Bayesian Estimation of Long-Run Risk Models
Using Sequential Monte Carlo

Andras Fulop† Jeremy Heng‡ Junye Li§ Hening Liu¶

First version: July 2019; This version: October 2020

Abstract

We propose a likelihood-based Bayesian method that exploits up-to-date sequential Monte Carlo methods to efficiently estimate long-run risk models in which the conditional variance of consumption growth follows either an autoregressive (AR) process or an autoregressive gamma (ARG) process. We use the U.S. quarterly consumption and asset returns data from the postwar period to implement estimation. Our findings are: (1) informative priors on the preference parameters can help to improve model performance; (2) expected consumption growth has a very persistent component, whereas consumption volatility is less persistent; (3) while the ARG-based model performs better than the AR-based one statistically, the latter could fit asset returns better; and (4) the solution method matters more for estimation in the AR-based model than in the ARG-based model.

Keywords: Asset Pricing, Long-Run Risk, Autoregressive Gamma Process, Log-linearization, Projection Methods, Particle Filters, Sequential Monte Carlo Sampler

JEL Classification: C11, C32, C58, E44, G12

*We thank George Tauchen (the Guest Editor) and anonymous referees for many helpful comments. Any remaining errors are ours. This work was funded by CY Initiative of Excellence (grant "Investissements d’Avenir" ANR-16-IDEX-0008).
†ESSEC Business School, Paris-Singapore. Email: fulop@essec.edu.
‡ESSEC Business School, Paris-Singapore. Email: heng@essec.edu.
§Corresponding Author. School of Management, Fudan University. Email: li.junye@fudan.edu.cn.
¶Alliance Manchester Business School. Email: hening.liu@manchester.ac.uk.

Electronic copy available at: https://ssrn.com/abstract=3573235
My view of the literature is that work “explaining the equity premium puzzle” is dying out (p.266) ... Really, the most natural thing to do with the consumption-based model is to estimate it and test it, as one would do for any economic model (p.267). — John H. Cochrane (2007)

1. Introduction

Since the seminal paper by Bansal and Yaron (2004), the long-run risk (LRR) model has attracted enormous attention (e.g., Hansen, Heaton, and Li, 2008; Bansal, Kiku, and Yaron, 2012; Beeler and Campbell, 2012, among others) and has become a benchmark in the consumption-based asset pricing literature. Models with long-run risks attempt to explain variations of asset prices by assuming a representative agent who has recursive preferences (Epstein and Zin, 1989; Weil, 1989) and persistent expected consumption growth and consumption volatility processes. Assuming an elasticity of intertemporal substitution (EIS) greater than 1, the typical long-run risk model can explain various stylized facts widely documented in the asset pricing literature including high equity premium and volatility of returns, a low and smooth risk-free rate, and return predictability by the price-dividend ratio.

Thus far, the vast majority of existing studies on long-run risks follow the calibration approach, i.e., choosing values of key parameters in a model to match a few moments of fundamentals and asset returns, lacking basis of the core theory of inference from statistics (Ghysels and Tauchen, 2016). Studies on structural estimation of long-run risk models remain very limited. Efficient econometric estimation of models with long-run risks is challenging primarily because global solutions to these models deliver asset prices that are highly non-linear functions of state variables and also because data on fundamentals are often observed in low frequency. Even though a few studies have tried to implement econometric estimation of long-run risk models using real aggregate economic and asset returns data (Bansal, Gallant, and Tauchen, 2007; Bansal, Kiku, and Yaron, 2016; Gallant, Jahan-Parvar, and Liu, 2019; Shorfheide, Song, and Yaron, 2018; Chen, Winkler, and Wasyk, 2019), these works either use moment-based econometric methods

Electronic copy available at: https://ssrn.com/abstract=3573235
or indirect inference methods, which do not fully exploit information from the likelihood function implied by the original asset pricing model, or heavily rely on the log-linearization method of Campbell and Shiller (1988) to solve for asset prices, or a combination of both approximations.\footnote{Pohl, Schmedders, and Wilms (2018) show that log-linearization of long-run risk models could generate large numerical errors when state variables are persistent. They advocate using projection methods to solve models with long-run risks to account for higher-order effects.}\footnote{Gallant (2016) explores the consequences of an assertion that the moment functions of a structural model follow a distribution and shows that if the moment functions have one of the properties of a pivotal, then this assertion coupled with a proper prior permits Bayesian inference. Gallant, Jahan-Parvar, and Liu (2019) follows this direction. However, Geweke (2016) raises some concerns about this approach and prefers direct use of the full likelihood function. Shorfheide, Song, and Yaron (2018) and this paper are two examples in this direction.} Pohl, Schmedders, and Wilms (2018) show that log-linearization of long-run risk models could generate large numerical errors when state variables are persistent. They advocate using projection methods to solve models with long-run risks to account for higher-order effects.

In this paper, we propose a likelihood-based Bayesian method that uses up-to-date sequential Monte Carlo (SMC) methods to estimate consumption-based models with long-run risks. The long-run risk model proposed by Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) assumes an autoregressive (AR) process for the conditional variance of consumption growth, which can possibly take negative values. As shown in Pohl, Schmedders, and Wilms (2018), this modeling feature, when combined with log-linearization, can lead to significant numerical bias in the solutions. To tackle this issue, we also consider an alternative, autoregressive gamma (ARG) process (Gourieroux and Jasiak, 2006), to model the conditional variance of consumption growth. This process ensures positivity of conditional variance. Furthermore, in addition to the commonly used log-linearization method, we also employ the collocation projection method to solve these two long-run risk models. As a result, we have four statistical models: two log-linearized models (LRR-AR-LL and LRR-ARG-LL) and two nonlinear models (LRR-AR and LRR-ARG).

To implement the estimation using our SMC-based Bayesian method, the first task is to deal with the dependence between the latent states and the fixed parameters. We marginalize out the former by estimating the likelihood function using a particle filter and then run a simulation routine targeting the marginal posterior distribution of the fixed parameters. One approach to achieve this aim is the pseudo-marginalization method...
pioneered by Beaumont (2003), Andrieu and Roberts (2009), and Andrieu, Doucet, and Holenstein (2010). This approach satisfies the “exact approximation” property, i.e., it targets the right posterior distribution despite the Monte Carlo error in the likelihood estimate. However, it is well-understood that central to a successful application of such an approach is the effective control of variance of the likelihood estimate (Doucet et al., 2015). Given that our models can be nonlinear and non-Gaussian, we design an efficient particle filter that approximates the optimal proposal distribution as per Doucet, Godsill, and Andrieu (2000) by running the (unscented) Kalman filter for each particle based on the past state particle and the current observations to obtain an unbiased and efficient estimate of the model likelihood. The second task is to design efficient proposals to explore the marginal distribution of the parameters. To do so, we employ the density-tempered SMC sampler based on Del Moral, Doucet, and Jasra (2006) and Duan and Fulop (2015). This method approximates a target distribution with a population of simulated points and allows one to adapt proposal moves to the simulation output in an iterative manner. We implement Monte Carlo simulation studies to show that our SMC-based Bayesian method can deliver reliable and efficient parameter estimates for the long-run risk models under consideration.

We implement model estimation using U.S. real quarterly data on consumption, dividends, market and risk-free returns ranging from 1947:Q1 to 2018:Q1. In the long-run risk models, the preference parameters in the Epstein-Zin’s utility function play an important role in determining asset prices but they are usually difficult to estimate jointly with parameters in the consumption and dividend processes. For this reason, we impose three types of priors on the risk-aversion parameter $\gamma$ and the EIS parameter $\psi$: non-informative priors, informative priors, and fixed values typically used in calibrations ($\gamma = 10$ and $\psi = 1.5$). We find that when non-informative priors are employed, the preference parameter estimates are insensitive to the variance specifications regardless of the solution methods. The risk-aversion parameter estimates in the two nonlinear models (LRR-AR and LRR-ARG) seem economically too large, whereas in the two linear models (LRR-AR-LL and LRR-ARG-LL), they are economically reasonable. Our findings
based on non-informative priors complement existing studies on the structural estimation of equilibrium asset pricing models, which commonly impose informative priors on the preference parameters (see, e.g., Shorfheide, Song, and Yaron (2018) and Gallant, Jahan-Parvar, and Liu (2019)).

However, when we introduce informative priors, the preference parameter estimates in the two nonlinear models become economically reasonable but very different across the variance specifications: the risk-aversion/EIS parameter estimate is much smaller/larger in the LRR-AR model than in the LRR-ARG model, whereas they are still quite similar in the two linear models. The long-run risk component in expected consumption growth is in general very persistent, whereas the volatility process is less persistent. While the persistence of the long-run risk component is roughly the same in the AR- and ARG-based models, the persistence of the conditional variance of consumption growth is much weaker in the ARG-based models than in the AR-based ones.

Notably, we find that the introduction of informative priors improves model performance both economically and statistically. Either imposing non-informative priors or fixing the preference parameters worsens model performance in comparison to the case of informative priors. The ARG-based long-run risk model, whether linear or nonlinear, always performs better statistically than the corresponding AR-based long-run risk model. However, we find that the LRR-AR model with informative priors performs the best economically among all the models considered, in terms of fitting market returns and risk-free rates.

Another related and interesting question is: given the observed data, how does the solution method of long-run risk models influence the estimation results? This question is still left unanswered in existing studies. We find that the estimation results for the LRR-AR model are significantly different from those of the LRR-AR-LL model. The LRR-AR model features a lower risk aversion, a higher EIS and a more persistent conditional variance process. Comparing these statistical models suggests that using global methods to solve long-run risk models can yield dramatic gains. Nevertheless, the conclusion is more obscure for the ARG-based models due to relatively weak persistence of the
estimated variance process.

In contrast to moment-based econometric methods, our SMC-based Bayesian method can directly provide filtered long-run risk and stochastic volatility components that take into account both parameter and state uncertainties. The filtered long-run risk component mainly falls down during recession periods, but also occasionally falls down in boom periods; the filtered consumption volatility clearly captures the volatility reduction in the periods of 1950’s and 1960’s and in the Great Moderation. The model-fit results indicate that in general, the long-run risk models can price persistent and less-volatile risk-free returns better than less-persistent and volatile market returns.

Our paper is closely related to Schorfheide, Song, and Yaron (2018) who consider a Bayesian approach based on a particle Markov chain Monte Carlo (MCMC) method to estimate their extended long-run risk model. However, our paper differs from their study in important aspects. First, Schorfheide, Song, and Yaron (2018) heavily rely on log-linearization and linearization of the log-volatility process to find linear functions for equilibrium asset prices, whereas we use both log-linearization and a projection method to solve for asset prices and examine the impact of the solution methods on estimation results. Second, the Bayesian method employed in Schorfheide, Song, and Yaron (2018) is specific to their extended long-run risk model, making it hard to generalize to alternative models with flexible specifications of consumption dynamics. In contrast, our SMC-based Bayesian approach is quite generic and can be used to estimate flexible asset pricing models that can be cast into the framework of nonlinear and/or non-Gaussian state-space models. Third, in Schorfheide, Song, and Yaron (2018), there exist nontrivial differences between the approximate model used to solve for asset prices and the model used in estimation. In addition, the measurement error variance of risk-free returns is fixed at 1% of the sample variance. As such, their estimation results are hard to interpret economically. In contrast, to maintain consistency, our paper uses the same model for solving asset prices and implementing estimation, and treats variances of measurement errors as free parameters.

The rest of the paper is organized as follows. Section 2 presents the long-run risk mod-
els and their solutions. Section 3 presents our Bayesian methodology based on sequential Monte Carlo methods. Section 4 provides Monte Carlo simulation studies. Section 5 presents the data, estimation results and asset pricing implications. Section 6 concludes the paper. The log-linear solution to the ARG-based model is given in the Appendix.

2. Model Framework

2.1. Preferences

We consider an endowment economy with a representative agent who has recursive preferences as in Epstein and Zin (1989) and Weil (1989). The agent maximizes her life-time utility, which is recursively given by,

\[ V_t = \left( (1 - \delta) C_t^{\frac{1-\gamma}{\psi}} + \delta \left[ E_t (V_{t+1}^{1-\gamma}) \right]^{\frac{1}{1-\gamma}} \right), \]

where \( C_t \) is consumption at time \( t \), \( 0 < \delta < 1 \) is the agent’s time preference parameter, \( \gamma \) is the relative risk aversion parameter, \( \psi \) is the elasticity of intertemporal substitution (EIS), \( \theta = \frac{1 - \gamma}{1 - 1/\psi} \), and \( E_t \) denotes conditional expectation with respect to information up to time \( t \). This class of preferences allows for a separation between risk aversion and the EIS. When \( \gamma > 1/\psi \), the agent prefers early resolution of uncertainty; when \( \gamma < 1/\psi \), she prefers late resolution of uncertainty; and when \( \theta = 1 \), she has the standard constant relative risk aversion (CRRA) preferences and is neutral to the time of resolution of uncertainty. The agent’s utility maximization is subject to the following budget constraint,

\[ W_{t+1} = (W_t - C_t) R_{w,t+1}, \]

where \( W_t \) is the wealth of the agent, and \( R_{w,t} \) is the return on the wealth portfolio.

For any asset \( i \) with ex-dividend price \( P_{i,t} \) and dividend \( D_{i,t} \), the standard Euler equation holds, i.e.,

\[ E_t [M_{t+1} R_{i,t+1}] = 1, \]

where \( R_{i,t+1} = (P_{i,t+1} + D_{i,t+1})/P_{i,t} \), and \( M_t \) is the stochastic discount factor. In particular,
for the risk-free asset, we have $R_{f,t} = 1/\mathbb{E}_t[M_{t+1}]$. For the recursive utility function defined in Equation (1), the stochastic discount factor is given by

$$M_{t+1} = \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\phi}} \left( \frac{V_{t+1}}{\mathbb{E}_t (V_{t+1}^{1-\gamma})} \right)^{\frac{1}{\psi-\gamma}}. \tag{3}$$

Epstein and Zin (1989) show that the wealth-consumption ratio, $W_t/C_t$, can be expressed in terms of the value function $V_t$,

$$\frac{W_t}{C_t} = \frac{1}{1 - \delta} \left( \frac{V_t}{C_t} \right)^{1-1/\psi}, \tag{4}$$

which allows us to reformulate the stochastic discount factor given in Equation (3) using the return on the wealth portfolio as follows,

$$M_{t+1} = \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta-1}. \tag{5}$$

Thus, the Euler equation (2) indicates that for the return on the wealth portfolio, $R_{w,t}$, we have

$$\mathbb{E}_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta-1} \right] = 1. \tag{6}$$

2.2. Fundamentals

Following Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012), we assume that the log-consumption growth, $\Delta c_{t+1} \equiv \ln \left( \frac{C_{t+1}}{C_t} \right)$, consists of a persistent component, $x_t$, and a transitory component,

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{c,t+1}, \tag{7}$$

$$x_{t+1} = \rho x_t + \phi_x \sigma_t \eta_{x,t+1}. \tag{8}$$
and that dividends are imperfectly correlated with consumption and their log-growth rate, $\Delta d_{t+1} \equiv \ln \left( \frac{D_{t+1}}{D_t} \right)$, has the following dynamics,

$$\Delta d_{t+1} = \mu_d + \Phi x_t + \phi_d \sigma \eta_{d,t+1} + \phi_d \sigma \eta_{d,t+1},$$  \hspace{1cm} (9)

where $\eta_{c,t}, \eta_{x,t},$ and $\eta_{d,t}$ are i.i.d normal $N(0,1)$, and $\sigma_t^2$ is the conditional variance of consumption growth.

In the standard long-run risk model (Bansal and Yaron, 2004; Bansal, Kiku, and Yaron, 2012, among others), the conditional variance, $\sigma_t^2$, is assumed to follow the autoregressive (AR) process,

$$\sigma_{t+1}^2 = \bar{\sigma}^2 + \nu (\sigma_t^2 - \bar{\sigma}^2) + \phi_s \eta_{s,t+1},$$  \hspace{1cm} (10)

where $\nu$ governs the persistence, $\bar{\sigma}^2$ captures the long-run mean, and $\eta_{s,t}$ is an i.i.d standard normal, independent of the other shocks in the model. We label the standard long-run risk model as “LRR-AR”.

However, the above variance process can take negative values.\(^2\) Therefore, we also exploit the non-negative autoregressive gamma (ARG) process, proposed by Gourieroux and Jasiak (2006),\(^3\) to model the conditional variance, $\sigma_t^2$,

$$\sigma_{t+1}^2 \sim \text{Gamma}(\phi_s + z_t, c), \hspace{1cm} z_t \sim \text{Poisson} \left( \frac{\nu \sigma_{t-1}^2}{c} \right),$$  \hspace{1cm} (11)

where $\text{Gamma}(\cdot)$ and $\text{Poisson}(\cdot)$ denote the gamma distribution and the Poisson distribution, respectively, $\nu$ controls the persistence, $c$ determines the scale, and to ensure positivity of conditional variances, the Feller condition, $\phi_s > 1$, needs to be satisfied. The stationary distribution of the ARG process is $\text{Gamma}(\phi_s, c/(1 - \nu))$ with the long-run mean given by $\bar{\sigma}^2 = \phi_s c/(1 - \nu)$. As shown by Gourieroux and Jasiak (2006) and Creal (2017), the transition density of $\sigma_t^2$ is a noncentral gamma distribu-

\(^2\)To avoid negative variance values, we restrict the variance process to the positive half line when solving and estimating the model.

\(^3\)Zviadadze (2017) applies the autoregressive gamma process to model consumption variance and study the term structure of currency returns.
tion and its conditional mean and variance are given by

\[ E[\sigma_t^2 | \sigma_{t-1}^2] = \bar{\sigma}^2 (1 - \nu) + \nu \sigma_{t-1}^2 \]

and

\[ \text{Var}[\sigma_t^2 | \sigma_{t-1}^2] = \frac{(1 - \nu) \bar{\sigma}^2}{\phi_s} \left( (1 - \nu) \bar{\sigma}^2 + 2 \nu \sigma_{t-1}^2 \right), \]

respectively. We label this alternative long-run risk model as “LRR-ARG”.

2.3. Model Solutions

**Log-linearization.** Bansal and Yaron (2004) first applied the log-linear approximation method of Campbell and Shiller (1988) to solve the long-run risk model. Log-linearization has been widely employed to solve models with long-run risks (see, e.g., Bansal, Kiku, and Yaron, 2012; Bansal, Kiku, and Yaron, 2016; Beeler and Campbell, 2012; Schorfheide, Song, and Yaron, 2018). We refer readers to these references for the solution to the LRR-AR model. For the LRR-ARG model, we provide the log-linearized solution in the Appendix. We follow Bansal, Kiku and Yaron (2007, 2016) and use a recursive method to numerically solve fixed-point problems and obtain long-run means of the wealth-consumption ratio and the price-dividend ratio. Bansal, Kiku and Yaron (2007, 2016) argue that the approximate analytical solutions are accurate when the recursive method is employed to solve the log-linear approximation.

**Projection Method.** In a recent paper, Pohl, Schmedders, and Wilms (2018) argue that log-linearization of models with long-run risks may generate large numerical errors when state variables are persistent. Thus, we also solve the two models using the collocation projection method (Judd, 1992) that can account for higher-order effects.

We denote the current state of the economy by \( z \) and the state of the next period by \( z' \) (i.e., \( z = \{x, \sigma^2\} \)). The models are solved in two steps. First, we use the projection method to solve the Euler equation for the wealth portfolio to obtain the wealth-consumption ratio. According to Equation (6), the solution function to the log wealth-consumption ratio, \( \varphi_w(z) \equiv \ln \left( \frac{W(z)}{C(z)} \right) \), satisfies

\[
E \left[ \exp \left( \theta \left( \ln \delta + \left( 1 - \frac{1}{\psi} \right) \Delta c(z'|z) + \varphi_w(z') - \ln (e^{\varphi_w(z)} - 1) \right) \right) \bigg| z \right] = 1, 
\]
as the log-return on the wealth portfolio is given by

\[ r_w(z'|z) \equiv \ln \left( \frac{W(z')}{W(z) - C(z)} \right) = \varphi_w(z') - \ln(e^{\varphi_w(z)} - 1) + \Delta c(z'|z). \]  

(13)

We approximate the solution function by \( \hat{\varphi}_w(z) = \sum_{k=0}^{n} \alpha_{w,k} \Lambda_k(z) \), where \( \Lambda_k(z), k = 0, \ldots, n, \) is a set of basis functions and \( \alpha_{w,k}, k = 0, \ldots, n, \) is a set of unknown coefficients to be determined.

Second, given the solution to the wealth-consumption ratio in the first step, we apply the projection method again to solve for the price-dividend ratio. According to Equations (2) and (5), the log price-dividend ratio, \( \varphi(z) \equiv \ln \left( \frac{P(z)}{D(z)} \right) \), satisfies

\[ \mathbb{E} \left[ \exp \left( \theta \ln \delta - \frac{\theta}{\psi} \Delta c(z'|z) + (\theta - 1) r_w(z'|z) + r(z'|z) \right) \right| z \] = 1,  

(14)

where \( r(z'|z) \) is the log-return on an asset with the dividend growth rate of \( \Delta d(z'|z) \),

\[ r(z'|z) = \ln \left( e^{\varphi(z')} + 1 \right) - \varphi(z) + \Delta d(z'|z). \]

(15)

As before, we approximate the solution function to the log price-dividend ratio by \( \hat{\varphi}(z) = \sum_{k=0}^{n} \alpha_k \Lambda_k(z) \), where \( \alpha_k, k = 0, \ldots, n, \) is a set of unknown coefficients to be determined.

We apply the collocation projection method and approximate the solution functions \( \varphi_w(z) \) and \( \varphi(z) \) using Chebyshev polynomials (Judd, 1998). Replacing the true solution functions by their respective approximations in Equations (12) and (14), we define two residual functions that contain the conditional expectation operator. In the LRR-AR model, given that the underlying innovation shocks are Gaussian, we use the Gauss-Hermite quadrature to compute conditional expectations, whereas in the LRR-ARG model, we use the importance sampling method to compute conditional expectations. The collocation projection method leads to a square system of nonlinear equations, which can be solved using the standard nonlinear equation solvers to obtain the coefficients \( \alpha_{w,k} \) and \( \alpha_k \).

Borovicka & Stachurski Conditions. Borovicka and Stachurski (2020) find exact necessary
and sufficient conditions for existence and uniqueness of solutions to a class of recursive utility-based models. These conditions are not necessarily fulfilled for arbitrary parameter values. In our implementation, we impose such conditions as extra restrictions on parameters when solving and estimating the models.

3. Econometric Framework

Our models of interest can be cast into the framework of nonlinear and/or non-Gaussian state-space models. There are two state variables: the long-run risk component, $x_t$, whose dynamics is given in Equation (8), and the stochastic volatility component, $\sigma_t^2$, whose dynamics is given in Equation (10) or (11). Furthermore, there are four observations including the consumption growth rates ($\Delta c_t$), the dividend growth rates ($\Delta d_t$), the market returns ($r_{m,t}$), and the risk-free returns ($r_{f,t}$). The dynamics of consumption and dividend growth rates are given in Equations (7) and (9), respectively. Assuming that market and risk-free returns are collected with measurement errors, the dynamics of market and risk-free returns are given by,

$$
\begin{align*}
    r_{m,t} &= f(x_t, \sigma_t^2, x_{t-1}, \sigma_{t-1}^2, \Delta d_t, \Theta) + \sigma_m \eta_{m,t}, \\
    r_{f,t} &= g(x_t, \sigma_t^2, \Theta) + \sigma_f \eta_{f,t},
\end{align*}
$$

where $\Theta$ denotes the parameter set of the long-run risk model, $r_{m,t}$ and $r_{f,t}$ are the observed market and risk-free returns, $\sigma_m$ and $\sigma_f$ are the standard deviations of the respective measurement errors that are assumed to follow independent standard normal distributions, and $f(\cdot)$ and $g(\cdot)$ are two functions determining the model-implied market and risk-free returns: they are linear when we use the log-linearization method to solve the models and are highly nonlinear when we use the projection method to solve the models.

For $T$ time periods, we denote all observations as $y_{1:T} = \{\Delta c_t, \Delta d_t, r_{m,t}, r_{f,t}\}_{t=1}^T$ and the latent states as $z_{1:T} = \{x_t, \sigma_t^2\}_{t=1}^T$. Our aim is to compute the joint posterior distri-
bution of parameters and latent states, \( p(\Theta, z_{1:T}|y_{1:T}) \), which can be decomposed into

\[
p(\Theta, z_{1:T}|y_{1:T}) = p(z_{1:T}|\Theta, y_{1:T})p(\Theta|y_{1:T}),
\]

(18)

where \( p(z_{1:T}|\Theta, y_{1:T}) \) solves the state smoothing problem, and \( p(\Theta|y_{1:T}) \) addresses the parameter inference task.

In the following, we propose using SMC methods to implement model estimation. Specifically, we first design an efficient particle filter that approximates the filtering distribution and provides us with an unbiased estimate of the likelihood function. We then rely on a SMC sampler to estimate the posterior distribution of the model parameters.

### 3.1. Sequential Monte Carlo for State Filtering and Likelihood Estimation

As our state-space models could be highly nonlinear and non-Gaussian, we rely on particle filters to perform state filtering given the static parameters. For notational convenience, dependence on \( \Theta \) is suppressed in most of this subsection. The basic idea is to approximate the filtering distribution \( p(z_t|y_{1:t}) \) with an empirical distribution, denoted as \( \hat{p}(z_t|y_{1:t}) \), supported on a number of particles in the state-space.

Given \( M \) samples, \( \{z_t^{(i)}; i = 1, 2, \ldots, M\} \), approximating the filtering distribution \( p(z_{t-1}|y_{1:t-1}) \) at time \( t - 1 \), the recursion

\[
p(z_t|y_{1:t}) \propto \int p(y_t|z_t, z_{t-1})p(z_t|z_{t-1})p(z_{t-1}|y_{1:t-1})dz_{t-1}
\]

(19)

prompts the following importance sampling strategy. First, draw samples \( \{z_t^{(i)}; i = 1, 2, \ldots, M\} \) from a known and easy-to-sample proposal transition density, \( m_t(z_t|z_{t-1}, y_t) \). Second, attach importance weights, \( w_t \), to account for the difference between the target and the proposal, i.e. for \( i = 1, 2, \ldots, M \),

\[
w_t^{(i)} = \frac{p(y_t|z_t^{(i)}, z_{t-1}^{(i)})p(z_t^{(i)}|z_{t-1}^{(i)})}{m_t(z_t^{(i)}|z_{t-1}^{(i)}, y_t)}.
\]

(20)

The normalized weights are given by \( W_t^{(i)} = w_t^{(i)}/\sum_{j=1}^{M} w_t^{(j)} \). Finally, to deal with the
particle degeneracy problem, we focus our computational efforts on the most promising particles by resampling from the weighted particle approximation whenever the effective sample size, $ESS_t = 1/\sum_{i=1}^{M} (W_t^{(i)})^2$, is smaller than some prespecified threshold. Writing $\{a_t^{(i)}; i = 1, 2, \ldots, M\}$ as the sampled ancestor indices, we obtain an approximation of the filtering distribution $p(z_t|y_{1:t})$ with equally weighted samples

$$\hat{p}(z_t|y_{1:t}) = \frac{1}{M} \sum_{i=1}^{M} \delta_{z_t^{(i)}}(z_t), \quad (21)$$

where $\delta_z(\cdot)$ denotes the Dirac measure centred on $z$.

Particle filters also provide an estimate of the likelihood of the observations

$$\hat{p}(y_{1:t}|\Theta) = \prod_{l=2}^{t} \hat{p}(y_l|y_{1:l-1}, \Theta)\hat{p}(y_1|\Theta), \quad (22)$$

where $\hat{p}(y_l|y_{1:l-1}, \Theta) = \frac{1}{M} \sum_{i=1}^{M} w_l^{(i)}$. Importantly, the likelihood estimate in Equation (22) produced by a particle filter is unbiased (Del Moral, 2004):

$$E[\hat{p}(y_{1:t}|\Theta)] = p(y_{1:t}|\Theta), \quad (23)$$

where the expectation is taken with respect to all random quantities generated in a particle filter.

The most commonly used particle filter is the bootstrap filter of Gordon, Salmond, and Smith (1993), which simply takes the state transition law as the proposal, i.e.

$$m_t(z_t|z_{t-1}, y_t) = p(z_t|z_{t-1}).$$

As the proposal does not take the latest information into account, this leads to poor performance when the observations are informative about the latent states. Our model has this feature as asset returns can contain rich information about the long-run risk and its stochastic volatility. Doucet, Godsill, and Andrieu (2000) show that the optimal proposal transition density minimizing the variance of the importance weights in Equation (20), has the form $m_t^{*}(x_t|x_{t-1}, y_t) = p(x_t|x_{t-1}, y_t)$, which conditions on both past states and new observations.

Although the optimal proposal density of Doucet, Godsill, and Andrieu (2000) is
typically analytically intractable, we can obtain an approximation that is locally optimal. For the log-linearized models, our approximation is based on applying the Kalman filter (KF) to each particle, and for the nonlinear models, our approximation employs the unscented Kalman filter (UKF) (van der Merwe et al., 2001; Li, 2011) to each particle. The resulting KF/UKF-based particle filter has the following algorithmic steps:

- **Step 1:** initialization at \( t = 0 \): draw a set of particles \( \{ z_0^{(i)}; i = 1, \ldots, M \} \) from the initial distribution \( p(z_0) \) and assign each particle a weight of \( 1/M \);

- **Step 2:** for time step \( t = 1, \ldots, T \) and for each particle \( i = 1, \ldots, M \):
  - for log-linearized models, run KF based on \( z_{t-1}^{(i)} \) and new observation \( y_t \) to obtain mean \( \bar{z}_t^{(i)} \) and variance \( P_t^{(i)} \);
  - for nonlinear models, run UKF based on \( z_{t-1}^{(i)} \) and new observation \( y_t \) to obtain mean \( \bar{z}_t^{(i)} \) and variance \( P_t^{(i)} \);
  - sample \( z_t^{(i)} \sim N(\bar{z}_t^{(i)}, P_t^{(i)}) \);
  - update the weight for each particle using Equation (20) and compute the normalized weight;
  - resample to obtain equally weighted new particles \( \{ z_t^{(i)}; i = 1, \ldots, M \} \).

3.2. **Sequential Monte Carlo Sampler for Parameter Estimation**

According to Bayes’ rule, the posterior distribution of model parameters is given by

\[
p(\Theta|y_{1:T}) \propto p(y_{1:T}|\Theta)p(\Theta),
\]

where the first term on the right-hand side is the likelihood and the second term is simply the prior. We take a hierarchical approach to target the posterior distribution of the fixed parameters: for a given set of model parameters \( \Theta \) drawn from some proposal, we run a particle filter to estimate the likelihood, \( p(y_{1:T}|\Theta) \), and approximate the smoothing distribution of the latent states, \( p(z_{1:T}|\Theta, y_{1:T}) \). It has been recognized in the literature that the unbiasedness property of likelihood estimates from a particle filter, given in Equation
allows one to construct pseudo-marginal methods (Beaumont, 2003; Andrieu and Roberts, 2009; Andrieu, Doucet, and Holenstein, 2010) that enables exact inference of model parameters despite the presence of estimation noise in the likelihood estimate.

The main idea of the pseudo-marginal approach is as follows. Define an auxiliary variable $u_t$, which includes all random variables generated by a particle filter at time $t$; these include all $M$ proposed states and ancestor indices from resampling. Let $\psi(u_{1:T}|\Theta, y_{1:T})$ denote the corresponding joint distribution of all auxiliary variables given fixed parameters $\Theta$. We then define the extended posterior distribution of model parameters and auxiliary variables as

$$
\tilde{p}(\Theta, u_{1:T}|y_{1:T}) \propto p(\Theta)\hat{p}(y_{1:T}|\Theta)\psi(u_{1:T}|\Theta, y_{1:T}).
$$

By the unbiasedness property of likelihood estimates from a particle filter, this extended posterior admits the desired posterior, $p(\Theta|y_{1:T})$, as a marginal distribution

$$
\int \tilde{p}(\Theta, u_{1:T}|y_{1:T})du_{1:T} = p(\Theta|y_{1:T}).
$$

Hence, we have the following exact-approximation property: if we can design a sampler whose output converges to the extended posterior $\tilde{p}(\Theta, u_{1:T}|y_{1:T})$, it will also deliver consistent approximations of its marginal distribution $p(\Theta|y_{1:T})$ for any fixed $M$, despite the estimation error in the likelihood.

There is a burgeoning literature on pseudo-marginal methods that target the extended posterior in Equation (25). One such class of algorithms is particle MCMC methods (Andrieu, Doucet, and Holenstein, 2010) that construct ergodic Markov chains with $\tilde{p}(\Theta, u_{1:T}|y_{1:T})$ as its stationary distribution. We will consider another class of algorithms, known as $SMC^2$ methods (Chopin, Jacob, and Papaspiliopoulos, 2013; Fulop and Li, 2013; Duan and Fulop, 2015), which target $\tilde{p}(\Theta, u_{1:T}|y_{1:T})$ using SMC samplers (Del Moral, Doucet, and Jasra, 2006) and can be easily parallelized. In particular, we employ the tempered marginalized SMC sampler of Duan and Fulop (2015).

The main idea of any SMC sampler (Del Moral, Doucet, and Jasra, 2006) is to begin
with an easy-to-sample distribution and traverse through a sequence of intermediate
distributions that lead to the desired target distribution. Following Duan and Fulop
(2015), we construct a sequence of $I$ distributions bridging between the extended prior
\[ \pi_1(\Theta, u_{1:T}) = p(\Theta)\psi(u_{1:T}|\Theta, y_{1:T}) \]
and the extended posterior \[ \pi_I(\Theta, u_{1:T}) = \tilde{p}(\Theta, u_{1:T}|y_{1:T}) \]
by defining
\[
\pi_i(\Theta, u_{1:T}) = \frac{\gamma_i(\Theta, u_{1:T})}{Z_i}, \tag{27}
\]
\[
\gamma_i(\Theta, u_{1:T}) = p(\Theta)\tilde{p}(y_{1:T}|\Theta)\xi_i\psi(u_{1:T}|\Theta, y_{1:T}), \tag{28}
\]
where $0 = \xi_1 < \xi_2 < \cdots < \xi_I = 1$ is a tempering sequence and $Z_i = \int \gamma_i(\Theta, u_{1:T})d(\Theta, u_{1:T})$
is the normalizing constant for the $i = 1, 2, \ldots, I$ distribution. Evidently, the last distri-
bution of the sequence is exactly the extended posterior $\tilde{p}(\Theta, u_{1:T}|y_{1:T})$.

We initialize $N$ equally weighted samples \{\(\Theta^{(n)}, u_{1:T}^{(n)}\)\}_{n=1}^N from $\pi_1(\Theta, u_{1:T})$ by sam-
ping $\Theta^{(n)} \sim p(\Theta)$ from the prior, and then running a particle filter with $M$ state par-
ticles to obtain $u_{1:T}^{(n)} \sim \psi(u_{1:T}|\Theta, y_{1:T})$ for $n = 1, 2, \ldots, N$. Given $N$ equally weighted
samples \{\(\Theta^{(n)}, u_{1:T}^{(n)}\)\}_{n=1}^N approximating the intermediate distribution $\pi_{i-1}(\Theta, u_{1:T})$, for
\(i = 2, 3, \ldots, I\), we move on to approximate $\pi_i(\Theta, u_{1:T})$ by weighting each parameter par-
ticle $\Theta^{(n)}$ by $s_i^{(n)} = \tilde{p}(y_{1:T}|\Theta^{(n)})\xi_i - \xi_{i-1}$ for $n = 1, 2, \ldots, N$. The next tempering level $\xi_i$
can be prespecified or chosen adaptively to prevent parameter particle degeneracy. Following
Del Moral, Doucet, and Jasra (2012), we set the value of $\xi_i$ to ensure that the parameter
effective sample size (ESS) is approximately some prespecified threshold. This can be im-
plemented using a simple bisection method. Thereafter, we resample from the weighted
particle approximation to obtain a new set of equally weighted samples \{\(\Theta^{(n)}, u_{1:T}^{(n)}\)\}_{n=1}^N
that approximate $\pi_i(\Theta, u_{1:T})$. We can also estimate the ratio of normalizing constants
$Z_i/Z_{i-1}$ using $\hat{Z}_i/Z_{i-1} = \frac{1}{N} \sum_{n=1}^{N} s_i^{(n)}$.

With repeated weighting and resampling, the support of the parameter particles would
gradually deteriorate, leading to the well-known particle impoverishment problem. To
resolve this issue, Gilks and Berzuini (2001) introduced particle diversity by moving
resampled particles according to a MCMC method with $\pi_i(\Theta, u_{1:T})$ as its stationary
distribution. In our context, this rejuvenation step can be achieved using a type of particle MCMC method known as particle marginal Metropolis-Hastings (PMMH). For each parameter particle, a PMMH iteration involves proposing a new parameter \( \Theta^* \), and accepting this proposal with a Metropolis-Hastings acceptance probability that requires estimating the likelihood \( p(y_{1:T} | \Theta^*) \) with a particle filter. To ensure adequate rejuvenation of the support, we perform several iterations of PMMH and measure its efficiency by monitoring the acceptance rate.

As output of the algorithm, we have a set of equally weighted samples \( \{ \Theta^{(n)}, u_{1:T}^{(n)} \}_{n=1}^N \) approximating the extended posterior \( \hat{p}(\Theta, u_{1:T} | y_{1:T}) \) and the ratio of normalizing constants estimates \( \{ \hat{Z}_i / Z_{i-1} \}_{i=2}^I \). Using the output, we can approximate expectations of the form \( E[\varphi(\Theta) | y_{1:T}] \) using the empirical average \( \hat{\varphi} = \frac{1}{N} \sum_{n=1}^N \varphi(\Theta^{(n)}) \) and estimate the marginal likelihood using \( \hat{p}(y_{1:T}) = \prod_{i=2}^I \hat{Z}_i / \hat{Z}_{i-1} \). As the above algorithm can be seen as a standard SMC sampler (Del Moral, Doucet, and Jasra, 2006) applied to the sequence of distributions (27) on the extended space, many general convergence results are available for the estimators \( \hat{\varphi} \) and \( \hat{p}(y_{1:T}) \) (Del Moral, 2004). One additional complication, however, is the use of adaptation to determine the tempering schedule \( \{ \xi_i \}_{i=1}^I \) and to construct the parameter proposals within PMMH. Although such adaptive SMC samplers can improve algorithmic efficiency and reduce the number of user-specified tuning parameters, standard convergence results do not directly apply. Thanks to recent theoretical advances in Beskos et al. (2016), under appropriate regularity assumptions, it follows that the estimators \( \hat{\varphi} \) and \( \hat{p}(y_{1:T}) \) are consistent and satisfy central limit theorems (at the usual Monte Carlo rate) as the number of parameter particles \( N \) goes to infinity, for any fixed number of state particles \( M \). Although these convergence properties hold for any \( M \), it is important to note that the choice of \( M \) impacts the efficiency of these estimators. If \( M \) is too small, large variance in the likelihood estimates may cause the acceptance rate of PMMH rejuvenation steps to be close to zero (Andrieu and Vihola, 2015). While there are theoretical guidelines on how to select \( M \) to optimize the asymptotic efficiency of estimators based on PMMH Markov chains (Doucet et al., 2015; Sherlock et al., 2015), to the best of our knowledge, analogous results for our estimators based on SMC\(^2\) are
not available. Hence in our applications, we select $M$ to ensure the reasonable acceptance rate of all PMMH rejuvenation steps.

4. Monte Carlo Simulation Studies

4.1. Choice of prior distributions

Our SMC-based Bayesian method introduced in Section 3 requires initialization from the prior distributions. Given that the likelihood function from the particle filter is a complicated nonlinear function of the fixed parameters in this approach, conjugate priors are not available. Thus, we simply assume normal distributions as priors; however, if a parameter under consideration has a bounded support, we choose a uniform distribution or a truncated normal distribution as its prior. Under these choices, simulation from the prior distribution is straightforward.

The preference parameters in the Epstein-Zin utility function play an important role in the long-run risk models. The prior on the subjective discount rate, $\delta$, captures the agent’s belief about the magnitude of the risk-free rate. Since the existing literature has determined that a relatively large value is necessary, we assume a uniform prior of $U(0.95, 0.999)$ for $\delta$. For the risk-aversion parameter, $\gamma$, and the EIS parameter, $\psi$, we consider the following three types of priors:

- Case I: non-informative priors.
- Case II: informative priors. The long-run risk literature usually advocates that the risk-aversion parameter, $\gamma$, should be smaller than 10 (Mehra and Prescott, 1985), and the EIS parameter, $\psi$, should be larger than 1, with a typical range of 1 to 3 (Hansen and Singleton, 1982). Therefore, we choose priors for these two parameters that largely cover their respective economically meaningful regions. Both studies in Shorfheide, Song, and Yaron (2018) and Gallant, Jahan-Parvar, and Liu (2019) also impose informative priors on these two parameters.
- Case III: fix $\gamma$ and $\psi$ at the values commonly used in calibration studies. The typical values for $\gamma$ and $\psi$ are 10 and 1.5, respectively. See, e.g., Bansal and Yaron
Table 1: Prior Distributions

Panel A: Preference Parameters

<table>
<thead>
<tr>
<th>Θ</th>
<th>Distr. Form</th>
<th>Support</th>
<th>Hyper-parameters</th>
<th>Case I</th>
<th>Case II</th>
<th>Case III</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ</td>
<td>Uniform</td>
<td>(0, 1)</td>
<td>(0.950, 0.999)</td>
<td>(0.950, 0.999)</td>
<td>(0.950, 0.999)</td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(8.00, 10.0)</td>
<td>(8.00, 2.00)</td>
<td>(10.00, 1e-8)</td>
<td></td>
</tr>
<tr>
<td>ψ</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(2.00, 4.00)</td>
<td>(2.00, 0.30)</td>
<td>(1.50, 1e-8)</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Other Parameters

<table>
<thead>
<tr>
<th>Θ</th>
<th>D. Form</th>
<th>Support</th>
<th>Hyper</th>
<th>Θ</th>
<th>D. Form</th>
<th>Support</th>
<th>Hyper</th>
</tr>
</thead>
<tbody>
<tr>
<td>µ</td>
<td>Normal</td>
<td>(-∞, ∞)</td>
<td>(µ, 1e-8)</td>
<td>µ_d</td>
<td>Normal</td>
<td>(-∞, ∞)</td>
<td>(µ_d, 1e-8)</td>
</tr>
<tr>
<td>ρ</td>
<td>Uniform</td>
<td>(-1, 1)</td>
<td>(0, 1.00)</td>
<td>Φ</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(2.00, 4.00)</td>
</tr>
<tr>
<td>φ_x</td>
<td>Uniform</td>
<td>(0, 1)</td>
<td>(0, 1.00)</td>
<td>φ_d</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(2.00, 4.00)</td>
</tr>
<tr>
<td>σ</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(0.004, 0.005)</td>
<td>φ_d</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(5.00, 6.00)</td>
</tr>
<tr>
<td>ν</td>
<td>Uniform</td>
<td>(-1, 1)</td>
<td>(0, 1.00)</td>
<td>σ_m</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(0.03, 0.10)</td>
</tr>
<tr>
<td>φ_s (AR)</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(5e-6, 1e-3)</td>
<td>σ_f</td>
<td>Tr. Normal</td>
<td>(0, ∞)</td>
<td>(0.003, 0.01)</td>
</tr>
<tr>
<td>φ_s (ARG)</td>
<td>Tr. Normal</td>
<td>(1, ∞)</td>
<td>(2.00, 4.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table presents the exact distributional form, the support, and the hyper-parameters of the prior distribution for each parameter. We assume normal distributions as priors; however, if a parameter under consideration has a bounded support, we choose a truncated normal distribution or a uniform distribution as its prior. We consider three types of priors on γ and ψ: Case I: noninformative priors, Case II: informative priors, and Case III: fix γ and ψ at calibration values used in the literature.

(2004), Bansal, Kiku, and Yaron (2012), and Beeler and Campbell (2012).

Furthermore, for the two parameters governing the unconditional means of consumption and dividend growth rates, µ and µ_d, we use informative priors, i.e., prior means are given by their respective empirical averages, and prior standard deviations are given by very small values. For the other parameters, we simply assume non-informative priors, especially for the two parameters governing persistence of the long-run risk component and the stochastic volatility process. Borovicka and Stachurski (2020) provide necessary and sufficient conditions for existence and uniqueness of solutions of a class of models with homothetic recursive utility. For each set of parameters, we compute the test value provided by Borovicka and Stachurski (2020) and check whether their condition for the models considered in this paper is satisfied. Any parameter set that violates the condition is given negative infinity as likelihood value. Table 1 presents the exact distributional form, the support, and the hyper-parameters of the prior distribution for each parameter.

When bridging between the prior and posterior distributions, we adopt Gaussian
mixture proposals with a fixed number of components in the PMMH steps. Each Gaussian mixture distribution is fitted on the existing parameter particle population that satisfies the Borovicka and Stachurski (2020) condition using the expectation-maximization (EM) algorithm. We note that the ability to adapt proposal distributions on the fly using information from intermediate target densities is the key to the efficiency of our algorithm and also a key advantage of the sequential Monte Carlo framework. Because of this, we find that in both simulation studies and real data application, almost all proposed parameter particles would satisfy the Borovicka and Stachurski (2020) condition after several steps of resample-move.

4.2. Monte Carlo on Parameter Estimation

We now move to Monte Carlo studies on parameter estimation, which aim to show that our SMC-based Bayesian method delivers reliable and efficient parameter estimates. We take the LRR-AR model, which is nonlinear and Gaussian, and the LRR-ARG model, which is nonlinear and non-Gaussian and may be more challenging, as our examples. From each model, we generate 100 sequences of quarterly observations for market returns, risk-free returns, consumption growth rates, and dividend growth rates, with a sample size of 600. The true values of model parameters used to generate the data are given in Table 2.

We run our SMC-based Bayesian method on each simulated dataset for each model. To alleviate the small sample size issue, we use the priors of Case II in Table 1 to initialize algorithm and a Gaussian mixture with 8 components as the proposal within PMMH. We set the number of parameter particles, $N$, equal to 1,024, and the number of state particles, $M$, equal to 200. At each distribution-bridging step, we select the next tempering level so that when the parameter ESS falls slightly below $N/2$, we resample from the weighted particle approximation and rejuvenate the particles using 5 PMMH iterations.

Table 2 presents the results of the simulation study for each model. We report the mean/median of the posterior means for each parameter and its root mean square error.
Table 2: Monte Carlo Simulation Studies

<table>
<thead>
<tr>
<th>Θ</th>
<th>LRR-AR Mean</th>
<th>LRR-ARG Mean</th>
<th>LRR-AR Median</th>
<th>LRR-ARG Median</th>
<th>LRR-AR RMSE</th>
<th>LRR-ARG RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ</td>
<td>0.9950</td>
<td>0.9950</td>
<td>0.9951</td>
<td>0.0003</td>
<td>0.9951</td>
<td>0.9951</td>
</tr>
<tr>
<td>γ</td>
<td>8.0000</td>
<td>8.4322</td>
<td>8.5783</td>
<td>1.0446</td>
<td>8.4547</td>
<td>8.4199</td>
</tr>
<tr>
<td>ψ</td>
<td>2.0000</td>
<td>2.0543</td>
<td>2.0753</td>
<td>0.1429</td>
<td>2.1172</td>
<td>2.1031</td>
</tr>
<tr>
<td>μ</td>
<td>0.0050</td>
<td>0.0050</td>
<td>0.0050</td>
<td>0.0000</td>
<td>0.0050</td>
<td>0.0050</td>
</tr>
<tr>
<td>ρ</td>
<td>0.9800</td>
<td>0.9714</td>
<td>0.9732</td>
<td>0.0134</td>
<td>0.9488</td>
<td>0.9594</td>
</tr>
<tr>
<td>φ_x</td>
<td>0.1000</td>
<td>0.1175</td>
<td>0.1166</td>
<td>0.0240</td>
<td>0.1677</td>
<td>0.1401</td>
</tr>
<tr>
<td>σ</td>
<td>0.0050</td>
<td>0.0050</td>
<td>0.0050</td>
<td>0.0002</td>
<td>0.0052</td>
<td>0.0051</td>
</tr>
<tr>
<td>ν</td>
<td>0.9000</td>
<td>0.8255</td>
<td>0.8630</td>
<td>0.1373</td>
<td>0.8309</td>
<td>0.8646</td>
</tr>
<tr>
<td>φ_s (AR)</td>
<td>3.50E-6</td>
<td>4.01E-6</td>
<td>3.95E-6</td>
<td>1.07E-6</td>
<td>——</td>
<td>——</td>
</tr>
<tr>
<td>φ_s (ARG)</td>
<td>2.0000</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>1.5195</td>
<td>1.6877</td>
</tr>
<tr>
<td>μ_d</td>
<td>0.0050</td>
<td>0.0050</td>
<td>0.0046</td>
<td>0.0006</td>
<td>0.0050</td>
<td>0.0050</td>
</tr>
<tr>
<td>Φ</td>
<td>2.5000</td>
<td>2.5159</td>
<td>2.4833</td>
<td>0.3661</td>
<td>2.5105</td>
<td>2.4522</td>
</tr>
<tr>
<td>φ_{dc}</td>
<td>2.5000</td>
<td>2.5177</td>
<td>2.5075</td>
<td>0.2491</td>
<td>2.4789</td>
<td>2.4753</td>
</tr>
<tr>
<td>φ_d</td>
<td>5.0000</td>
<td>5.0235</td>
<td>5.0189</td>
<td>0.2245</td>
<td>5.2361</td>
<td>5.3781</td>
</tr>
<tr>
<td>σ_m</td>
<td>0.0300</td>
<td>0.0327</td>
<td>0.0325</td>
<td>0.0042</td>
<td>0.0360</td>
<td>0.0362</td>
</tr>
<tr>
<td>σ_f</td>
<td>0.0030</td>
<td>0.0030</td>
<td>0.0030</td>
<td>0.0001</td>
<td>0.0030</td>
<td>0.0030</td>
</tr>
</tbody>
</table>

Note: This table present the results of the Monte Carlo simulation studies on parameter estimation. We generate 100 sequences of quarterly observations, with a sample size of 600, for market returns, risk-free returns, consumption growth rates, and dividend growth rates from the LRR-AR and LRR-ARG models, respectively, using the true parameters reported in the table. For each dataset in each model, we run our SMC-based Bayesian method to estimate the fixed parameters. The parameter estimates are taken to be the posterior means. The table reports the mean parameter estimates across the simulation runs and the root mean square errors of these estimates compared to their true values.

For most of the parameters in both models, we find that the mean/median is close to the corresponding true value and the RMSE is small. This is particularly true for the preference parameters (δ, γ, and ψ) and the long-run component parameters (ρ and φ_x). We also find that the two measurement error standard deviations, σ_m and σ_f, can also be well-identified.

5. Empirical Results and Asset Pricing Implications

5.1. Data

For model estimation, we consider the U.S. real quarterly data from 1947:Q1 to 2018:Q1. Following Bansal, Kiku, and Yaron (2012) and Beeler and Campbell (2012), we measure real consumption as the sum of seasonally adjusted real nondurable and services con-
sumption per capita from the Bureau of Economic Analysis. The consumption growth rates are then constructed by taking the first difference of log-real consumption. Our measure of the real risk-free rates is constructed from the nominal 90-day U.S. Treasury bill rates and the quarterly consumer price index (CPI), obtained from the Wharton Research Data Services (WRDS) Treasury and Inflation database. We first construct the \textit{ex post} real risk-free rates by deflating the nominal T-bill rates using the CPI inflation rates, and then regress the \textit{ex post} real risk-free rates on the nominal rates and past inflation rates. The predicted values from this regression yield the \textit{ex ante} risk-free rates.

Our stock market returns are proxied by the value-weighted returns on the stock portfolio of NYSE/AMEX/NASDAQ ("VWRETD", including dividends) from the Center for Research in Security Prices (CRSP). We construct real stock returns by deflating nominal returns using the CPI inflation rates. Lastly, we construct the dividend growth rates from the difference between the value-weighted returns including and excluding dividends (VWRETD and VWRETX). As usual, we smooth out dividend series by aggregating four quarters’ (including the current quarter) values of the dividend index series. We then obtain the real dividend growth rates by deflating the nominal rates using the inflation rates. Our construction method closely follows Robert Shiller and the consumption-based long-run risk studies such as Bansal and Yaron (2004) and Schorfheide, Song and Yaron (2018). The contemporaneous correlation between consumption growth and dividend growth, both of which are sampled at quarterly frequency, is about 0.16 in the data. Bansal, Kiku, and Yaron (2012) report a correlation of 0.55 for the annual data. The difference arises mainly because we consider a different data frequency and focus on the post-war sample.

Table 3 reports summary statistics of the data. The mean annualized real consumption growth rate is about 1.9% and its annualized standard deviation is about 1.0%. In contrast, the real dividend growth rate has a higher mean and is more volatile than the real consumption growth rate: its mean is about 2.5% and its standard deviation is about 5.2%. The average annualized real market return is about 7% and its annualized volatility is about 16.3%, whereas the mean annualized real risk-free rate is less than 1% and its
Table 3: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>AC(1)</th>
<th>JB Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta c_t$</td>
<td>1.877</td>
<td>0.996</td>
<td>-0.384</td>
<td>4.213</td>
<td>0.294</td>
<td>0.001</td>
</tr>
<tr>
<td>$\Delta d_t$</td>
<td>2.539</td>
<td>5.173</td>
<td>0.692</td>
<td>8.543</td>
<td>0.331</td>
<td>0.001</td>
</tr>
<tr>
<td>$r_{m,t}$</td>
<td>6.975</td>
<td>16.26</td>
<td>-0.859</td>
<td>4.598</td>
<td>0.079</td>
<td>0.001</td>
</tr>
<tr>
<td>$r_{f,t}$</td>
<td>0.819</td>
<td>1.548</td>
<td>-1.450</td>
<td>10.23</td>
<td>0.821</td>
<td>0.001</td>
</tr>
<tr>
<td>$r_{m,t} - r_{f,t}$</td>
<td>6.157</td>
<td>16.23</td>
<td>-0.881</td>
<td>4.612</td>
<td>0.078</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Note: The table presents summary statistics of the data used for model estimation. The data include the consumption growth rates ($\Delta c_t$), the dividend growth rates ($\Delta d_t$), the market returns ($r_{m,t}$), and the risk-free returns ($r_{f,t}$). The summary statistics of the equity risk premia ($r_{m,t} - r_{f,t}$) are also included. Mean and Std are annualized and in percentage. JB Test represents the Jarque-Bera normality test. The data are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters.

Annualized volatility is only about 1.5%. The equity risk premium is on average about 6.2% and its volatility is about 16.2%. All variables except the dividend growth rate are left-skewed, and all of them are leptokurtic. The risk-free returns are persistent with the first-order autocorrelation of 0.8, whereas the market returns are not persistent with the first-order autocorrelation of only 0.08. The autocorrelations of the consumption and dividend growth rates are 0.29 and 0.33, respectively. The Jarque-Bera tests reject the null hypothesis of normality for all series (with very small p-values). Figure 1 plots the time-series data of all variables.

5.2. Model Estimation

5.2.1. Statistical Analysis and Diagnosis

In estimation using real data, we set the number of parameter particles ($N$) equal to 1,024 and the number of state particles ($M$) equal to 200. We select the tempering schedule so that the parameter ESS falls slightly below $N/2$ (at each distribution-bridging step), in which case, we resample and perform 10 PMMH iterations to ensure adequate rejuvenation of the parameter supports. Following Bansal, Kiku, and Yaron (2012) and Schorfheide, Song, and Yaron (2018), we fix the quarterly subjective discount rate $\delta$ at 0.9968, which is equivalent to the monthly value of 0.9989 used by Bansal, Kiku, and Yaron (2012).\footnote{Schorfheide, Song, and Yaron (2018) fix the monthly subjective discount rate at 0.999 in their estimation.} We estimate all the long-run risk models relying on the priors presented...
An important quantity in our proposed methodology is the acceptance rate, which measures the efficiency of each rejuvenation step. Figure 2 display the acceptance rate of the last PMMH iteration at each bridging distribution when estimating the nonlinear AR-based long-run risk model based on the three different types of priors on $\gamma$ and $\psi$. We observe that the acceptance rate is quite high before the tempering coefficient, $\xi$, reaches the level of about 0.50, and thereafter, it declines a little and remains high (around 0.20) when $\xi$ approaches 1 in Case II and Case III; however, whenever we use the non-informative priors on $\gamma$ and $\psi$ in Case I, the acceptance rate is less than 0.10 when $\xi$ reaches 1. Also note that as usual for SMC samplers, that smaller increments in tempering levels, and hence more resampling and rejuvenation steps, are taken at the initial stages of the algorithm.

Figure 3 illustrates the sequence of bridging distributions for each parameter when
Case I

Case II

Case III

Figure 2: Acceptance Rates

Note: This figure plots the acceptance rate of the last PMMH iteration at each bridging distribution determined by $\xi$ for the three types of priors on $\gamma$ and $\psi$ considered. The data are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters.

the non-informative priors on $\gamma$ and $\psi$ are used in the nonlinear AR-based long-run risk model. The means and (5, 95)% quantiles of each bridging distribution are plotted. We can see that in the beginning when the tempering level is equal to zero, we only have prior information given by distributions with large dispersions. As the tempering procedure progresses, information from the data is slowly reflected in the estimates. This is evident from the shrinkage of the (5, 95)% credible intervals. At the end of the procedure when the tempering level reaches one, we obtain the posterior distributions of all model parameters. All parameters have narrow posterior (5, 95)% credible intervals. Notably, we observe that almost all parameter estimates reach their respective stable values after the tempering level becomes larger than 0.6, indicating that they can be well-identified by our SMC-based Bayesian method.

We further examine how the Monte Carlo error in the particle filter affects the parameter estimates in our SMC-based Bayesian method. To proceed, we estimate the
nonlinear AR-based long-run risk model using the non-informative priors on $\gamma$ and $\psi$ for 15 times. Table 4 presents means and standard deviations of the posterior means of the model parameters resulting from those 15 runs. We see that for all model parameters, the standard deviations of the posterior means are very small relative to the means of the posterior means. This result suggests that the Monte Carlo error in the particle filter has little effect and our estimates are stable and robust.

### 5.2.2. Parameter Estimates of the Nonlinear Models

In Table 5, we present parameter estimates of the nonlinear AR- and ARG-based long-run risk models (LRR-AR and LRR-ARG) using the three types of priors introduced in
Table 4: Stability of Parameter Estimates

<table>
<thead>
<tr>
<th>Θ</th>
<th>Mean</th>
<th>Std</th>
<th>Θ</th>
<th>Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ</td>
<td>15.88</td>
<td>0.869</td>
<td>φ_{sx}</td>
<td>5.9e-6</td>
<td>5.0e-7</td>
</tr>
<tr>
<td>ψ</td>
<td>1.247</td>
<td>0.019</td>
<td>μ_{d}</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>μ</td>
<td>0.005</td>
<td>0.000</td>
<td>Φ</td>
<td>0.966</td>
<td>0.024</td>
</tr>
<tr>
<td>ρ</td>
<td>0.984</td>
<td>0.003</td>
<td>φ_{dc}</td>
<td>1.236</td>
<td>0.232</td>
</tr>
<tr>
<td>φ_{x}</td>
<td>0.280</td>
<td>0.032</td>
<td>φ_{d}</td>
<td>5.278</td>
<td>0.185</td>
</tr>
<tr>
<td>σ</td>
<td>0.005</td>
<td>0.000</td>
<td>σ_{m}</td>
<td>7.9e-2</td>
<td>1.4e-3</td>
</tr>
<tr>
<td>ν</td>
<td>0.792</td>
<td>0.012</td>
<td>σ_{f}</td>
<td>4.4e-3</td>
<td>3.0e-4</td>
</tr>
</tbody>
</table>

*Note:* The table presents means and standard deviations of the posterior means of all parameters, resulting from 15 runs of our SMC-based Bayesian method for estimating the LRR-AR model using the noninformative priors on γ and ψ. The data used for model estimation include the consumption growth rates (∆c_t), the dividend growth rates (∆d_t), the market returns (r_{m,t}), and the risk-free returns (r_{f,t}). They are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters.

Table 1. We first examine the estimates in Case I, which exploits the non-informative priors on γ and ψ. We find that both models have very similar estimates of the preference parameters: the posterior mean of the risk-aversion parameter, γ, is about 15.93 (1.35) in the LRR-AR model and is about 15.36 (2.72) in the LRR-ARG model, and the posterior mean of the EIS parameter, ψ, is about 1.23 (0.03) in the LRR-AR model and is about 1.28 (0.07) in the LRR-ARG model. This suggests that the preference parameter estimates are insensitive to the volatility specifications.

The long-run risk component, x_t, is slightly more persistent in the LRR-AR model than in the LRR-ARG model, as the posterior mean of the persistence parameter, ρ, is about 0.983 (0.002) in the former and is about 0.967 (0.01) in the latter. The variance process, σ^2_t, is much more persistent in the LRR-AR model than in the LRR-ARG model: the posterior mean of ν is about 0.80 (0.02) in the former and is only about 0.65 (0.05) in the latter. The dividend growth rate dynamics are similar in both models.

As for model performance, we find that the LRR-ARG model fits both market and risk-free returns slightly better than the LRR-AR model does: the posterior means of σ_{m} and σ_{f} are, respectively, 0.0776 and 0.0043 in the LRR-ARG model and 0.0784 and 0.0045 in the LRR-AR model. Furthermore, the log-marginal likelihood estimate (LMLH) is about 3163.6 in the LRR-ARG model, whereas it is about 3132.8 in the LRR-AR model, suggesting that the LRR-ARG model performs statistically better than the LRR-AR.
Table 5: Parameter Estimates of the Nonlinear Models

<table>
<thead>
<tr>
<th></th>
<th>Case I</th>
<th></th>
<th>Case II</th>
<th></th>
<th>Case III</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AR</td>
<td>ARG</td>
<td>AR</td>
<td>ARG</td>
<td>AR</td>
<td>ARG</td>
</tr>
<tr>
<td>Θ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>15.926</td>
<td>15.357</td>
<td>9.751</td>
<td>36.83</td>
<td>10.000</td>
<td>10.000</td>
</tr>
<tr>
<td></td>
<td>(1.3476)</td>
<td>(2.7227)</td>
<td>(0.3730)</td>
<td>(1.3131)</td>
<td>(——)</td>
<td>(——)</td>
</tr>
<tr>
<td>ψ</td>
<td>1.2341</td>
<td>1.2796</td>
<td>2.759</td>
<td>1.3245</td>
<td>1.5000</td>
<td>1.5000</td>
</tr>
<tr>
<td></td>
<td>(0.0319)</td>
<td>(0.0730)</td>
<td>(0.2218)</td>
<td>(0.0962)</td>
<td>(——)</td>
<td>(——)</td>
</tr>
<tr>
<td>μ</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.9829</td>
<td>0.9671</td>
<td>0.9853</td>
<td>0.9786</td>
<td>0.9984</td>
<td>0.9710</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.0108)</td>
<td>(0.0025)</td>
<td>(0.0053)</td>
<td>(0.0010)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>φ</td>
<td>0.2859</td>
<td>0.3087</td>
<td>0.3252</td>
<td>0.3214</td>
<td>0.2405</td>
<td>0.3300</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.0451)</td>
<td>(0.0356)</td>
<td>(0.0422)</td>
<td>(0.0245)</td>
<td>(0.0492)</td>
</tr>
<tr>
<td>σ</td>
<td>0.0048</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0048</td>
<td>0.0049</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>ν</td>
<td>0.8010</td>
<td>0.6508</td>
<td>0.9067</td>
<td>0.7017</td>
<td>0.8134</td>
<td>0.7265</td>
</tr>
<tr>
<td></td>
<td>(0.0174)</td>
<td>(0.0496)</td>
<td>(0.0116)</td>
<td>(0.0406)</td>
<td>(0.0099)</td>
<td>(0.0360)</td>
</tr>
<tr>
<td>φ</td>
<td>5.84E-06</td>
<td>1.6427</td>
<td>3.27E-06</td>
<td>1.4973</td>
<td>4.39E-06</td>
<td>1.4687</td>
</tr>
<tr>
<td></td>
<td>(4.98E-07)</td>
<td>(0.1922)</td>
<td>(2.57E-07)</td>
<td>(0.1639)</td>
<td>(3.23E-07)</td>
<td>(0.1584)</td>
</tr>
<tr>
<td>µ</td>
<td>0.0063</td>
<td>0.0063</td>
<td>0.0063</td>
<td>0.0063</td>
<td>0.0063</td>
<td>0.0063</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>Φ</td>
<td>0.9605</td>
<td>0.9520</td>
<td>0.9610</td>
<td>0.9248</td>
<td>0.9580</td>
<td>0.9085</td>
</tr>
<tr>
<td></td>
<td>(0.0445)</td>
<td>(0.0877)</td>
<td>(0.1107)</td>
<td>(0.0668)</td>
<td>(0.2281)</td>
<td>(0.0730)</td>
</tr>
<tr>
<td>φ</td>
<td>1.7035</td>
<td>1.1219</td>
<td>0.9680</td>
<td>1.1064</td>
<td>1.1254</td>
<td>1.1159</td>
</tr>
<tr>
<td></td>
<td>(0.3098)</td>
<td>(0.3352)</td>
<td>(0.2823)</td>
<td>(0.3449)</td>
<td>(0.3251)</td>
<td>(0.3633)</td>
</tr>
<tr>
<td>σ</td>
<td>5.2540</td>
<td>5.4879</td>
<td>5.7569</td>
<td>5.5749</td>
<td>5.2115</td>
<td>5.6236</td>
</tr>
<tr>
<td></td>
<td>(0.2145)</td>
<td>(0.3147)</td>
<td>(0.2723)</td>
<td>(0.3329)</td>
<td>(0.2127)</td>
<td>(0.3359)</td>
</tr>
<tr>
<td>σ_m</td>
<td>0.0784</td>
<td>0.0776</td>
<td>0.0663</td>
<td>0.0722</td>
<td>0.0731</td>
<td>0.0773</td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0030)</td>
<td>(0.0043)</td>
<td>(0.0032)</td>
<td>(0.0033)</td>
<td>(0.0030)</td>
</tr>
<tr>
<td>σ_f</td>
<td>0.0045</td>
<td>0.0043</td>
<td>0.0036</td>
<td>0.0043</td>
<td>0.0036</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td>(0.0002)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>LMLH</td>
<td>3132.8</td>
<td>3163.6</td>
<td>3150.5</td>
<td>3162.7</td>
<td>3155.4</td>
<td>3164.8</td>
</tr>
<tr>
<td>LLH</td>
<td>3174.6</td>
<td>3194.5</td>
<td>3210.0</td>
<td>3198.1</td>
<td>3199.8</td>
<td>3188.8</td>
</tr>
</tbody>
</table>

Note: This table presents the posterior means and standard deviations of model parameters for LRR-AR and LRR-ARG resulting from our SMC-based Bayesian method that is initialized using the priors given in Table 1 for the three cases considered in the paper. Case I refers to the noninformative priors on γ and ψ; Case II refers to the informative priors on on γ and ψ; and Case III refers to fixing the values of γ and ψ. LMLH represents the log-marginal likelihood and LLH the log-likelihood at the posterior means. The data used for model estimation include the consumption growth rates (Δc_t), the dividend growth rates (Δd_t), the market returns (r_m,t), and the risk-free returns (r_f,t). They are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters.

However, the risk-aversion parameter estimates in both models are economically too large in Case I. Therefore, we reestimate these two models by imposing the informative model does.

Electronic copy available at: https://ssrn.com/abstract=3573235
priors on $\gamma$ and $\psi$ as in Case II. Now we find that the preference parameter estimates become reasonable and drastically different in the two models: the posterior mean of $\gamma$ is about 5.98 (0.37) in the LRR-AR model, whereas it is larger in the LRR-ARG model, 9.37 (1.31), and the posterior mean of $\psi$ is about 2.48 (0.22) in the LRR-AR model, larger than that in the LRR-ARG model, 1.32 (0.10). Using a Bayesian indirect inference method and relying on the informative priors, Gallant, Jahan-Parvar, and Liu (2019) estimate the same nonlinear LRR-AR model using the U.S. annual data and find an estimate of $\gamma$ of about 8.4 and an estimate of $\psi$ of about 1.7.

The long-run risk component is almost as persistent as in Case I in the LRR-AR model, whereas it becomes moderately more persistent than in Case I in the LRR-ARG model. For both models, the persistence of the variance process becomes stronger in Case II than in Case I: the posterior mean of $\nu$ is about 0.907 (0.012) in the LRR-AR model and is about 0.702 (0.041) in the LRR-ARG model; similar to Case I, the variance process is much less persistent in the LRR-ARG model than in the LRR-AR model.

Interestingly, we find that when the informative priors are imposed $\gamma$ and $\psi$, the LRR-AR model fits both market and risk-free returns better than the LRR-ARG model does: the posterior means of $\sigma_m$ and $\sigma_f$ are, respectively, 0.0663 and 0.0036 in the LRR-AR model whereas 0.0772 and 0.0043 in the LRR-ARG model. However, statistically speaking, the LRR-ARG model is still better than the LRR-AR model as the log-marginal likelihood estimate (LMLH) is about 3162.7 in the LRR-ARG model while 3150.5 in the LRR-AR model. Furthermore, we find that by imposing the informative priors on $\gamma$ and $\psi$, the performance of each model improves both economically and statistically. For both models, the estimates of $\sigma_m$ and $\sigma_f$ become smaller in Case II than in Case I. Given that the marginal likelihood is very sensitive to the priors used, we use the log likelihood estimate (LLH) evaluated on the parameter posterior means for each model to compare its statistical performance between the two different cases. We notice that the log-likelihood of LRR-AR is 3210.0 in Case II, larger than that in Case I, 3174.6; and the log-likelihood of LRR-ARG is 3198.1 in Case II, also larger than that in Case I, 3194.5.

Finally, we investigate how the two models perform when we fix the preference pa-
Figure 4: Posterior Distributions

Note: The figure plots the posterior distributions of the preference parameter, $\gamma$ and $\psi$, and the persistence parameters, $\rho$ and $\nu$, for the three cases considered in the paper. The left panels are for LRR-AR, and the right panels for LRR-ARG.

Parameters, $\gamma$ and $\psi$, at the values commonly used in calibrations, 10 and 1.5, respectively (Case III). For each model, even though the other parameter estimates are largely similar, the model performance worsens modestly with comparison to that in Case II both economically and statistically.

Figure 4 presents the posteriors distributions of the preference parameters ($\gamma$ and $\psi$) and the persistence parameters ($\rho$ and $\nu$) resulting from all three cases. Clearly, we see that priors can have significant impact on parameter estimates and such impact seems more dramatic when estimating the LRR-AR model compared to the LRR-ARG model.
5.2.3. Parameter Estimates of the Log-linearized Models

We now take a look at parameter estimates of the log-linearized models. Table 6 presents parameter estimates of the linear AR- and ARG-based long-run risk models (LRR-AR-LL and LRR-ARG-LL) using the three types of priors. We have some interesting findings.

First, even with the non-informative priors in Case I, the preference parameter estimates in both linearized models are economically reasonable. For example, the posterior mean of the risk-aversion parameter, $\gamma$, is about 8.76 (2.88) in the LRR-AR-LL model and is about 8.60 (1.65) in the LRR-ARG-LL model, and the posterior mean of the EIS parameter, $\psi$, is about 1.18 (0.05) in the LRR-AR-LL model and is about 1.21 (0.04) in the LRR-ARG-LL model. Again as in the nonlinear models, the preference parameter estimates are insensitive to the volatility specifications. Using moment-based estimation methods, Bansal, Gallant and Tauchen (2007) and Bansal, Kiku, and Yaron (2016) estimate the similar LRR-AR-LL model. Bansal, Gallant and Tauchen (2007) fix $\psi$ at 2 and obtain an estimate of $\gamma$ of about 7.1, and Bansal, Kiku, and Yaron (2016) get an estimate of $\gamma$ of about 9.7 and an estimate of $\psi$ of about 2.2. Using a particle MCMC method, Schorfheide, Song, and Yaron (2018) estimate a linearized long-run risk model that features multiple stochastic volatility processes and find an estimate of $\gamma$ of 8.89 and an estimate of $\psi$ of 1.30.

Second, imposing informative priors on the preference parameters again improves model performance, but the effect is not as strong as that in the nonlinear models. The log-likelihood estimate in the LRR-AR-LL model is about 3158.2 in Case I, whereas it becomes 3160.1 in Case II; and the log-likelihood estimate in the LRR-ARG-LL model is about 3198.4 in Case I, whereas it becomes 3204.1 in Case II. However, economic improvement is not that clear. Fixing $\gamma$ and $\psi$ as in Case III worsens model performance. Again as in the nonlinear models, the marginal likelihood estimates show that the LRR-ARG-LL model performs better than the LRR-AR-LL model in all three cases.

Third, it seems that log-linearization has larger impact on the estimation of the AR-based long-run risk model than on the estimation of the ARG-based model. Figure 5 presents the posterior distributions of the preference parameters and the persistence.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case I AR</th>
<th>Case I ARG</th>
<th>Case II AR</th>
<th>Case II ARG</th>
<th>Case III AR</th>
<th>Case III ARG</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma )</td>
<td>8.7569 (2.8817)</td>
<td>8.5980 (1.6521)</td>
<td>7.5248 (1.3313)</td>
<td>7.5043 (1.1042)</td>
<td>10.000 (——)</td>
<td>10.000 (——)</td>
</tr>
<tr>
<td>( \psi )</td>
<td>1.1771 (0.0459)</td>
<td>1.2133 (0.0412)</td>
<td>1.2098 (0.0609)</td>
<td>1.2422 (0.0518)</td>
<td>1.5000 (——)</td>
<td>1.5000 (——)</td>
</tr>
<tr>
<td>( \mu )</td>
<td>0.0047 (0.0000)</td>
<td>0.0047 (0.0000)</td>
<td>0.0047 (0.0000)</td>
<td>0.0047 (0.0000)</td>
<td>0.0047 (0.0000)</td>
<td>0.0047 (0.0000)</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.9897 (0.0051)</td>
<td>0.9902 (0.0031)</td>
<td>0.9907 (0.0037)</td>
<td>0.9911 (0.0028)</td>
<td>0.9772 (——)</td>
<td>0.9786 (——)</td>
</tr>
<tr>
<td>( \phi_x )</td>
<td>0.3663 (0.0500)</td>
<td>0.2561 (0.0359)</td>
<td>0.3498 (0.0447)</td>
<td>0.3256 (0.0366)</td>
<td>0.3512 (——)</td>
<td>0.2831 (——)</td>
</tr>
<tr>
<td>( \varphi )</td>
<td>0.0046 (0.0002)</td>
<td>0.0047 (0.0001)</td>
<td>0.0046 (0.0002)</td>
<td>0.0048 (0.0001)</td>
<td>0.0046 (——)</td>
<td>0.0048 (——)</td>
</tr>
<tr>
<td>( \nu )</td>
<td>0.8450 (0.0405)</td>
<td>0.6818 (0.0448)</td>
<td>0.8553 (0.0369)</td>
<td>0.7118 (0.0412)</td>
<td>0.8695 (——)</td>
<td>0.7449 (——)</td>
</tr>
<tr>
<td>( \phi_s )</td>
<td>5.10E-06 (8.75E-07)</td>
<td>1.7959 (0.2328)</td>
<td>5.01E-06 (0.0437)</td>
<td>1.7564 (0.0328)</td>
<td>4.70E-06 (——)</td>
<td>1.6574 (——)</td>
</tr>
<tr>
<td>( \mu_d )</td>
<td>0.0063 (0.0000)</td>
<td>0.0063 (0.0000)</td>
<td>0.0063 (0.0000)</td>
<td>0.0063 (0.0000)</td>
<td>0.0063 (——)</td>
<td>0.0063 (——)</td>
</tr>
<tr>
<td>( \Phi )</td>
<td>1.0210 (0.0631)</td>
<td>0.9855 (0.0463)</td>
<td>1.0180 (0.0701)</td>
<td>0.9682 (0.0459)</td>
<td>1.0546 (——)</td>
<td>0.9696 (——)</td>
</tr>
<tr>
<td>( \phi_{dc} )</td>
<td>1.2595 (0.3252)</td>
<td>1.1301 (0.3117)</td>
<td>1.2750 (0.3311)</td>
<td>1.1758 (0.2901)</td>
<td>1.1884 (——)</td>
<td>1.1278 (——)</td>
</tr>
<tr>
<td>( \phi_d )</td>
<td>5.5960 (0.3215)</td>
<td>5.4722 (0.3090)</td>
<td>5.5810 (0.3290)</td>
<td>5.4985 (0.2970)</td>
<td>5.6844 (——)</td>
<td>5.6654 (——)</td>
</tr>
<tr>
<td>( \sigma_m )</td>
<td>0.0765 (0.0029)</td>
<td>0.0769 (0.0029)</td>
<td>0.0759 (0.0033)</td>
<td>0.0772 (0.0032)</td>
<td>0.0751 (——)</td>
<td>0.0767 (——)</td>
</tr>
<tr>
<td>( \sigma_f )</td>
<td>0.0053 (0.0003)</td>
<td>0.0039 (0.0003)</td>
<td>0.0053 (0.0003)</td>
<td>0.0040 (0.0003)</td>
<td>0.0058 (——)</td>
<td>0.0044 (——)</td>
</tr>
<tr>
<td>LMLH</td>
<td>3124.5</td>
<td>3170.9</td>
<td>3125.5</td>
<td>3170.9</td>
<td>3123.9</td>
<td>3169.9</td>
</tr>
<tr>
<td>LLH</td>
<td>3158.2</td>
<td>3198.4</td>
<td>3160.1</td>
<td>3204.1</td>
<td>3152.2</td>
<td>3190.3</td>
</tr>
</tbody>
</table>

**Note:** This table presents the posterior means and standard deviations of model parameters for LRR-AR-LL and LRR-ARG-LL resulting from our SMC-based Bayesian method that is initialized using the priors given in Table 1 for the three cases considered in the paper. Case I refers to the noninformative priors on \( \gamma \) and \( \psi \); Case II refers to the informative priors on on \( \gamma \) and \( \psi \); and Case III refers to fixing the values of \( \gamma \) and \( \psi \). LMLH represents the log-marginal likelihood and LLH the log-likelihood on the posterior means. The data used for model estimation include the consumption growth rates (\( \Delta c_t \)), the dividend growth rates (\( \Delta d_t \)), the market returns (\( r_{m,t} \)), and the risk-free returns (\( r_{f,t} \)). They are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters.
between the LRR-AR and LRR-AR-LL models are more apparent than those between the LRR-ARG and LRR-ARG-LL models. The main reason may be because the variance process is much less persistent in the LRR-ARG model than in the LRR-AR model. Pohl, Schmedders, and Wilms (2018) argue that log-linearization of the long-run risk model could generate large numerical errors when the state variables are persistent.

Fourth, comparing linear and nonlinear models, we find that log-linearization of the LRR-AR model makes model performance worse both economically and statistically. However, we find that the log-linearized LRR-ARG model performs slightly better than the nonlinear LRR-ARG model.

In summary, the LRR-AR model with the informative priors is the best model based on the economic measure, whereas the LRR-ARG-LL model with the informative priors is the best model based on the statistical criterion. Therefore, in what follows, we focus on these two specifications.

5.3. Asset Pricing Implications

5.3.1. Filtered Long-Run Component and Stochastic Volatility

Figure 6 presents the posterior means and (5, 95)% credible intervals of the filtered long-run risk component \(x_t\) (upper panels) and stochastic volatility \(\sigma_t\) (lower panels) that take into account parameter uncertainty. The left panels are for the LRR-AR model, and the right panels for the LRR-ARG-LL model. We first examine the estimates for the LRR-AR model. We notice that the long-run risk component has a much smaller variation than consumption growth rates: its (5, 95)% credible intervals are mainly within the bounds of -1% to 1%, whereas consumption growth rates vary between -1.3% to 2%. The long-run risk component falls down mainly in recession periods; however, its decreases also occasionally happen in boom periods.

The filtered volatility slowly declines after 1952, consistent with the empirical observation that consumption volatility becomes smaller in 1950’s and 1960’s. It increases mildly in 1970’s and has a dramatic drop-down in 1981-1983, the period related to the

---

\(^5\)Whether nonlinear models or log-linearized models better fit the data is completely an empirical issue, given that both types of models are approximations of the true data generating process.
Figure 5: **Posteriors from Linear and Nonlinear Models**

*Note:* The figure plots the posterior distributions of the preference parameter, $\gamma$ and $\psi$, and the persistence parameters, $\rho$ and $\nu$, for the three cases considered in the paper. The left panels are for LRR-AR (solid line) and LRR-AR-LL (dashed line), and the right panels for LRR-ARG (solid line) and LRR-ARG-LL (dashed line).

Great Moderation, when the 5% quantiles reach the lower bound of volatility. We also observe that consumption volatility reaches low levels again in 1996-1999; since then, it slowly goes up and reaches a high level in around 2005, and after a short period of decrease, it reaches another high level around 2012-2013.

The filtered latent states suggest that asset returns play an important role in identifying the key state dynamics in the long-run risk model, namely both the long-run risk and stochastic volatility components. For instance, the early 1980s witnessed an increase in the risk-free rate. This stylized fact helps identify the rise in the long-run risk component.
Figure 6: The Filtered Latent States

Note: This figure plots the filtered long-run consumption component ($x_t \times 100$) and the filtered consumption volatility ($\sigma_t \times 100$) that take into account parameter uncertainty from LRR-AR and LRR-ARG-LL. The mean and (5, 95)% quantiles are reported at each time point. The data used for model estimation include the consumption growth rates ($\Delta c_t$), the dividend growth rates ($\Delta d_t$), the market returns ($r_{m,t}$), and the risk-free returns ($r_{f,t}$). They are in quarterly frequency ranging from 1947:Q1 to 2018:Q1, with a total of 285 quarters. Shaded bars represent the NBER recession periods.

and the drop in the stochastic volatility component for the same period, where the former is due to the intertemporal substitution effect and the latter is due to the precautionary savings effect.

The main features of the filtered long-run risk and stochastic volatility components in LRR-ARG-LL are similar to those in LRR-AR. We also notice some differences, which mostly originate from the difference in the volatility specifications. The ARG process never reaches zero. Therefore, the volatility drop in the Great Moderation period in LRR-ARG-LL is not as dramatic as in LRR-AR. Furthermore, the volatility process in LRR-ARG-LL is clearly less persistent than that in LRR-AR, as the former features more frequent small variations than the latter.

Our SMC-based Bayesian method allows for structural estimation of equilibrium asset pricing models and thus provides a better way to quantify the impact of the latent states on asset prices than the calibration approach does. According to the long-run risk model,
a rise in stochastic volatility or a decline in the conditional expectation of consumption growth leads to a lower price-dividend ratio and therefore higher future returns, when the level of risk aversion is reasonable and the EIS is greater than one. In addition, a rise in expected consumption growth results in a higher risk-free return because of the intertemporal substitution effect, and a rise in consumption volatility also raises the risk-free return due to the precautionary savings effect. The filtered latent states shown in Figure 6 not only reproduce the dynamics of fundamentals but also adequately capture significant variations of market and risk-free returns in historical episodes.

5.3.2. Market Returns and Risk-Free Returns

Figure 7 plots the posterior means and (5, 95)% credible intervals of the fitted market returns (upper panels) and risk-free returns (lower panels) in LRR-AR (left panels) and LRR-ARG-LL (right panels), overlaying with the observed returns. We see that the measurement errors are larger in fitting market returns than they are in fitting risk-free returns in both models. In LRR-AR, we see that the posterior means of the fitted market returns can track movements of the observed ones; but they are unable to match substantial changes in market returns. However, we do notice that nearly all the observed market returns are within the (5, 95)% credible intervals of the fitted values.

In contrast, the posterior means of the fitted risk-free returns can very closely track movements of the observed ones, except in the beginning of 1980s when risk-free rates reach historically high levels, and the (5, 95)% credible intervals are quite narrow and contain almost all observed risk-free returns, suggesting that the long-run risk model can capture well the dynamics of risk-free returns.

When we move to LRR-ARG-LL, we observe that the market returns cannot be fitted well, as many of the observed returns are outside of their corresponding (5, 95)% credible intervals. The fitting of the risk-free returns is better, but incomparable to that in LRR-AR. This is especially apparent when the risk-free returns reach historically high levels. Such results are consistent with our estimates of the measurement error standard deviations, $\sigma_m$ and $\sigma_f$, of both models in Table 5 and Table 6.
5.3.3. Asset Return Moments

We now investigate the ability of the estimated model to match the unconditional moments of asset returns and consumption and dividend growth rates. In contrast to calibrations or moment-based estimation methods that try to select parameter values or obtain parameter estimates to match moments as close as possible, our SMC-based Bayesian method is likelihood-based and delivers both parameter and state particles that take into account uncertainties related to parameter and state estimates.

Table 7 reports the model-implied unconditional means and standard deviations of the market return, the risk-free return, the equity risk premium, the consumption growth rate, and the dividend growth rate corresponding to LRR-AR and LRR-ARG-LL. First, we find that for the market returns, the data-implied unconditional mean and standard deviation are within the corresponding model-implied (5, 95)% credible intervals in LRR-AR, whereas in LRR-ARG-LL, both data-implied unconditional mean and standard deviation
Table 7: Asset Return Moments

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(r_m)</td>
<td>6.975</td>
<td>4.776</td>
<td>6.002</td>
<td>7.716</td>
<td>3.064</td>
<td>4.280</td>
<td>5.685</td>
</tr>
<tr>
<td>Std(r_m)</td>
<td>16.26</td>
<td>11.81</td>
<td>14.10</td>
<td>17.09</td>
<td>7.085</td>
<td>8.461</td>
<td>10.61</td>
</tr>
<tr>
<td>E(r_f)</td>
<td>0.819</td>
<td>0.600</td>
<td>0.737</td>
<td>0.868</td>
<td>0.575</td>
<td>0.720</td>
<td>0.862</td>
</tr>
<tr>
<td>Std(r_f)</td>
<td>1.548</td>
<td>1.303</td>
<td>1.379</td>
<td>1.450</td>
<td>1.120</td>
<td>1.204</td>
<td>1.280</td>
</tr>
<tr>
<td>E(r_{m,t} - r_{f,t})</td>
<td>6.157</td>
<td>4.015</td>
<td>5.265</td>
<td>7.034</td>
<td>2.327</td>
<td>3.562</td>
<td>5.015</td>
</tr>
<tr>
<td>Std(r_{m,t} - r_{f})</td>
<td>16.23</td>
<td>11.87</td>
<td>14.15</td>
<td>17.19</td>
<td>7.119</td>
<td>8.424</td>
<td>10.48</td>
</tr>
<tr>
<td>E(\Delta c)</td>
<td>1.877</td>
<td>1.786</td>
<td>1.929</td>
<td>2.078</td>
<td>1.759</td>
<td>1.909</td>
<td>2.060</td>
</tr>
<tr>
<td>Std(\Delta c)</td>
<td>0.996</td>
<td>0.653</td>
<td>0.732</td>
<td>0.812</td>
<td>0.634</td>
<td>0.723</td>
<td>0.816</td>
</tr>
<tr>
<td>E(\Delta d)</td>
<td>2.539</td>
<td>2.492</td>
<td>2.549</td>
<td>2.635</td>
<td>2.469</td>
<td>2.537</td>
<td>2.578</td>
</tr>
<tr>
<td>Std(\Delta d)</td>
<td>5.173</td>
<td>0.892</td>
<td>1.067</td>
<td>1.427</td>
<td>0.782</td>
<td>1.123</td>
<td>1.583</td>
</tr>
</tbody>
</table>

Note: This table reports (5, 50, 95)% quantiles of the model-implied unconditional means and standard deviations of market return and risk-free returns, consumption growth rates, and dividend growth rates from LRR-AR and LRR-ARG. Our estimation strategy delivers both parameter particles and state particles, based on which we can compute the posterior distributions of asset returns that take into account both parameter and state uncertainties. The model-implied moments can then be obtained from these posterior distributions.

are outside the corresponding model-implied (5, 95)% credible interval. The inability of the LRR-ARG-LL model in matching equity premium and volatility of returns is mainly because of the low estimates of the persistence parameter in the variance process, which dampens the impact of time-varying uncertainty on asset prices. We also observe that for the risk-free returns in both models, the data-implied unconditional mean are within the corresponding model-implied (5, 95)% credible intervals, whereas the data-implied standard deviations are not in the corresponding model-implied (5, 95)% credible intervals, but very close to the 95% quantiles. These results are in fact consistent with the measurement error standard deviation estimates ($\sigma_m$ and $\sigma_f$) in Table 5 and Table 6, and Figure 7.

Second, we find that both models can capture the unconditional means of the consumption and dividend growth rates, as the data-implied unconditional mean are within the corresponding model-implied (5, 95)% credible intervals. Furthermore, even though both models can fairly match the unconditional standard deviation of the consumption growth rate, as the data-implied unconditional standard deviation is very close to the
corresponding model-implied 95% quantile, they can not capture the unconditional standard deviation of the dividend growth rate, as the data-implied unconditional standard deviation is much larger than the corresponding model-implied 95% quantile.

6. Conclusion

Since the seminal paper by Bansal and Yaron (2004), the long-run risk model has become a benchmark in consumption-based asset pricing models. However, even though it has attracted a great deal of attention, most works simply follow the calibration approach and thus the long-run risk models have not been fully tested against real economic data. A number of recent studies have tried to implement econometric estimation of the long-run risk models using real aggregate economic and asset prices data (Bansal, Gallant, and Tauchen, 2007; Bansal, Kiku, and Yaron, 2016; Gallant, Jahan-Parvar, and Liu, 2019; Shorfheide, Song, and Yaron, 2018). However, these works either use moment-based econometric methods or indirect inference, which do not fully exploit information from the likelihood function of the original asset pricing models, or heavily rely on log-linearization to solve for asset prices, or a combination of both approximations.

In this paper, we examine two types of long-run risk models: the first model assumes an autoregressive process for the conditional variance of consumption growth, and the second model uses a positive autoregressive gamma process to model the conditional variance of consumption growth. We solve the models using both the log-linearization method and the collocation projection method. We propose a likelihood-based Bayesian method that exploits up-to-date sequential Monte Carlo methods to efficiently estimate the nonlinear long-run risk models.

Using the U.S. data on consumption, dividends, market return and risk-free rates, we find that the informative priors on the preference parameters are helpful to improve the model performance. We also find that the expected consumption growth has a very persistent component, whereas the consumption volatility is less persistent than that assumed in calibration studies. The estimation of the log-linearized models may deliver different results, suggesting that the solution method matters for structural estimation of
asset pricing models. In addition, we find the impact of log-linearization is larger on the estimation of the AR-based model than that of the ARG-based model, as the estimated ARG-based model has a much less persistent volatility process.

APPENDIX

A. Log-linearization of the ARG-based Long-Run Risk Model

In this appendix, we characterize the log-linear solution to the long-run risk model with an ARG process for the conditional variance. In general, the log-linear solution to the long-run risk model relies on the approximations to the log-return on the wealth portfolio

\[ r_{t+1}^W = wc_{t+1} + \Delta c_{t+1} - \kappa_0 - \kappa_1 wc_t, \]  

(A.1)

where

\[ wc_t = \ln \left( \frac{W_t}{C_t} \right), \quad \kappa_1 = \frac{e^{wc}}{e^{wc} - 1}, \quad \kappa_0 = \ln \left( e^{wc} - 1 \right) - \kappa_1 wc, \]

and the log-return on the market portfolio

\[ r_{m,t+1} = \kappa_{m0} + \kappa_{m1} pd_{t+1} - pd_t + \Delta d_{t+1}, \]  

(A.2)

where

\[ pd_t = \ln \left( \frac{P_t}{D_t} \right), \quad \kappa_{m1} = \frac{e^{pd}}{e^{pd} + 1}, \quad \kappa_{m0} = \ln \left( e^{pd} + 1 \right) - \kappa_{m1} pd. \]

The Euler equation is

\[ E_t \left[ \exp \left( \theta \ln \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{t+1}^W + r_{i,t+1} \right) \right] = 1. \]
The dynamics of consumption growth and dividend growth is given by

\[\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{c,t+1},\]
\[x_{t+1} = \rho x_t + \phi_x \sigma_t \eta_{x,t+1},\]
\[\Delta d_{t+1} = \mu_d + \Phi x_t + \phi_d \sigma_t \eta_{d,t+1} + \phi_d \sigma_t \eta_{d,t+1},\]

where the conditional variance \(\sigma_t^2\) follows an ARG process with the scale parameter \(c\), persistence parameter \(\nu\) and degree of freedom \(\phi_s\). As usual, we solve for the wealth-consumption ratio first and determine the stochastic discount factor. We then solve for the price-dividend ratio.

We conjecture that the solution to the log-wealth-consumption ratio is \(wc_t = A_0 + A_1 x_t + A_2 \sigma_t^2\), where \(A_0, A_1,\) and \(A_2\) are coefficients to be determined. We substitute the solution into the Euler equation and obtain

\[1 = E_t \left[ \exp \left\{ \theta \ln \delta + \theta (1 - 1/\psi) (\mu + x_t + \sigma_t \eta_{c,t+1}) - \theta \kappa_0 
+ \theta (A_0 + A_1 x_{t+1} + A_2 \sigma_t^2) - \theta \kappa_1 (A_0 + A_1 x_t + A_2 \sigma_t^2) \right\} \right]
= E_t \left[ \exp \left\{ \theta \ln \delta + \theta (1 - 1/\psi) (\mu + x_t + \sigma_t \eta_{c,t+1}) - \theta \kappa_0 + \theta A_0 
+ \theta A_1 (\rho x_t + \phi_x \sigma_t \eta_{x,t+1}) - \theta \kappa_1 (A_0 + A_1 x_t + A_2 \sigma_t^2) \right\} \right] E_t \left[ \exp \left\{ \theta A_2 \sigma_t^2 \right\} \right], \tag{A.3}\]

where the second equality follows from the condition independence of \(\sigma_t^2\) and other innovation shocks. According to Lemma 1 in the Appendix of Gourieroux and Jasiak (2006), the last conditional expectation term is given by

\[E_t \left[ \exp \left\{ \theta A_2 \sigma_t^2 \right\} \right] = (1 - c \theta A_2)^{-\phi_s} \exp \left( \frac{\nu \theta A_2}{1 - c \theta A_2} \sigma_t^2 \right), \tag{A.4}\]

where \(c = \bar{\sigma}^2 (1 - \nu) / \phi_s\).

Collecting and matching coefficients yields

\[A_1 = \frac{1 - 1/\psi}{\kappa_1 - \rho}, \tag{A.5}\]
\[ \theta (1 - \kappa_1) A_0 = \phi_s \ln (1 - \theta A_2) + \theta \kappa_0 - \theta \ln \delta - \theta (1 - \psi) \mu, \] (A.6)

\[ \kappa_1 \theta A_2^2 + \left[ -\frac{1}{2} \theta \theta^2 (1 - \psi)^2 - \frac{1}{2} \psi (\theta A_1 \phi_x)^2 - \kappa_1 + \nu \right] A_2 
+ \frac{1}{2} \theta (1 - \psi)^2 + \frac{1}{2} \theta (A_1 \phi_x)^2 = 0. \] (A.7)

The equation for \( A_2 \) is quadratic and has two real roots if its discriminant \( \text{Disc} = (B_\sigma^2 - 4A_\sigma C_\sigma) > 0 \), where

\[
\begin{align*}
A_\sigma &= \kappa_1 \theta, \\
B_\sigma &= -\frac{1}{2} \theta \theta^2 (1 - \psi)^2 - \frac{1}{2} \psi (\theta A_1 \phi_x)^2 - \kappa_1 + \nu, \\
C_\sigma &= \frac{1}{2} \theta (1 - \psi)^2 + \frac{1}{2} \theta (A_1 \phi_x)^2.
\end{align*}
\]

We choose the root that satisfies the requirement of stochastic stability (Hansen, 2012)

\[ A_2 = -B_\sigma + \text{sign}(B_\sigma) \sqrt{B_\sigma^2 - 4A_\sigma C_\sigma} \] (A.8)

Because the derived solutions to \( A_0, A_1 \) and \( A_2 \) depend on the approximating constants \( \kappa_0 \) and \( \kappa_1 \), which in turn depend on the long-run mean of the wealth-consumption ratio, we use a recursive method to numerically solve a fixed-point problem

\[
\bar{w} = A_0 (\bar{w}) + A_2 (\bar{w}) \sigma^2
\] (A.9)

where the dependence of the solution coefficients \( A_0, A_1 \) and \( A_2 \) on \( \bar{w} \) is shown above.

We conjecture that the solution to the log-price-dividend ratio is \( pd_t = A_{m0} + A_{m1} x_t + A_{m2} \sigma_t^2 \). By substituting the conjectured solution into the Euler equation, collecting and matching coefficients, we obtain

\[ A_{m1} = \frac{\Phi - 1/\psi}{1 - \kappa_{m1} \rho}, \] (A.10)
\[ 0 = \theta \ln \delta - (\theta - 1) \kappa_0 + (\theta - 1 - \theta/\psi) \mu + (\theta - 1) A_0 - (\theta - 1) \kappa_1 A_0 + \kappa_{m0} + \kappa_{m1} A_m - A_m + \mu_d - \phi_s \ln \left(1 - c \left(\theta - 1 \right) A_2 + \kappa_{m1} A_{m2}\right), \quad (A.11) \]

\[ 0 = \frac{1}{2} \left(\theta - 1 - \theta/\psi + \phi_{dc}\right)^2 + \frac{1}{2} \left((\theta - 1) A_1 + \kappa_{m1} A_m\right)^2 \phi_x^2 + \frac{1}{2} \phi_d^2 - (\theta - 1) \kappa_1 A_2 - A_m + \frac{\nu \left[(\theta - 1) A_2 + \kappa_{m1} A_{m2}\right]}{1 - c \left[\left(\theta - 1\right) A_2 + \kappa_{m1} A_{m2}\right]}. \quad (A.12) \]

The equation for \( A_{m2} \) is quadratic and can be written as

\[ A^m_{m2} A^2_{m2} + B^m_{m2} A_{m2} + C^m_{m2} = 0. \quad (A.13) \]

If its discriminant \( Disc = (B^m_{m2})^2 - 4A^m_{m2}C^m_{m2} > 0 \), we choose the root that maintains stochastic stability:

\[ A_{m2} = \frac{-B^m_{m2} + \text{sign} (B^m_{m2}) \sqrt{(B^m_{m2})^2 - 4A^m_{m2}C^m_{m2}}}{2A^m_{m2}}. \quad (A.14) \]

Again, because the derived solutions to \( A_{m0}, A_{m1} \) and \( A_{m2} \) depend on the approximating constants \( \kappa_{m0} \) and \( \kappa_{m1} \), which in turn depend on the long-run mean of the price-dividend ratio, we use a recursive method to numerically solve a fixed-point problem

\[ \dot{pd} = A_0 (pd) + A_2 (pd) \hat{\sigma}^2 \quad (A.15) \]

where the dependence of the solution coefficients \( A_{m0}, A_{m1} \) and \( A_{m2} \) on \( \dot{pd} \) is shown above.

The log risk-free rate is given by

\[ r_{f,t} = -\ln \left(\mathbb{E}_t [M_{t+1}]\right) \]

and

\[ \mathbb{E}_t [M_{t+1}] = \mathbb{E}_t \left[\exp \left(\theta \ln \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{t+1}^W\right)\right]. \]
We can substitute the log-linear approximation to $r_t^{W}$ into the Euler equation and obtain the solution to $r_{f,t}$.

REFERENCES


