Annals of Financial Economics
Vol. 13, No. 4 (December 2018) 1850019 (15 pages)

DOI: 10.1142/S2010495218500197

TESTING THE LONG-RUN RISK MODEL: A KALMAN FILTER APPROACH

JIANQIU WANG[*,‡] and KE WU[†,§]

[*]*PBC School of Finance, Tsinghua University*
Beijing 100083, P. R. China

[†]*Hanqing Advanced Institute, Renmin University of China*
Beijing 100872, P. R. China
[‡]*wangjq@pbcsf.tsinghua.edu.cn*
[§]*ke.wu@ruc.edu.cn*

Published 13 December 2018

This paper reevaluates the Long-Run Risk model proposed by Bansal and Yaron (2004) using the Kalman filter and Maximum Likelihood estimation method. Our findings show that the persistence of the small long-run predictable component in the consumption growth process is the key for the model performance. In our estimation exercises, if we relax the persistence restriction on the long-run risk parameter and adopt a Maximum Likelihood estimate, the Long-Run Risk model still requires a relative risk aversion at around 70 to fit the US data. However, we do not find strong empirical support for the persistence restriction from the data.

Keywords: Equity premium puzzle; long-run risk model; Kalman filter.

1. Introduction

During recent decades, various consumption-based asset pricing models have become work horses in the macro-finance literature.[1] Mehra and Prescott (1985) documented a famous equity premium puzzle, i.e., it is difficult for the consumption-based asset pricing models with power utility function to justify a 6% annual equity premium and the low risk-free rate. As documented in their paper, the model requires an unreasonably high relative risk aversion (**RRA**) (>50) to generate the 6% equity premium in the U.S equity market. Second, with high relative risk aversion the model creates an unreasonably high risk-free rate, which

‡ Corresponding author.
[1] Please see Cochrane (2017) for a most recent and comprehensive survey on consumption-based macro-finance literature.

is the risk-free rate puzzle emphasized by Weil (1989). Third, for the conditional expectation and variance of the model-generated Stochastic Discount Factor (**SDF**) to be consistent with the data, i.e., the mean and variance combination has to be within the Hansen-Jagannathan (Hansen and Jagannathan, 1991) Bound, it requires an even higher relative risk aversion (> 250).

Many papers after Mehra and Prescott (1985) tried to solve the equity premium puzzle from different perspectives. Some notable papers includes Epstein and Zin (1989), Campbell and Cochrane (1999), Bansal and Yaron (2004), Barro (2006), etc. Epstein and Zin (1989) proposed a more flexible utility preference that preserves some desirable features of the power utility, such as the scale-invariance, but allows for a separation between the elasticity of inter-temporal substitution (EIS) and relative risk aversion. With Epstein–Zin (Epstein and Zin, 1989, 1991) preferences, consumption-based asset pricing models still require very high risk aversion to generate 6% equity premium but it does not create an unreasonably high risk-free rate, since EIS and risk aversion are not linked in Epstein–Zin preferences. Thus, Epstein–Zin preferences can solve the risk-free rate puzzle, i.e., generate low risk-free rate with reasonably low relative risk aversion, but it cannot completely solve the equity premium puzzle.

Bansal and Yaron (2004) proposed a long-run risk (LRR) model in order to solve the equity premium puzzle. There are two key ingredients for LRR model to match the data. First is the adoption of Epstein–Zin recursive preference, which breaks the link between the elasticity of inter-temporal substitution and relative risk aversion. The other ingredient is the introduction of a predictable, tiny but highly persistent unobserved component into the consumption growth process (the "Long-Run risk"). With reasonable model parameters, i.e., elasticity of inter-temporal substitution equal to 1.5 and relative risk aversion set at 10, the LRR model is able to generate patterns found in the US data. However, the model has two shortcomings. First, it imposes strong assumptions about the dynamics of the unobserved component embedded in the consumption growth process. It is assumed that the predictable component follows an AR(1) process that is highly persistent (with AR(1) coefficient $\rho = 0.98$) but with very small magnitude (very tiny conditional standard deviation σ_x). Second, calibration loses information contained in the data as the calibrated parameters only match to certain data moments. In this paper, we try to re-evaluate the empirical performance of the long-run risk model when all the parameters are econometrically estimated instead of being calibrated. Since the long-run risk component is not unobservable from the consumption data, a natural solution is to form a state-space representation and estimate the consumption growth process via the Kalman filtering techniques. The state-space form also facilitates the calculation of the likelihood of the observed data given certain model parameters and makes the maximum likelihood methods easily applicable in parameter estimation. We also employ Metropolis-Hastings algorithm to construct a posterior distribution and confidence intervals for the ML estimators.

The Kalman filter has been widely used in economics and finance research, when some latent factors are involved in modelling. For example, Babbs and Nowman (1999) use the Kalman filter to deal with unobservable state variables in general linear Gaussian model of the bond term structure. Schwartz and Smith (2000) apply the Kalman filter to estimate a latent two-factor model of commodity prices that better accommodates both the short-term variation and long-term dynamics of price movements. Please see Harvey (1990) for a comprehensive review of Kalman filtering and its applications in econometrics.

Our results indicate that a highly persistent unobserved component ($\rho > 0.98$) in the consumption growth process does not have a strong support from the data. Imposing some restriction on the parameter values is necessary in order for the LRR model to generate patterns consistent with the data. If we relax the restrictions and use the estimated parameter values, the LRR model still requires very high relative risk aversion to generate 6% annual equity premium. The rest of the paper is organized as follows. In Sec. 2, we develop the model settings and derive the analytical solution. We explain the data used in the paper and discuss model performance in Sec. 3. Main estimation results are in Sec. 4 and Sec. 5 concludes.

2. Theoretical Framework

Consistent with Bansal and Yaron (2004), we assume that a representative agent has Epstein-Zin preference.

$$U_t = \left((1-\beta)C_t^{1-\eta} + \beta\left(E_t\left(U_{t+1}^{1-\gamma}\right)\right)^{\frac{1-\eta}{1-\gamma}}\right)^{\frac{1}{1-\eta}}. \tag{1}$$

For the model to yield a closed-form analytical solution, we further assume the EIS $\eta = 1$. Then Eq. (1) will take the form

$$U_t = C_t^{1-\beta}\left(E_t U_{t+1}^{1-\gamma}\right)^{\frac{\beta}{1-\gamma}}. \tag{2}$$

Take logarithm on both sides of Eq. (2),

$$u_t = (1-\beta)c_t + \frac{\beta}{1-\gamma}\log\left(E_t(\exp\left[(1-\gamma)u_{t+1}\right])\right), \tag{3}$$

where lowercase letters refer to natural logs.

The log of consumption growth contains a tiny but persistent unobserved component, denoted as x_t. It follows a stationary $AR(1)$ process with $\rho < 1$.

$$c_{t+1} = c_t + \mu_c + x_t + \sigma_c\varepsilon_{t+1}^c, \tag{4}$$

$$x_{t+1} = \rho x_t + \sigma_x\varepsilon_{t+1}^x, \tag{5}$$

where μ_c is the unconditional mean of consumption growth, σ_c, σ_x stand for the variance of consumption growth and long-run risk term respectively. The innovation ε_t^x and ε_t^c both follow an *i.i.d.* standard normal distribution and are independent of each other. The analytical solution of u_t can be solved by using guess and verify method.[2] Note that our model setup is a simplified version of the model in Bansal and Yaron (2004). We also solved the model with a dividend growth process as in Bansal and Yaron (2004) model, but the dividend term does not show up in the analytical solution due to cancellation, so the results remain the same.

$$u_t = \frac{\beta}{1-\beta}\left(\mu_c + (1-\gamma)\left(\frac{\sigma_c^2}{2} + \frac{\beta^2}{2(1-\beta\rho)^2}\sigma_x^2\right)\right) + c_t + \frac{\beta}{1-\beta\rho}x_t. \tag{6}$$

One can also show that the SDF M_{t+1} will equal to

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-1}\left[\frac{\exp\left[(1-\gamma)u_{t+1}\right]}{E_t[\exp\left((1-\gamma)u_{t+1}\right)]}\right], \tag{7}$$

which can be written in terms of c_{t+1} and x_{t+1} as follows:

$$\begin{aligned} m_{t+1} = {} & \log\beta - (\mu_c + x_t) - \frac{1}{2}(1-\gamma)^2\left(\sigma_c^2 + \left(\frac{\beta\sigma_x}{1-\beta\rho}\right)^2\right) \\ & + (1-\gamma)\frac{\beta\sigma_x}{1-\beta\rho}\varepsilon_{t+1}^x - \gamma\sigma_c\varepsilon_{t+1}^c. \end{aligned} \tag{8}$$

Its unconditional mean and standard deviation will be[3]

$$E(M_{t+1}) = \beta\exp\left(-\mu_c + \left(\gamma - \frac{1}{2}\right)\sigma_c^2 + \frac{\sigma_x^2}{2(1-\rho^2)}\right), \tag{9}$$

$$\sigma_{M_{t+1}} = E(M_{t+1})\cdot\sqrt{\exp\left(\frac{\sigma_x^2}{1-\rho^2} + (\gamma\sigma_c)^2 + \left(\frac{(1-\gamma)\beta\sigma_x}{1-\beta\rho}\right)^2\right) - 1}. \tag{10}$$

3. Data and Model Performance

The model is estimated based on quarterly data. We collect the US per capita consumption on both non-durable goods and services, quarterly real returns on value weighted S&P 500 index and 3-month Treasury bills, etc. from the Federal Reserve Economic Data (FRED) database. The Risk free rate, dividends and value-weighted market return data of the same period are from CRSP. We also include

[2] See Appendix A for details.
[3] See Appendix B for details.

6 Fama-French portfolios as possible assets available in the market. All data series ranges from 1948**Q2** to 2011**Q3**.

Given all data, we construct the Hansen-Jagannathan bound for the SDF in the market (i) only has one market portfolio and risk free asset, (ii) has all assets in (i) plus the 6 Fama-French Portfolios (Fama and French, 1989). As given in Fig. 1, the hyperbolic curves show the relation between the mean and minimum variance of SDF. SDFs are consistent with the observed data only when their unconditional means and variances fall in the hyperbolic area.

Case I: CRRA. If we use the simple CRRA power utility function, the SDF in this case takes the form of

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}. \tag{11}$$

By assuming the $\log C_t$ follow drifted random walk

$$c_{t+1} = c_t + \mu_c + \sigma_c \varepsilon^c_{t+1}, \quad \varepsilon^c_{t+1} \sim i.i.d.\ \text{Normal}\,(0, 1), \tag{12}$$

it can be shown

$$\begin{aligned} E(M_{t+1}) &= \beta \exp\left[-\gamma\mu_c + \frac{1}{2}(\gamma\sigma_c)^2 \right], \\ \text{Var}(M_{t+1}) &= E(M_{t+1}) \cdot \sqrt{\exp(\gamma^2\sigma_c^2) - 1}. \end{aligned} \tag{13}$$

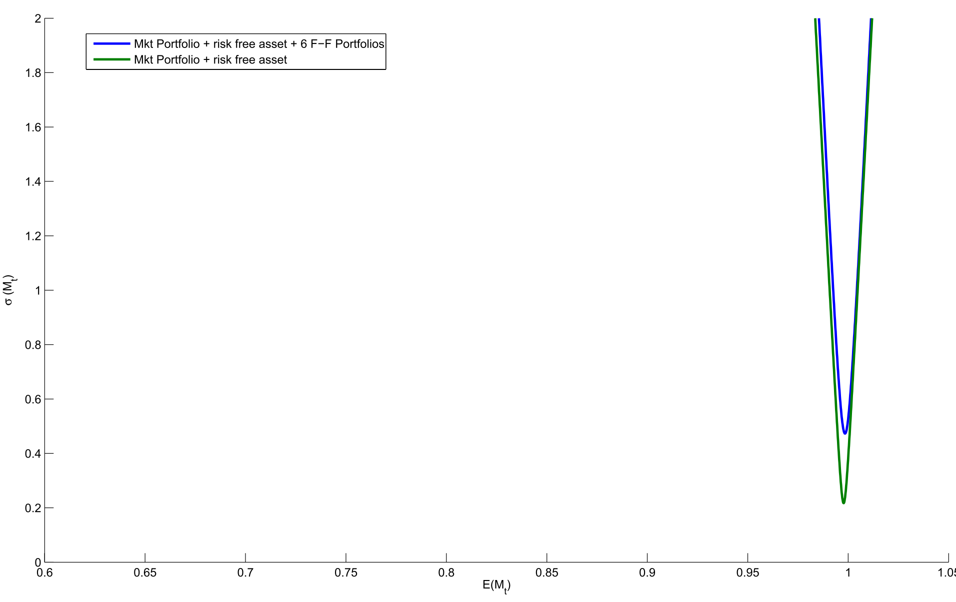

Figure 1. The HJ bound for the SDF in two markets.

Table 1. Parameter setting.

Variable	Model Panel (a) CRRA and Epstein-Zin	 Panel (b) Long-run risk
β	0.995	0.995
μ_c	0.0042	0.0042
σ_c	0.0062	0.0062
ρ	—	0.98
σ_x	—	0.00062

Note: This table provides the calibrated parameter setting for all three models. The discount factor β are set to be 0.995 for all three models. For CRRA utility function and Epstein–Zin preference, we estimate μ_c and σ_c from Eq. (12) and present the results in Panel (a). Panel (b) shows the parameter setting for the Long-run risk model. We follow the setting in Bansal and Yaron (2004) for parameter ρ and σ_x.

We estimate μ_c and σ_c^2 from the consumption data, as shown in Table 1. Using Eq. (13), we can plot the mean and variance of M_{t+1} for different γ's in Fig. 2(a). It is clear that the CRRA utility performs very badly. In order to have a mean-variance pair fallen into the Hansen–Jagannathan Bound, one almost has to use a γ be as large as 220. Although statistically a small γ may still be acceptable (a formal

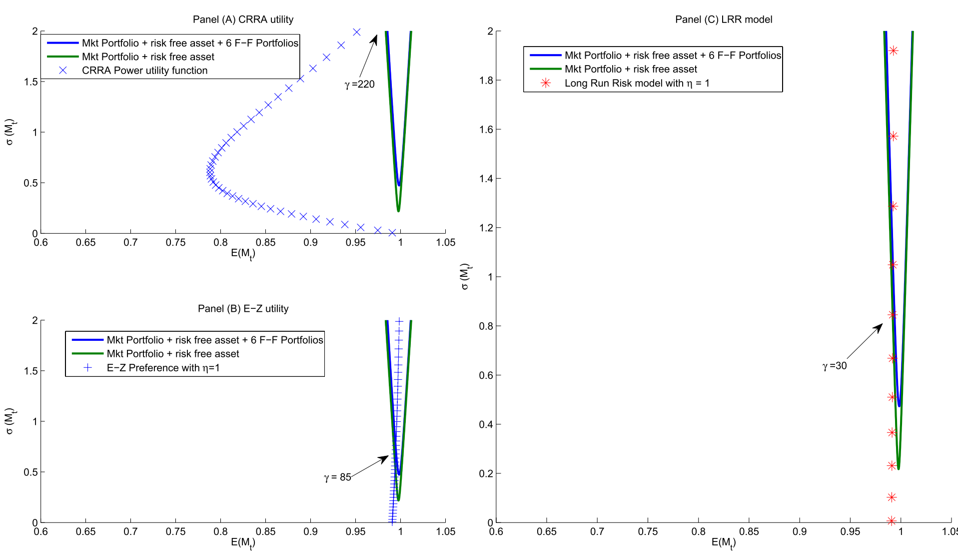

Figure 2. Performances of different models.

testing approach can be accomplished by using the Hansen and Jagannathan distance as in (Hansen and Jagannathan, 1997)), according to other literature however, we will still have to have a large RRA as pointed in Mehra and Prescott (1985). The reason behind this result is the parameter γ determines both EIS and RRA of the CRRA utility function. Due to the low volatility in consumption growth, we have to have a large γ to generate the average 6% equity premium.

Case II: Epstein–Zin. In Epstein–Zin preference setting with $\eta = 1$, the SDF is given by[4]

$$
\begin{aligned}
E(M_{t+1}) &= \beta \exp\left[-\mu_c + \sigma_c^2\left(\gamma - \frac{1}{2}\right)\right], \\
\sigma_{M_{t+1}} &= E(M_{t+1}) \cdot \sqrt{\exp(\gamma^2\sigma_c^2) - 1}.
\end{aligned}
\tag{14}
$$

Same as before, we use the estimates μ_c and σ_c^2 from Table 1 Panel (a), and plot the mean and variance of M_{t+1} for different γ's in Fig. 2(b). Since for Epstein–Zin preferences, EIS and RRA are determined by parameter η and γ separately, we only need a $\gamma \simeq 85$ for the model to match the data. Although this result shows a big progress, we still need an unreasonable large RRA γ to generate the equity premium.

Case III: Long-Run risk. The long-run risk, denoted by x_t in the model, comes from the long run economy uncertainty (conditional volatility of consumption). The parameter settings are shown in Table 1, Panel (b). The values of μ_c and σ_c^2 are obtained in the same way as we did in Panel (a). For ρ and σ_x^2, we use the calibrated value from Bansal and Yaron (2004). As Bansal and Yaron claim, by adding this small but persistent component x_t, the consumption volatility will increase dramatically in the long run (this uncertainty will contribute as much as 39% of variance of the pricing kernel according to the variance decomposition in Bansal and Yaron (2004)). In this model, we only need $\gamma \simeq 30$ to fit our model with data, as shown in Fig. 2(c).

4. Estimation of the Long-Run Risk Component

Although the x_t component seems to have a sound economic background (undetectable economy uncertainty in the long run) and this setting does improve the model performance, the assumption is still questionable. Bansal and Yaron (2004) assumed that x_t follows a very persistent AR(1) process and calibrated the value of ρ to be 0.98. Meanwhile, in order to make the short-run conditional consumption volatility small, they calibrated $\sigma_x \simeq \frac{\sigma_c}{10}$. In this section, we will

[4] See Appendix B for details.

estimate ρ and σ_x and plot both distributions to check whether the original assumptions are consistent with the data under normality assumptions. Noticing that the variable x_t is latent, we use the Kalman Filter to calculate the conditional likelihood of $x_{t|t}$.

4.1. *The state-space form*

Let the state vector $z_t' = (c_t\ x_t)'$. According to Hamilton (1994), the measurement equation is simply

$$c_t = \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{H} \cdot \underbrace{\begin{pmatrix} c_t \\ x_t \end{pmatrix}}_{z_t}. \tag{15}$$

The state equation will be

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{A_0} \underbrace{\begin{pmatrix} c_{t+1} \\ x_{t+1} \end{pmatrix}}_{z_{t+1}} = \begin{pmatrix} \mu_c \\ 0 \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & \rho \end{pmatrix}}_{A_1} \underbrace{\begin{pmatrix} c_t \\ x_t \end{pmatrix}}_{z_t} + \underbrace{\begin{pmatrix} \sigma_c \varepsilon_{t+1}^c \\ \sigma_x \varepsilon_{t+1}^x \end{pmatrix}}_{\nu_{t+1}}, \tag{16}$$

The innovation ν_{t+1} follows *i.i.d.* normal distribution with mean 0 and variance-covariance matrix Q given by

$$Q = E(\nu_t \nu_t') = \begin{pmatrix} \sigma_c^2 & 0 \\ 0 & \sigma_x^2 \end{pmatrix}, \quad \forall t. \tag{17}$$

4.2. *Main results*

We estimate all parameters μ_c, ρ, σ_c and σ_x from the state-space form (Eq. (16)) by Kalman Filter. To check the robustness of MLE, we consider three cases. In the first case, we did not specify any restriction beyond natural ones to the model. In the second case, we impose a mild restriction on the persistence feature of x_t by letting $\rho \in [0.9, 1)$. For the last case, we assume x_t follows an extremely persistent AR(1) process by letting the value of $\rho \in [0.98, 1)$ as in Bansal and Yaron (2004). For all three cases, we set the initial value of parameters equal to their baseline calibration, i.e.,

$$\begin{aligned} \mu_c &= 0.0042, \quad & \sigma_c &= 0.0062, \\ \rho &= 0.98, \quad & \sigma_x &= 0.00062. \end{aligned} \tag{18}$$

4.2.1. *Natural restriction*

In this case, we only impose natural restrictions on each parameter, i.e., we let standard deviation σ_c and σ_x be greater than 0, and require the AR correlation coefficient $\rho \in (0, 1)$ to ensure stationarity. As the first step, we use the the data to

Table 2. The MLE of LRR model under different settings.

	Natural restriction	Mild restriction	Extreme restriction
μ_c	0.004274	0.004239	0.003894
σ_c	0.004628	0.004999	0.005185
ρ	0.741086	0.90	0.98
σ_x	0.002719	0.001826	0.001422

Note: This table shows the maximum likelihood estimation of LRR model under different setting for all three models. The left column shows the result under natural restriction. The middle column imposes a more restricted constrain on the correlation coefficient that $\rho \in [0.9, 1)$. In the right column, we further assume that $\rho \in [0.98, 1)$.

obtain the MLE of all four parameters from the Kalman Filter by setting the parameter initial values equal to their baseline calibrations in Eq. (18). The results are presented in the first column of Table 2. It can easily be seen that the estimated consumption growth $\hat{\mu}_c$ is quite close to its baseline calibration 0.004215, but the estimation of the rest three parameters is significantly different from the baseline calibration. In the baseline Eq. (18), $\sigma_x \simeq \frac{\sigma_c}{10}$ while the MLE shows $\hat{\sigma}_x \simeq \frac{\hat{\sigma}_c}{2}$. This result contradicts Bansal and Yaron's assumption that x_t is not detectable in the short run since now $\hat{\sigma}_c$ and $\hat{\sigma}_x$ are of the same order of magnitude. The short-run consumption volatility will definitely be affected given the fairly large x_t component. One can also notice that $\hat{\rho} = 0.741$ is far less than its calibration, showing that even if the dynamics of the intangible component x_t follows Eq. (5), the data suggests this process is far less persistent than Bansal and Yaron (2004) assumed.

We also construct the empirical probability density functions (**PDF**) of all parameters using random walk Metropolis–Hasting algorithm. Initial values are set to be equal to their ML estimates. We generate a sample of size $50,000$ and plot the empirical PDFs of all four parameters in Fig. 3. The upper right panel shows the distribution of μ_c and the two panels in the bottom row show the PDFs of σ_x and σ_c, respectively. All three densities are symmetric and bell-shaped with the mean value close to the MLE, suggesting that $\mu_c, \sigma_c, \sigma_x$ are following normal distributions. In the upper left panel we present the empirical distribution of ρ. The peak of the distribution verifies our finding of MLE and its shape indicates most of density are located in within $(0.72, 0.8)$. Besides this, the shape of ρ does not carry any information on its underlying distribution. This finding doesn't support the assumptions made in Bansal and Yaron (2004) because data suggests it's very unlikely to have the persistence parameter ρ to be close to 0.98 (or any value greater than 0.9 even). We also notice that the distribution of σ_c and σ_x overlap

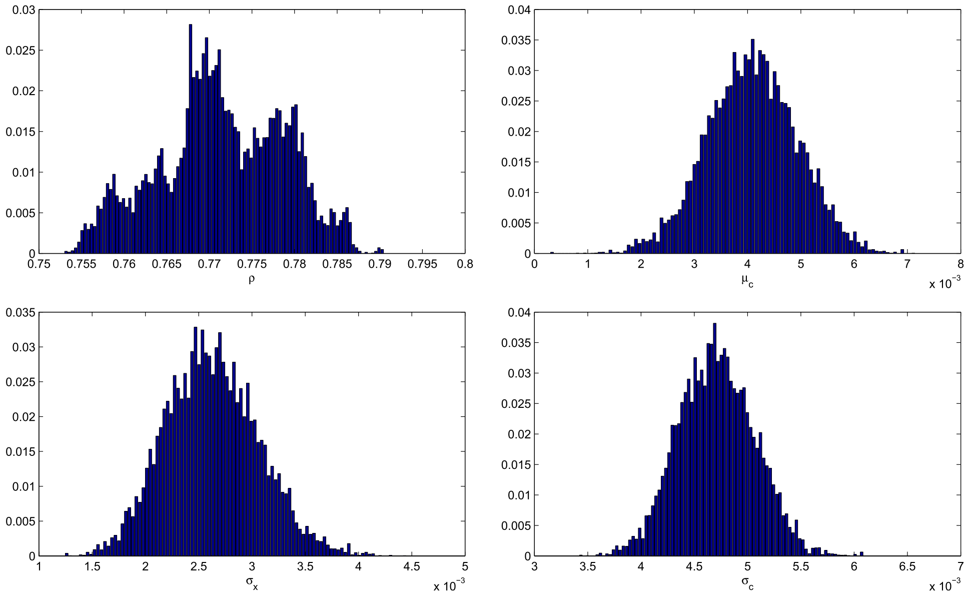

Figure 3. Empirical distribution for 4 parameters under the natural restriction $\rho \in (0, 1)$.

in $(0.35, 0.4)$, meaning we may fail to reject the hypothesis $\sigma_x = \sigma_c$. As we have already mentioned, this finding questions the intangibility assumption that x_t is a tiny component hard to be detected.

4.2.2. *Mild and extreme restrictions*

To investigate how μ_c, σ_c and σ_x are affected by the changes in ρ, we impose various restrictions on ρ and construct corresponding empirical distribution for each parameter. In the mild restriction case, we assume ρ can only take value in $[0.9, 1)$. The Kalman MLE results under this restriction are shown in the middle column of Table 2. Comparing with results in the first column, we find quite similar results: $\hat{\mu}_c$ approximately equal to its calibrated value but the rest three parameters deviate significantly. The value of $\hat{\sigma}_c$ increases and $\hat{\sigma}_x$ are decreasing comparing with the previous case, however $\hat{\sigma}_x \simeq \frac{\hat{\sigma}_c}{3}$ still indicates that σ_c and σ_x are approximately of the same order of magnitude. The $\hat{\rho}$ hit the lower bound at 0.9, still suggesting x_t is less persistent than it is in the original model. In sum, the point estimation suggests that x_t follows a less persistent and detectable process.

We also construct the empirical PDFs of all parameters in Fig. 4. In the upper left panel we present the PDF of ρ. One can read from the figure that even when we let ρ vary in $[0.9, 1)$, it is very unlikely to have ρ exceeded 0.945. The PDF of μ_c still have bell shape, however, the whole distribution started to shift to the left and has heavier tails. σ_x and σ_c still distributed like normal but now their bulks are

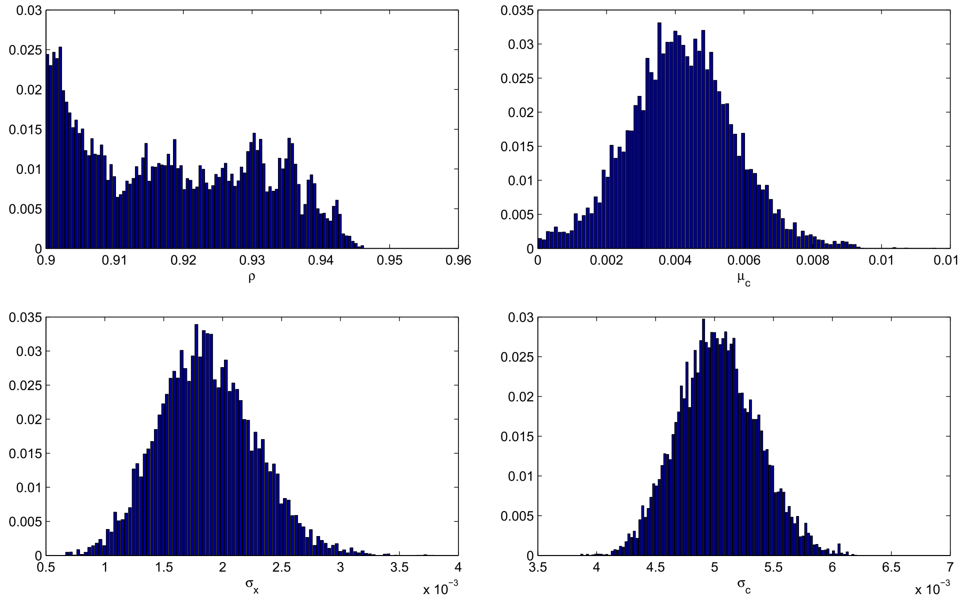

Figure 4. Empirical distribution for 4 parameters under the mild restriction $\rho \in [0.9, 1)$.

shifting to left and right, respectively. Although $\hat{\sigma}_x \simeq \frac{\hat{\sigma}_c}{3}$, we can reject the hypothesis $\sigma_x = \sigma_c$ given their PDFs.

For the extreme case, we let $\rho \in [0.98, 1)$. The MLE of all parameters are shown in the right column of Table 2. In this case, the likelihood is maximized when $\hat{\rho}$ takes the value at the lower bound of 0.98, $\hat{\sigma}_c$ continues increasing while $\hat{\sigma}_x$ continues decreasing; $\hat{\mu}_c$ deviates from the baseline calibration 0.0042. The empirical PDFs are shown in Fig. 5. Interestingly, the PDF of μ_c no longer possesses a normal shape, instead we now have a highly right-skewed distribution. The bulks of σ_x and σ_c continue shifting, and therefore, we retain the conclusion that we can reject $\sigma_x = \sigma_c$.

Figures 3–5 indicate that ρ is the key parameter which controls the dynamics of the LRR model. As ρ increases, The distributions of μ_c and σ_x shift to the left and the distribution of σ_c shifts to the right. The results show that the intangibility of x_t mostly comes from the persistence restriction imposed on ρ. The findings indicate that the data do not provide enough support for introducing this highly persistent x_t.

4.3. *Diagnostic test*

We present the model performance by using the parameter estimations from MLE under the natural restriction in Fig. 6. The figure shows that LRR model will not satisfy the Hansen–Jagannathan min-variance boundary unless we have the RRA γ approximately equal to 70. Compared to the performance of usual Epstein–Zin

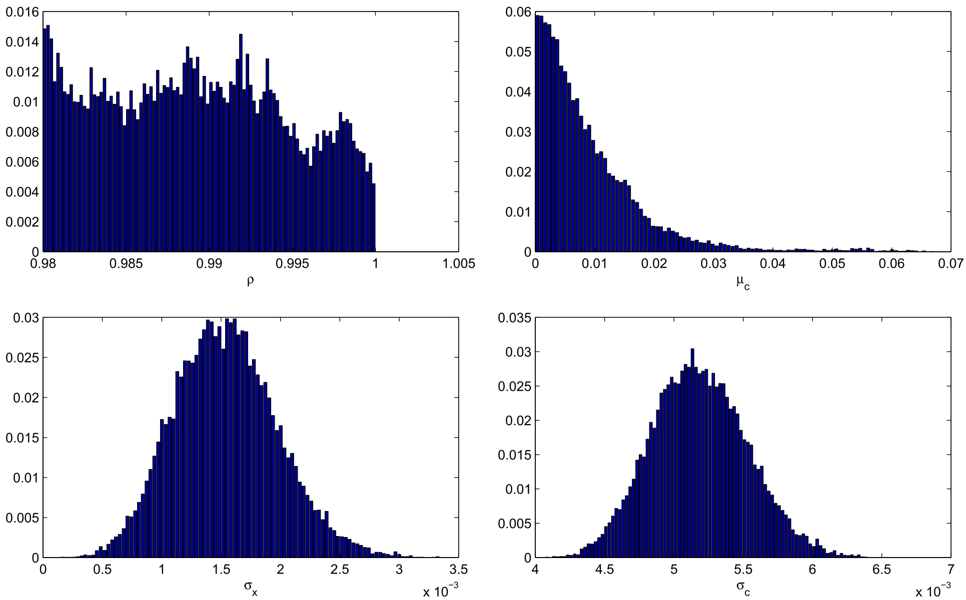

Figure 5. Empirical distribution for 4 parameters under the extreme restriction $\rho \in [0.98, 1)$.

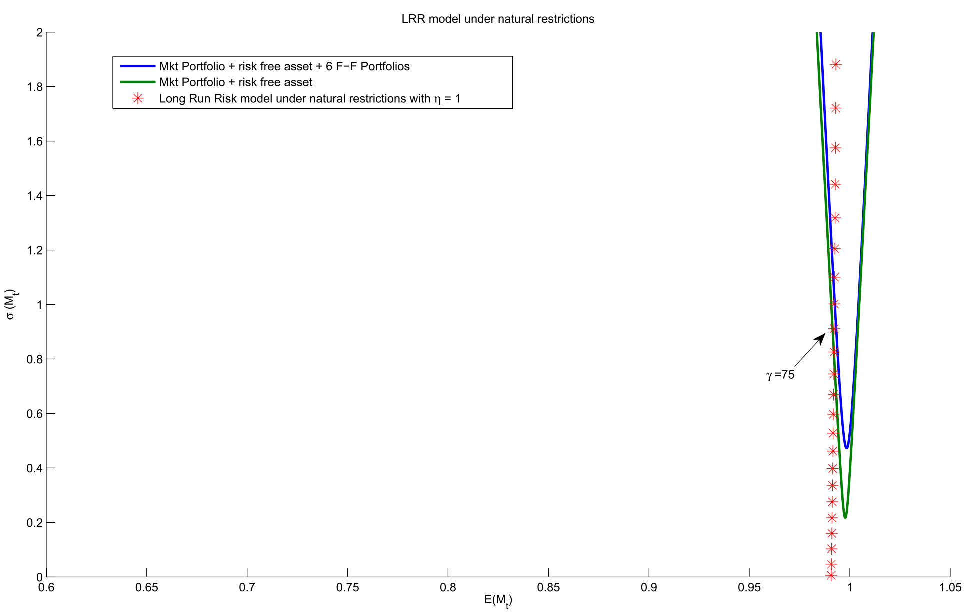

Figure 6. Model performance under natural restriction.

recursive preference in Panel (B) of Fig. 2, we do not find any significant improvement for the LRR model.

5. Conclusion

In this paper, we use the Kalman Filter and Maximum Likelihood to estimate and reevaluate the credibility of the long-run risk model as in Bansal and Yaron (2004). We show that the persistent parameter ρ is the key component that governs the dynamics of the LRR model. The model performance strongly relies on the persistence restriction imposed on ρ. Under the more reasonable restrictions, the distribution of σ_c and σ_x obtained from Metropolis–Hastings algorithm lead to rejection of the null hypothesis that $\sigma_x = \sigma_c$, which will put the small magnitude assumption under question. Also, under estimated parameter values, the LRR model still requires relative risk aversion to be around 70 for the model to fit the US data. Model performance improves only when we impose the persistence restriction on the unobserved component as Bansal and Yaron (2004) did in their calibration. However, imposing such restriction seems to lack empirical support from the data.

Acknowledgments

This research is supported by China Postdoctoral Science Foundation (Grant No. 2018M641302).

Appendix A. Value Function

Guess the solution to Eq. (3) takes the form $u_t = k_0 + k_1 c_t + k_2 x_t$, where k_0, k_1 and k_2 are constants. Forward the value function for one period then substitute into Eq. (3), we will have

$$u_t = (1-\beta)c_t + \frac{\beta}{1-\gamma}\log\left[E_t(\exp((1-\gamma)(k_0 + k_1 c_{t+1} + k_2 x_{t+1})))\right]. \tag{A.1}$$

Substitute Eqs. (4) and (5) in Eq. (A.1)

$$\begin{aligned} u_t =& (1-\beta)c_t + \frac{\beta}{1-\gamma}\log\left[\exp((1-\gamma)(k_0 + k_1(c_t+\mu_c) + (k_1 + \rho k_2)x_t))\right] \\ &+ \frac{\beta}{1-\gamma}\log E_t\left[\exp((1-\gamma)(k_1\sigma_c\varepsilon^c_{t+1} + k_2\sigma_x\varepsilon^x_{t+1}))\right] \\ =& \beta\left(k_0 + k_1\mu_c + (1-\gamma)\left(\frac{k_1^2\sigma_c^2 + k_2^2\sigma_x^2}{2}\right)\right) + (1-\beta+\beta k_1)c_t \\ &+ \beta(k_1 + k_2\rho)x_t. \end{aligned} \tag{A.2}$$

The last equality comes from the log normal assumption. Notice the LHS of Eq. (A.2) will be exactly $k_0 + k_1 c_t + k_2 x_t$ from our guess. By matching the coefficients of each variables in Eq. (A.2), we will have

$$
\begin{aligned}
k_0 &= \frac{\beta}{1-\beta}\left(\mu_c + (1-\gamma)\left(\frac{\sigma_c^2}{2} + \frac{\beta^2}{2(1-\beta\rho)^2}\sigma_x^2\right)\right),\\
k_1 &= 1,\\
k_2 &= \frac{\beta}{1-\beta\rho}.
\end{aligned}
$$

Appendix B. SDF

If we consider a general form of Epstein–Zin preference with EIS $= \eta$ and $rra = \gamma$, let $\mu_{t+1} = (E_t U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}$, the SDF can be written as

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\eta} \cdot \left(\frac{U_{t+1}}{\mu_{t+1}}\right)^{\eta-\gamma} \tag{B.1}$$

For the special case with $\eta = 1$, we have

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-1} \cdot \left(\frac{\exp[(1-\gamma)u_{t+1}]}{E_t[\exp((1-\gamma)u_{t+1})]}\right) \tag{B.2}$$

Take logarithm on both sides, we will have

$$m_{t+1} = \log\beta - \Delta c_{t+1} + (1-\gamma)u_{t+1} - \log E_t[\exp((1-\gamma)u_{t+1})]. \tag{B.3}$$

Substitute Eqs. (6), (4) and (5) in Eq. (B.4),

$$
\begin{aligned}
m_{t+1} = {}& \log\beta - (\mu_c + x_t) - \frac{1}{2}(1-\gamma)^2\left(\sigma_c^2 + \left(\frac{\beta\sigma_x}{1-\beta\rho}\right)^2\right)\\
& + (1-\gamma)\frac{\beta\sigma_x}{1-\beta\rho}\varepsilon_{t+1}^x - \gamma\sigma_c\varepsilon_{t+1}^c.
\end{aligned} \tag{B.4}
$$

Based on the normality assumption of innovation $\{\varepsilon_t^c, \varepsilon_t^x\}$, m_{t+1} will also follow normal distribution and M_{t+1} will follow log-normal distribution. Therefore, the unconditional mean and standard deviation of M_{t+1} will be

$$E(M_{t+1}) = \beta\exp\left(-\mu_c + \left(\gamma - \frac{1}{2}\right)\sigma_c^2 + \frac{\sigma_x^2}{2(1-\rho^2)}\right)$$

$$\sigma_{M_{t+1}} = E(M_{t+1}) \cdot \sqrt{\exp\left(\frac{\sigma_x^2}{1-\rho^2} + (\gamma\sigma_c)^2 + \left(\frac{(1-\gamma)\beta\sigma_x}{1-\beta\rho}\right)^2\right) - 1}.$$

References

Babbs, SH and KB Nowman (1999). Kalman filtering of generalized vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(1), 115–130.

Bansal, R and A Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4), 1481–1509.

Barro, RJ (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3), 823–866.

Campbell, JY and JH Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *The Journal of Political Economy*, 107(2), 205–251.

Cochrane, JH (2017). Macro-Finance. *Review of Finance*, 21(3), 945–985.

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

Epstein, LG and SE Zin (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99(2), 263–286.

Fama, EF and KR French (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1), 23–49.

Hamilton, JD (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hansen, LP and R Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99(2), 225–262.

Hansen, LP and R Jagannathan (1997). Assessing specification errors in stochastic discount factor models. *The Journal of Finance*, 52(2), 557–590.

Harvey, AC (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Mehra, R and EC Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15, 145–161.

Schwartz, E and JE Smith (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7), 893–911.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24, 401–421.