} s^{(2)}_{t} \\
&+ \begin{bmatrix} \left[\mu_{r,4} 1_{1\times M}, 0_{1\times(M-1)}\right] \\ 0_{1\times(2M-1)} \\ \left[\sqrt{1-\rho_d^2}\left[\frac{M-1}{M}, \cdots, \frac{1}{M}, 0\right], 0_{1\times(M-1)}\right] \\ 0_{1\times(2M-1)} \end{bmatrix} s^{(3)}_{t} \\
&+ \begin{bmatrix} \sum_{i=0}^{M-1}\left[\mu_{r,5}\sigma^2_{t-i} + \mu_{r,6}\sigma^2_{t-i-1} + \mu_{r,7}\tau^2_{t-i} + \mu_{r,8}\tau^2_{t-i-1}\right] \\ \sum_{i=1}^{M}\left[\mu_{f,2}\sigma^2_{t-i} + \mu_{f,3}\tau^2_{t-i}\right] \\ A_{\sigma,m}\sigma^2_t + A_{\tau,m}\tau^2_t \\ B^{(a)}_{VP,\sigma}\sigma^2_t + B^{(a)}_{VP,\tau}\tau^2_t \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma^{a}_{f,\epsilon}\epsilon^{a}_{f,(j)} \\ \sigma^{a}_{pd,\epsilon}\epsilon^{a}_{pd,(j)} \\ \sigma^{a}_{vp,\epsilon}\epsilon^{a}_{vp,(j)} \end{bmatrix} \\
&\equiv D^{(3)} + Z^{(3,1)} s^{(1)}_{t} + Z^{(3,2)} s^{(2)}_{t} + Z^{(3,3)} s^{(3)}_{t} + Z^{v,(3)} s^{v}_{t} + Z^{\tau,(3)} s^{\tau}_{t} + u^{(3)},\end{aligned}$$

where $\mu_{pd,0} \equiv A_{0,m} - \log M + \frac{1}{2}(M-1)\mu_d$, and

$$\underset{(M+1)}{s^{v}_{t}} = \begin{bmatrix} \sigma^2_t \\ \sigma^2_{t-1} \\ \vdots \\ \sigma^2_{t-M} \end{bmatrix}, \quad \underset{(M+1)}{s^{\tau}_{t}} = \begin{bmatrix} \tau^2_t \\ \tau^2_{t-1} \\ \vdots \\ \tau^2_{t-M} \end{bmatrix}, \quad u^{(3)} = \begin{bmatrix} 0 \\ \sigma^{a}_{f,\epsilon}\epsilon^{a}_{f,(j)} \\ \sigma^{a}_{pd,\epsilon}\epsilon^{a}_{pd,(j)} \\ \sigma^{a}_{vp,\epsilon}\epsilon^{a}_{vp,(j)} \end{bmatrix}.$$

Stacking the observed annual data into the vector y_t^o, we thus derive the following measurement equation:

$$y_t^o = D + Z s_t + Z^v s_t^v + Z^\tau s_t^\tau + u_t, \text{ where } u_t \sim N(0, R).$$

where $u_t = \left(0, 0, u_t^{(3)\prime}\right)^\prime$ and $s_t = \left(s_t^{(1)\prime}, s_t^{(2)\prime}, s_{\epsilon,t}^{(1)\prime}, s_t^{(3)\prime}, s_{\epsilon,t}^{(2)\prime}\right)^\prime$ is a $(8M \times 1)$ vector.

Next, we derive the state equation. It is straightforward to derive the transition equation for the vector of the state variables s_t as follows:

$$s_{t+1} = \Phi s_t + v_{t+1}, \text{ where } v_{t+1} \sim N(0, \Sigma_t^s).$$

We now derive the transition equation for the volatility terms. Define $\sigma_t = \sigma \exp(h_{\sigma,t})$ and $\tau_t = \tau \exp(h_{\tau,t})$. Recall that

$$\begin{aligned} \sigma_{t+1}^2 &= \sigma^2 + \rho_\sigma \left(\sigma_t^2 - \sigma^2\right) + \varphi_\sigma \epsilon_{\sigma,t+1} \\ &= \sigma^2 + \rho_\sigma \left(\sigma_t^2 - \sigma^2\right) + \varphi_\sigma \left(\widetilde{z}_t + \widetilde{\epsilon}_{\sigma,t+1}\right) \\ &\equiv \sigma^2 + \rho_\sigma \left(\sigma_t^2 - \sigma^2\right) + \widetilde{\varphi}_\sigma w_{\sigma,t+1}, \end{aligned}$$

where $\widetilde{\varphi}_\sigma = \varphi_\sigma \sqrt{1+\tau_t^2}$ and the disturbance term $\left(\widetilde{z}_t + \widetilde{\epsilon}_{\sigma,t+1}\right)$ is expressed equivalently as $\sqrt{1+\tau_t^2} w_{\sigma,t+1}$ with $w_{\sigma,t+1}$ being a standard normal white noise. As a result, by log-linearization, we have

$$h_{\sigma,t+1} \approx \rho_\sigma h_{\sigma,t} + \frac{\varphi_\sigma}{2\sigma^2} \sqrt{1+\tau_t^2} w_{\sigma,t+1}.$$

A similar expression for ambiguity can be derived as follows:

$$h_{\tau,t+1} = \rho_\tau h_{\tau,t} + \frac{\varphi_\tau}{2\tau^2} \sigma_t w_{\tau,t+1}.$$

Stacking $h_{\sigma,t}$ and $h_{\tau,t}$ into a vector h_t, we thus have

$$\begin{aligned} h_{t+1} &= \begin{bmatrix} h_{\sigma,t+1} \\ h_{\tau,t+1} \end{bmatrix} = \begin{bmatrix} \rho_\sigma & \\ & \rho_\tau \end{bmatrix} \begin{bmatrix} h_{\sigma,t} \\ h_{\tau,t} \end{bmatrix} + \begin{bmatrix} \frac{\varphi_\sigma}{2\sigma^2}\sqrt{1+\tau_t^2} & \\ & \frac{\varphi_\tau}{2\tau^2}\sigma_t \end{bmatrix} \begin{bmatrix} w_{\sigma,t+1} \\ w_{\tau,t+1} \end{bmatrix} \\ &\equiv \Psi h_t + \Sigma_t^h w_{t+1} \end{aligned}$$

In summary, the state-space representation with stochastic volatilities is given by:

$$y_{t+1}^o = D + Z s_{t+1} + Z^v s_{t+1}^v + Z^\tau s_{t+1}^\tau + u_{t+1}, \tag{30a}$$

$$s_{t+1} = \Phi s_t + v_{t+1}, \tag{30b}$$

$$h_{t+1} = \Psi h_t + \Sigma_t^h w_{t+1}, \tag{30c}$$

where $u_{t+1} \sim N(0, R)$, $v_{t+1} \sim N(0, \Sigma_t^s)$, and $w_{t+1} \sim N(0, I_2)$. Note that both Σ_t^s and Σ_t^h depends on h_t.

B.3 Particle Filter

Note that the state-space representation is linear conditional on h_t. Furthermore, the distribution $p\left(s_t \left| h_t, y^o_{1:t}\right.\right)$ is normal:

$$s_t \left| h_t, y^o_{1:t} \right. \sim N\left(s_{t|t}\left(h_t\right), P_{t|t}\left(h_t\right)\right)$$

We thus abbreviate the density of $s_t \left| h_t, y^o_{1:t}\right.$ by $p_N\left(s_t \left| s_{t|t}\left(h_t\right), P_{t|t}\left(h_t\right)\right.\right)$.

Note that the joint distribution of (h_t, s_t) can be factorized as $p\left(h_t, s_t \left| y^o_{1:t}\right.\right) = p\left(h_t \left| y^o_{1:t}\right.\right) p\left(s_t \left| h_t, y^o_{1:t}\right.\right)$. We can approximate $\left(h_t, s_t \left| y^o_{1:t}\right.\right)$ by the quadruplets $\left\{h^j_t, s^j_{t|t}, P^j_{t|t}, W^j_t\right\}_{j=1}^{J}$.

The implementation of the particle filter is based on Algorithm 13 in Herbst and Schorfheide (2015).

1. Initialization: draw the initial particles $\{h^j_{(-M):0}, W^j_0\}$, $j = 1, \cdots, J$ with $W^j_0 = 1$, and specify $s^j_{0|0}$ and $P^j_{0|0}$.

2. Recursion. For $t = 1, \cdots, T$

   (a) Forecasting h_t. Propagate the period $t-1$ particles $\{h^j_{t-1}, W^j_{t-1}\}$ by iterating the state-transition equation forward:

   $$h^j_t = \Psi h^j_{t-1} + \Sigma^h(h^j_{t-1}) w_t.$$

   Calculate

   $$s^{v,j}_t = \begin{bmatrix} \sigma^{2,j}_t \\ \sigma^{2,j}_{t-1} \\ \vdots \\ \sigma^{2,j}_{t-M} \end{bmatrix} = \begin{bmatrix} \sigma^2 \exp(2h^j_{\sigma,t}) \\ \sigma^2 \exp(2h^j_{\sigma,t-1}) \\ \vdots \\ \sigma^2 \exp(2h^j_{\sigma,t-M}) \end{bmatrix}, s^{\tau,j}_t = \begin{bmatrix} \tau^{2,j}_t \\ \tau^{2,j}_{t-1} \\ \vdots \\ \tau^{2,j}_{t-M} \end{bmatrix} = \begin{bmatrix} \tau^2 \exp(2h^j_{\tau,t}) \\ \tau^2 \exp(2h^j_{\tau,t-1}) \\ \vdots \\ \tau^2 \exp(2h^j_{\tau,t-M}) \end{bmatrix}.$$

   Forecasting s_t and y_t:

   $$\begin{aligned} s^j_{t|t-1} &= \Phi s^j_{t-1|t-1}, \\ P^j_{t|t-1} &= \Phi P^j_{t-1|t-1} \Phi' + \Sigma^s(h^j_{t-1}) \Sigma^s(h^j_{t-1})', \\ y^j_{t|t-1} &= D + Z s^j_{t|t-1} + Z^v s^{v,j}_t + Z^\tau s^{\tau,j}_t, \\ F^j_{t|t-1} &= Z P^j_{t|t-1} Z' + RR'. \end{aligned}$$

   Calculate predictive distribution:

   $$\begin{aligned} & p_N\left(y^o_t \left| y^j_{t|t-1}, F^j_{t|t-1}\right.\right) \\ = \ & (2\pi)^{-n/2} \left|F^j_{t|t-1}\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(y^o_t - y^j_{t|t-1}\right)' \left(F^j_{t|t-1}\right)^{-1} \left(y^o_t - y^j_{t|t-1}\right)\right\} \end{aligned}$$

   and the incremental weights

   $$w^j_t = p_N\left(y^o_t \left| y^j_{t|t-1}, F^j_{t|t-1}\right.\right)$$

(b) **Updating.** Compute

$$\begin{aligned} s^j_{t|t} &= s^j_{t|t-1} + P^j_{t|t-1} Z' \left(F^j_{t|t-1}\right)^{-1} \left(y^o_t - y^j_{t|t-1}\right) \\ P^j_{t|t} &= P^j_{t|t-1} - P^j_{t|t-1} Z' \left(F^j_{t|t-1}\right)^{-1} Z P^j_{t|t-1} \end{aligned}$$

and define the normalized weights:

$$W^j_t = \frac{w^j_t}{\sum_{j=1}^J w^j_t}$$

(c) **Selection.** Resample the particles via multinomial resampling using normalized weights W^j_t, and let $\{h^j_t\}_{j=1}^J$ denote the resampled draws and set $W^j_t = 1$ for $j = 1, \cdots, J$. Therefore, an approximation of $E\left[f\left(h_t, s_t\right)|Y_{1:t}\right]$ can be obtained by $\left\{h^j_t, s^j_{t|t}, P^j_{t|t}, W^j_t\right\}_{j=1}^J$.

3. **Likelihood Approximation.** The approximation of the log likelihood function is given by

$$\ln \widehat{p}\left(Y_{1:T} \,|\Theta\right) = \sum_{t=1}^T \ln\left(\frac{1}{J}\sum_{j=1}^J w^j_t\right)$$

In our implementation, we set $J = 1000$ and use 1000 rounds of simulations. In addition, following Schorfheide et al. (2018), we assume that variances of the measurement errors are one percent of variances of the corresponding moments in the data. Accordingly, we set $\sigma^a_\epsilon = 0.0015$, $\sigma^a_{d,\epsilon} = 0.0080$, $\sigma^a_{f,\epsilon} = 0.0028$, $\sigma^a_{pd,\epsilon} = 0.0450$, $\sigma^a_{vp,\epsilon} = 0.61$.

Figures & Tables

Table 1: Benchmark Calibration

Panel A: Benchmark Parameter Specification						
Preferences	γ	7.13	ρ	0.481	β	0.999
Consumption & Dividends	μ_c	0.0016	ρ_d	0.43		
	μ_d	0.0016	ϕ	3.83	φ_d	4.49
Expected Cons. Growth	ρ_x	0.9822	φ_x	0.0293		
Economic Volatility	σ	0.0073	ρ_σ	0.9987	φ_σ	$2.05E{-6}$

Panel B: Moments	Data	BKY
	Consumption & Dividends	
$\sigma(\Delta c)$	0.021	0.024
$AC1(\Delta c)$	0.473	0.430
$AC2(\Delta c)$	0.115	0.205
$\sigma(\Delta d)$	0.114	0.105
$AC1(\Delta d)$	0.198	0.391
$corr(\Delta c, \Delta d)$	0.589	0.573
	Asset Prices	
$E(pd)$	3.377	3.377
$\sigma(pd)$	0.450	0.390
$AC1(pd)$	0.948	0.913
$E(R_m - R_f)$	0.077	0.058
$\sigma(r_m)$	0.196	0.177
$E(r_f)$	0.005	0.013
	Predictability	
$corr(r_m, pd_{-1})$	−0.236	−0.056
$corr(\Delta c, pd_{-1})$	0.202	0.317
	Variance Premium	
$E(vp)$	8.285	0.781
$\sigma(vp)$	6.128	0

Note: This table reports the benchmark parameter specification in Panel A, which is based on the estimation results in Bansal et al. (2016) (see Column "Monthly" in Table VI). In Panel B, Columns "Data" and "BKY" report moments of the cash flow processes and asset prices in the data and implied by the standard LRR model using the benchmark parameter specification in Bansal et al. (2016), respectively. Furthermore, $E(\cdot)$, $\sigma(\cdot)$, $AC1(\cdot)$, and $corr(\cdot,\cdot)$ denote the mean, standard deviation, first-order autocorrelation, and correlation, respectively. Δc and Δd denote the annual log consumption and dividend growth rates, respectively. pd is the log of the annual price-dividend ratio, $r_m \equiv \log R_m$ is the continously compounded annual market return, and $r_f \equiv \log R_f$ is the logarithm of the annual risk-free rate. vp denotes the variance premium of the market return, annualized and in basis points.

Table 2: Moments in the Data and Models

Panel A: Calibrated Parameter Values				
		LRR	gLRR1	gLRR2
Risk Aversion	γ	8.548	4.700	4.576
IES	$1/\rho$	2.976	2.079	2.033
Ambiguity Aversion	η	–	32.508	30.807
Persistence of Ambiguity	ρ_τ	–	–	0.985
Volatility of Ambiguity	φ_τ	–	–	25

Panel B: Moments				
	Data	LRR	gLRR1	gLRR2
		Consumption & Dividends		
$\sigma(\Delta c)$	0.021	0.024	0.024	0.024
$AC1(\Delta c)$	0.473	0.430	0.430	0.430
$AC2(\Delta c)$	0.115	0.205	0.205	0.205
$\sigma(\Delta d)$	0.114	0.105	0.105	0.105
$AC1(\Delta d)$	0.198	0.391	0.391	0.391
$corr(\Delta c, \Delta d)$	0.589	0.573	0.573	0.573
		Asset Prices		
$E(pd)$	3.377	3.124	3.083	3.086
$\sigma(pd)$	0.450	0.424	0.320	0.322
$AC1(pd)$	0.948	0.924	0.884	0.885
$E(R_m - R_f)$	0.077	0.077	0.077	0.077
$\sigma(r_m)$	0.196	0.179	0.169	0.170
$E(r_f)$	0.005	0.005	0.005	0.005
		Predictability		
$corr(r_m, pd_{-1})$	-0.236	-0.092	-0.042	-0.046
$corr(\Delta c, pd_{-1})$	0.202	0.293	0.373	0.370
		Variance Premium		
$E(vp)$	8.285	1.107	8.285	8.285
$\sigma(vp)$	6.128	0	0	1.498

Note: This table reports calibrated values for key parameters in Panel A, and sample- and model-based moments of the cash flow processes and asset prices in Panel B. To generate model-based moments, we use the calibrated values for the parameters in Panel A in this table and the values in the benchmark specification from Panel A in Table 1 for the rest of the parameters. Furthermore, $E(\cdot)$, $\sigma(\cdot)$, $AC1(\cdot)$, and $corr(\cdot,\cdot)$ denote the mean, standard deviation, first-order autocorrelation, and correlation, respectively. Δc and Δd denote the annual log consumption and dividend growth rates, respectively. pd is the log of the annual price-dividend ratio, $r_m \equiv \log R_m$ is the continously compounded annual market return, and $r_f \equiv \log R_f$ is the logarithm of the annual risk-free rate. vp denotes the variance premium of the market return, annualized and in basis points. Columns under "LRR", "gLRR1", and "gLRR2" contain results for the standard and generalized LRR models with constant and stochastic ambiguity, respectively.

Table 3: Risk Premium vs. Ambiguity Premium under the Generalized LRR Model

Moment	Ambiguity-Neutral	Ambiguity-Averse
$E\left(R_m - R_f\right)$	0.046	0.077
$E\left(vp\right)$	1.921	8.285

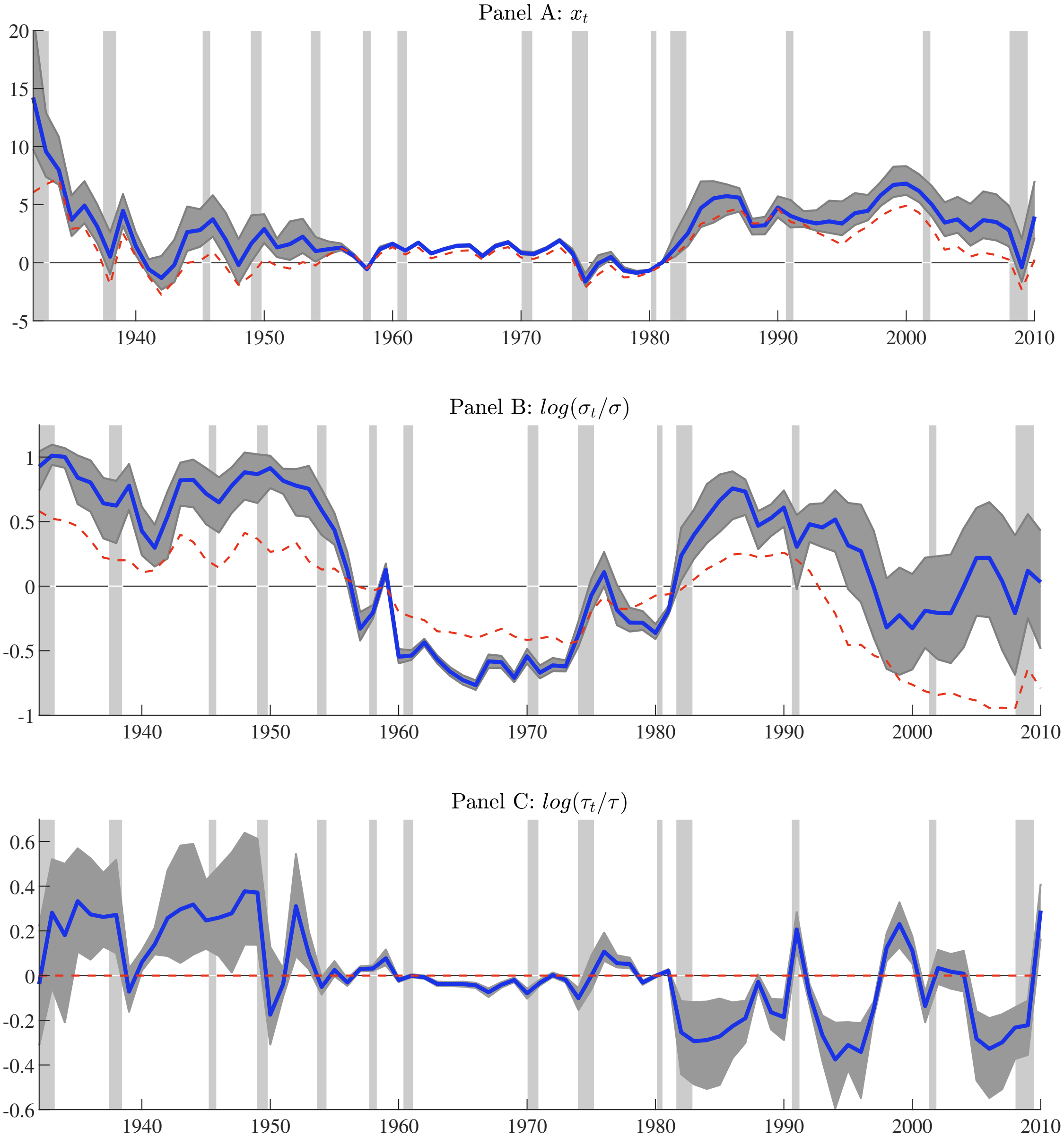
Note: This table reports both the equity premium and the variance premium implied by Case II of the generalized LRR model (i.e., gLRR2). Column "Ambiguity-Averse" reports the model-implied equity and variance premiums using the calibrated parameter values in Column "gLRR2" in Table 2 for an ambiguity-averse agent with $\eta = 30.807$ and $\gamma = 4.576$. To report the moments for an ambiguity-neutral agent, we set the coefficient of ambiguity aversion to be the same as the coefficient of risk aversion (i.e., $\eta = \gamma = 4.576$), and then report the model-implied equity and variance premiums in Column "Ambiguity-Neutral" in this table. The difference in the equity or variance premiums between these two cases represents the *ambiguity-premium component*.

Table 4: Variance Decomposition

	LRR			gLRR2		
	x_t	σ_t^2	τ_t^2	x_t	σ_t^2	τ_t^2
r_m	1.3	24.5	–	1.6	2.2	2.5
r_f	37.2	62.8	–	76.1	13.9	10.0
pd	20.7	79.3	–	74.8	22.9	2.3
vp	–	–	–	0.0	0.0	100.0

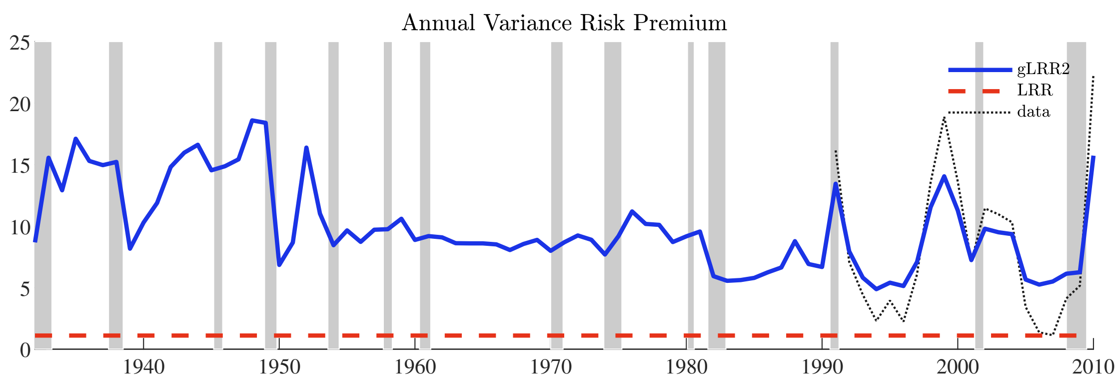
Note: This table reports the contribution of the fluctuations in growth prospects, x_t, economic volatility σ_t^2, and ambiguity τ_t^2 to the volatility of the stock market return, r_m, the risk-free rate, r_f, the price-dividend ratio, pd, and the variance premium, vp, for both the standard LRR model (see Column "LRR") and the generalized LRR model (see Column "gLRR2").

Figure 1: Estimated Latent States



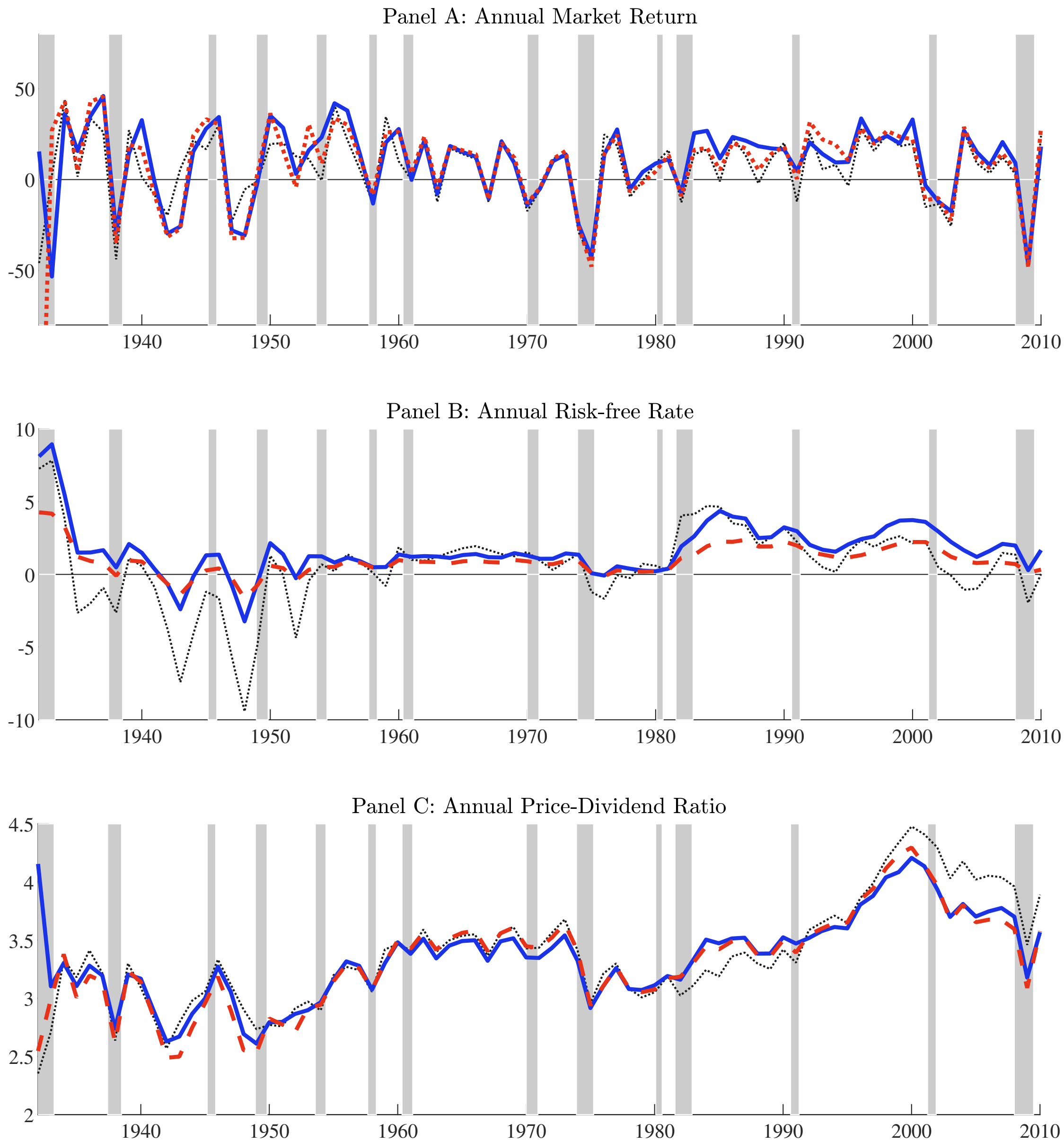
Note: This figure plots time series of the filtered latent variables with solid lines for the generalized LRR model and dashed lines for the standard LRR model, such as the long-run risks, x_t, in Panel A, economic volatility, $\log(\sigma_t/\sigma)$, in Panel B, and ambiguity, $\log(\tau_t/\tau)$, in Panel C. The latent variables are inferred from a non-linear state-space system via particle filtering (see Appendix B.2 for details).

Figure 2: The Variance Premium in the Data and Models



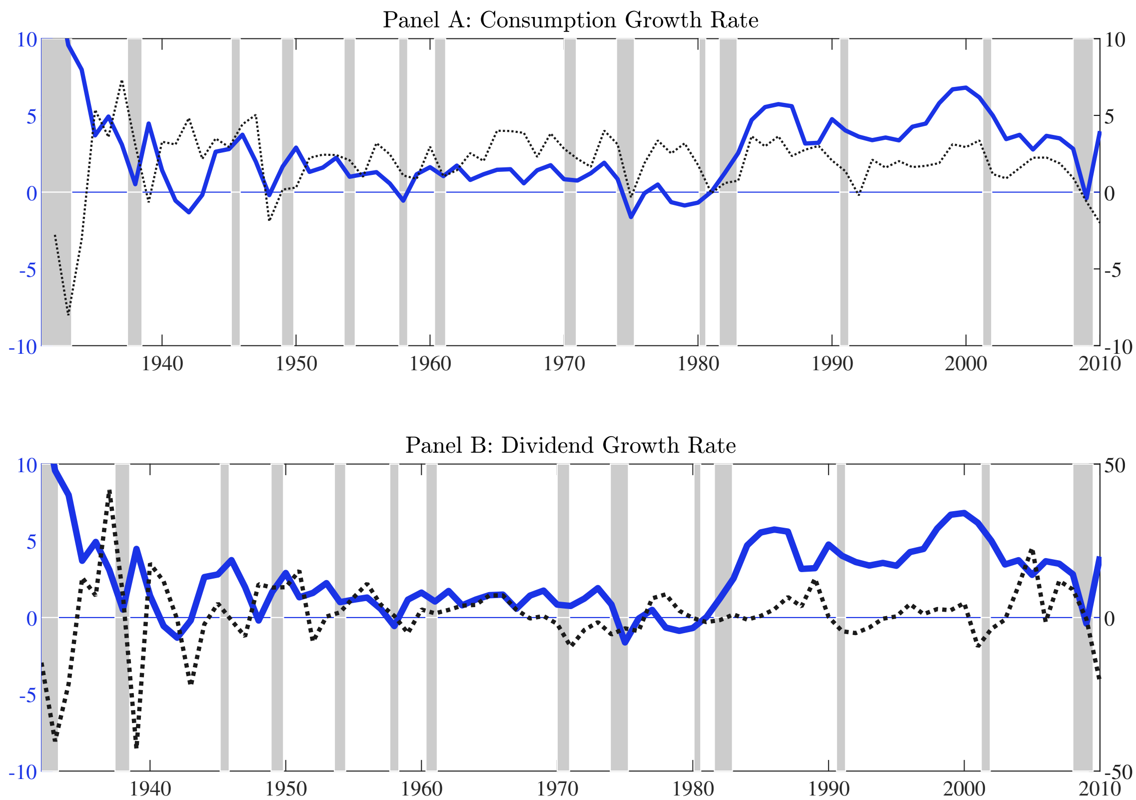
Note: This figure plots filtered time series of the variance premium between 1931 and 2009 with solid lines for the generalized LRR model, dashed lines for the standard LRR model, and dotted lines for the data.

Figure 3: Estimated Latent States



Note: This figure plots filtered time series of the stock market return (Panel A), the risk-free rate (Panel B), and the price-dividend ratio (Panel C) between 1931 and 2009 with solid lines for the generalized LRR model, dashed lines for the standard LRR model, and dotted lines for the data. The stock market return and the risk-free rate are expressed in percentages.

FIGURE 4: Estimated Latent States



NOTE: Panel A of this figure plots the realized consumption growth in the dotted line and the expected consumption growth (i.e., x_t) in the solid line (solid line) under the generalized LRR model. Panel B of this figure plots the realized dividend growth in the dotted line together with the long-run risks x_t (solid line).