A Bayesian estimation approach of random switching exponential smoothing with application to credit forecast

Citation for published version:
Wang, R, Wang, T, Qian, Z & Hu, S 2023, 'A Bayesian estimation approach of random switching exponential smoothing with application to credit forecast', *Finance Research Letters*, vol. 58, no. Part C, 104525. https://doi.org/10.1016/j.frl.2023.104525

Digital Object Identifier (DOI):
10.1016/j.frl.2023.104525

Published In:
Finance Research Letters

A Bayesian estimation approach of random switching exponential smoothing with application to credit forecast

Renhe Wang[a], Tong Wang[b], Zhiyong Qian[a] and Shulan Hu[a,*,4]

[a]Zhongnan University of Economics and Law, Wuhan, 430073, Hubei, Peoples Republic of China
[b]Business School, University of Edinburgh, Edinburgh, EH8 9JS, Scotland, United Kingdom

ARTICLE INFO

Keywords:
Random switching exponential smoothing
Precision-based algorithms
Bayesian estimation
Forecasting
Credit

ABSTRACT

We introduce an efficient Markov Chain Monte Carlo sampler in precision-based algorithms for the estimation of the Random Switching Exponential Smoothing model, a versatile forecasting mechanism for time series data characterized with changing trends. Through a series of simulation experiments, RC-MCMC exhibits superior parameter estimation accuracy, particularly for datasets featuring low persistence trends. Furthermore, an empirical evaluation using the Bank for International Settlements' quarterly time series data on the non-financial sector's total credit relative to GDP validates the findings. The out-of-sample results indicate that the proposed approach outperforms its counterparts in estimating and forecasting accuracy for trending time series data.

1. Introduction

Exponential smoothing is among the most commonly widespread class of methods for time series analysis and forecasting applied in the fields of economics, finance, and operations management (e.g. Hyndman et al., 2008; Gardner, 2006; Ord et al., 2017; Kourentzes et al., 2019). Following the improvement of Holt's original simple exponential smoothing, the "single source of error" (SSOE) framework was first introduced by Hyndman et al. (2008) and McKenzie and Gardner (2010), and then Sbrana and Silvestrini (2014) proposed a "multiple source of error" (MSOE) model that allows for random switching between simple exponential smoothing and trend exponential smoothing by setting the coefficients randomly, called random switching exponential smoothing. Therefore it is possible to control, in a flexible manner, the random changing dynamic behaviour of the time series. Sbrana and Silvestrini (2019) developed a new, fast and efficient method for estimating the model based the algebraic link between the model's structural parameters and the steady-state Kalman gain vector rather than the approach within a restricted region in the previous work.

Bayesian inference methods in state space models have been influenced by some studies (Carter and Kohn, 1994; De Jong and Shephard, 1995; Durbin and Koopman, 2002), e.g. Kalman Filter for linear state space models. Kim et al. (1998) focused on stochastic volatility models based on an auxiliary sampling approach and then used Gibbs to estimate this class of models, and some improved estimate methods for state space models have since been proposed (e.g. Moura and Turatti, 2014; Monache and Petrella, 2019). A series of papers by Chan further developed an efficient Markov chain Monte Carlo (MCMC) sampler for estimating this class of models (Chan and Jeliazkov, 2009; Chan, 2013; Chan and Grant, 2016), which also verified that the precision-based algorithms are stable and reliable. In this paper, we estimate random switching exponential smoothing in the MSOE framework through precision-based algorithms, denoted as the Random Coefficient MCMC method (RC-MCMC), which is based on the feature of banded precision matrices to enhance the efficiency of parameter estimation.

Within the simulation study, we compare the accuracy of parameter estimation and forecasting performance of the MSOE form model, as estimated through two different methods: the RC-MCMC approach and the direct method proposed by Sbrana and Silvestrini (2019), referred to as RC-SSPACE. Simulation results highlight the superior performance of the RC-MCMC in estimating structural parameters compared to the RC-SSPACE. While both methods display comparable forecasting performances overall, RC-SSPACE performance tends to lag slightly when dealing with low trend persistence. This observation is further substantiated in empirical application, which utilizes quarterly credit-to-GDP data published by the Bank for International Settlements (BIS). Through these results, we provide additional

*Corresponding author. School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan, 430073, People's Republic of China.
E-MAIL ADDRESS: hu_shulan@zuel.edu.cn (S. Hu)

evidence supporting the superiority of the RC-MCMC approach over its competitor RC-SSPACE, particularly in an economic context. The RC-MCMC method, when applied to random switching exponential smoothing, attains commendable estimation precision. In contrast to many estimation methods that require optimal computational steps and might result in parameter estimates that do not correspond with reality, the parameter estimates obtained through the RC-MCMC method utilizing a band-precision matrix are found to be more practically significant.

The rest of this paper is organized as follows: Section 2 introduces the model specification. In Section 3, an efficient posterior simulator is developed. In Section 4, some properties of the two methods are compared via Monte Carlo simulations. Section 5 presents an empirical application and the last section concludes the findings.

2. Random Coefficient State-Space Models

The random coefficient state space model based on the generalized MSOE framework was proposed by McKenzie and Gardner (2010), the specific form model (1) is as follows:

$$\begin{aligned} y_t &= l_{t-1} + A_t b_{t-1} + \epsilon_t,\ \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2), \\ l_t &= l_{t-1} + A_t b_{t-1} + \eta_t,\ \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \\ b_t &= A_t b_{t-1} + \xi_t,\ \xi_t \sim \mathcal{N}(0, \sigma_\xi^2), \end{aligned} \tag{1}$$

where $t = 1 \ldots T$, and $\mathcal{N}(\cdot,\cdot)$ denotes independent and identically normal distribution, y_t is the observation at time t, l_t is the stochastic trend, b_t is the slope of its stochastic trend, the term A_t is a sequence of independent, identically distributed binary random variables with probabilities:

$$P(A_t = 1) = \phi,\ P(A_t = 0) = 1 - \phi,\ 0 \le \phi \le 1. \tag{2}$$

The state space representation of this model can be expressed in the following form:

$$\begin{aligned} y_t &= \mathbf{z}_t' \boldsymbol{\alpha}_{t-1} + \epsilon_t, \\ \boldsymbol{\alpha}_t &= \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{R}\mathbf{u}_t, \end{aligned} \tag{3}$$

where $\mathbf{z}_t = \begin{bmatrix} 1 \\ A_t \end{bmatrix}$, $\mathbf{T}_t = \begin{bmatrix} 1 & A_t \\ 0 & A_t \end{bmatrix}$, $\boldsymbol{\alpha}_t = \begin{bmatrix} l_t \\ b_t \end{bmatrix}$, $\mathbf{R} = \begin{bmatrix} \sigma_\eta & 0 \\ 0 & \sigma_\xi \end{bmatrix}$, and $\mathbf{u}_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}_2\right)$, $\mathbf{I}_2$ is an 2×2 identity matrix, i.e., $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and we denote the parameter vector as $\boldsymbol{\theta} = (\phi, \sigma_\epsilon^2, \sigma_\eta^2, \sigma_\xi^2)'$.

3. Bayesian Estimation

3.1. Likelihood evaluation

We first derive the joint distribution of the observations $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)'$. For notational convenience, we define $\boldsymbol{l} = (l_1, l_2, \ldots, l_T)'$ and $\boldsymbol{b} = (b_1, b_2, \ldots, b_T)'$. Moreover, innovation is also represented in vector form $\boldsymbol{u}_{\epsilon-\eta} = (\epsilon_1 - \eta_1, \epsilon_2 - \eta_2, \ldots, \epsilon_T - \eta_T)'$, similarly, $\boldsymbol{u}_\eta = (\eta_1, \eta_2, \ldots, \eta_T)'$ and $\boldsymbol{u}_\xi = (\xi_1, \xi_2, \ldots, \xi_T)'$ from model (1).

The observation equation of model (1) can be rewritten in the following matrix form:

$$\boldsymbol{y} = \boldsymbol{l} + \boldsymbol{u}_{\epsilon-\eta}, \tag{4}$$

where $\boldsymbol{u}_{\epsilon-\eta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{S}_y)$, $\boldsymbol{S}_{\boldsymbol{y}} = diag(\sigma_\epsilon^2 + \sigma_\eta^2, \sigma_\epsilon^2 + \sigma_\eta^2, \ldots, \sigma_\epsilon^2 + \sigma_\eta^2)$. By a simple variable transformation, we can obtain $(\boldsymbol{y}|\boldsymbol{l}, \sigma_\epsilon^2, \sigma_\eta^2) \sim \mathcal{N}(\boldsymbol{l}, \boldsymbol{S}_y)$. The conditional log-likelihood function of model (1) can be expressed by the prediction error decomposition:

$$\log P(\boldsymbol{y}|\boldsymbol{l}, \sigma_\epsilon^2, \sigma_\eta^2) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log(\sigma_\epsilon^2 + \sigma_\eta^2) - (\boldsymbol{y} - \boldsymbol{l})' \boldsymbol{S}_{\boldsymbol{y}}^{-1} (\boldsymbol{y} - \boldsymbol{l}). \tag{5}$$

Next, regarding the measurement equation $\boldsymbol{l}$, the following relationship can be observed in the sequence:

$$\begin{aligned} l_1 &= l_0 + A_1 b_0 + \eta_1 = l_0 + b_1 + \eta_1 - \xi_1 \\ l_2 - l_1 &= A_2 b_1 + \eta_2 = b_2 + \eta_2 - \xi_2 \\ l_3 - l_2 &= A_3 b_2 + \eta_3 = b_3 + \eta_3 - \xi_3 \\ &\vdots \\ l_T - l_{T-1} &= A_T b_{T-1} + \eta_T = b_T + \eta_T - \xi_T; \end{aligned}$$

similarly, the vector sequence $\boldsymbol{b}$ also has a similar structure, $\mathbf{H}_l$ and $\mathbf{H}_b$ are the following lower triangular matrixes:

$$\mathbf{H}_l = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}_{T\times T}, \ \mathbf{H}_b = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -A_2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -A_3 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -A_T & 1 \end{pmatrix}_{T\times T}.$$

by using matrix notation:

$$\mathbf{H}_l \boldsymbol{l} = \boldsymbol{\mu}_l + \boldsymbol{u}_\eta, \ \mathbf{H}_b \boldsymbol{b} \quad = \boldsymbol{\mu}_b + \boldsymbol{u}_\xi, \tag{6}$$

by a simple variable transformation, we can obtain:

$$\boldsymbol{l}|\boldsymbol{b} \sim \mathcal{N}(\mathbf{H}_l^{-1}\boldsymbol{\mu}_l, \boldsymbol{S}_l), \quad \boldsymbol{b}|\boldsymbol{A} \sim \mathcal{N}(\mathbf{H}_b^{-1}\boldsymbol{\mu}_b, \boldsymbol{S}_b), \tag{7}$$

where $\boldsymbol{\mu}_l = \boldsymbol{b} + (l_0, 0, 0, \dots, 0)'$, $\boldsymbol{\mu}_b = (A_1 b_0, 0, 0, \dots, 0)'$, $\boldsymbol{S}_l = (\mathbf{H}_l' \Sigma_l^{-1} \mathbf{H}_l)^{-1}$, $\boldsymbol{S}_b = (\mathbf{H}_b' \Sigma_b^{-1} \mathbf{H}_b)^{-1}$, and $\Sigma_l = diag(\sigma_\eta^2 + \sigma_\xi^2, \sigma_\eta^2 + \sigma_\xi^2, \dots, \sigma_\eta^2 + \sigma_\xi^2)$, $\Sigma_b = diag(\sigma_\xi^2, \sigma_\xi^2, \dots, \sigma_\xi^2)$. It is important to note that the banded matrix $\mathbf{H}_l$ and $\mathbf{H}_b$ with bandwidth q contain only $(T - q/2)(q+1) < T(q+1)$ non-zero elements and the bandwidth q satisfy $q \ll T < T^2$. This greatly reduces the computational complexity of posterior analysis in estimation work.

3.2. Posterior analysis

Now we discuss an effective posterior sampler in model (1). Such estimation method was first introduced in the stochastic volatility state-space model by Kim et al. (1998), and have been further improved by algorithmic advances in Chan and Jeliazkov (2009) and Chan and Grant (2016).

We develop a MCMC algorithm in which posterior draws can be obtained by sequentially sampling from:

1. Step 1: $P(\phi|\boldsymbol{A})$;
2. Step 2: $P(\boldsymbol{A}|\boldsymbol{b}, \phi, \sigma_\xi^2)$;
3. Step 3: $P(\boldsymbol{l}|\boldsymbol{y}, \boldsymbol{A}, \sigma_\epsilon^2, \sigma_\eta^2)$;
4. Step 4: $P(\boldsymbol{b}|\boldsymbol{l}, \boldsymbol{A}, \sigma_\eta^2, \sigma_\xi^2)$;
5. Step 5: $P(\sigma_\epsilon^2|\boldsymbol{y}, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{A})$;
6. Step 6: $P(\sigma_\eta^2|\boldsymbol{l}, \boldsymbol{b}, \boldsymbol{A})$;
7. Step 7: $P(\sigma_\xi^2|\boldsymbol{b}, \boldsymbol{A})$.

In Step 1, according to conjugate prior property and $\phi \sim Beta(k_a, k_b)$, the conditional posterior distribution is given as follows:

$$(\phi \mid \boldsymbol{A}) \sim Beta\left(k_a + \sum_{t=1}^{T} A_t, \ k_b + T - \sum_{t=1}^{T} A_t\right), \tag{8}$$

In Step 2 of the MCMC sampler, the conditional posterior probabilities of each variable value of A_t are

$$\begin{aligned} P(A_t = 1 \mid \cdot) &\propto \phi exp(-(1/2)(\hat{\mathbf{H}}_b \boldsymbol{b})' \Sigma_b^{-1/2} (\hat{\mathbf{H}}_b \boldsymbol{b})), \\ P(A_t = 0 \mid \cdot) &\propto (1-\phi) exp(-(1/2)\boldsymbol{b}' \Sigma_b^{-1/2} \boldsymbol{b}), \end{aligned} \tag{9}$$

where

$$\hat{\mathbf{H}}_b = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}_{T\times T},$$

therefore $A_t = 1$ with probability $P\left(A_t = 1 \mid \boldsymbol{b}, \cdot\right) / \left(P\left(A_t = 0 \mid \boldsymbol{b}, \cdot\right) + P\left(A_t = 1 \mid \boldsymbol{b}, \cdot\right)\right)$, otherwise $A_t = 0$.

Next, for Step 3 $P(\boldsymbol{l}|\boldsymbol{y},\boldsymbol{A},\sigma_\epsilon^2,\sigma_\eta^2)$, based on the log-likelihood function (5), the conditional distribution of the sample can be calculated as follows:

$$(\boldsymbol{l}|\boldsymbol{y},\boldsymbol{A},\sigma_\epsilon^2,\sigma_\eta^2) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{l}},\hat{\Sigma}_{\boldsymbol{l}}), \tag{10}$$

where $\hat{\Sigma}_{\boldsymbol{l}} = (\mathbf{H}_{\boldsymbol{l}}'\Sigma_{\boldsymbol{l}}^{-1}\mathbf{H}_{\boldsymbol{l}} + \boldsymbol{S}_{\boldsymbol{y}}{}^{-1})^{-1}$ and $\hat{\boldsymbol{\mu}}_{\boldsymbol{l}} = \hat{\Sigma}_{\boldsymbol{l}}(\boldsymbol{S}_{\boldsymbol{y}}{}^{-1}\boldsymbol{y} + \boldsymbol{S}_{\boldsymbol{l}}^{-1}\mathbf{H}_{\boldsymbol{l}}^{-1}\boldsymbol{\mu}_{\boldsymbol{l}})$, for Step 4, by similarly calculating, we have

$$(\boldsymbol{b}|\boldsymbol{l},\boldsymbol{A},\sigma_\eta^2,\sigma_\xi^2) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{b}},\hat{\Sigma}_{\boldsymbol{b}}), \tag{11}$$

where $\hat{\Sigma}_{\boldsymbol{b}} = (\Sigma_l^{-1} + \boldsymbol{S}_{\boldsymbol{b}}{}^{-1})^{-1}$ and $\hat{\boldsymbol{\mu}}_{\boldsymbol{b}} = \hat{\Sigma}_{\boldsymbol{b}}(\Sigma_l^{-1}(\mathbf{H}_{\boldsymbol{l}}\boldsymbol{l} - (l_0, 0, 0, \dots, 0)') + \boldsymbol{S}_{\boldsymbol{b}}{}^{-1}\mathbf{H}_{\boldsymbol{b}}^{-1}\boldsymbol{\mu}_{\boldsymbol{b}})$.

Lastly, note that $P(\sigma_\epsilon^2)$, $P(\sigma_\eta^2)$ and $P(\sigma_\xi^2)$ are inverse-gamma densities and can therefore be sampled through standard methods:

$$(\sigma_\epsilon^2|\boldsymbol{y},\boldsymbol{l},\boldsymbol{b},\boldsymbol{A}) \sim IG(\nu_\epsilon + (T-1)/2, \tilde{S}_\epsilon),$$
$$(\sigma_\eta^2|\boldsymbol{l},\boldsymbol{b},\boldsymbol{A}) \sim IG(\nu_\eta + (T-1)/2, \tilde{S}_\eta),$$
$$(\sigma_\xi^2|\boldsymbol{b},\boldsymbol{A}) \sim IG(\nu_\xi + (T-1)/2, \tilde{S}_\xi),$$

where $\tilde{S}_\epsilon = S_\epsilon + \sum_{t=2}^T (y_t - l_{t-1} - A_t b_{t-1})^2/2$, $\tilde{S}_\eta = S_\eta + \sum_{t=2}^T (l_t - l_{t-1} - A_t b_{t-1})^2/2$ and $\tilde{S}_\xi = S_\xi + \sum_{t=2}^T (b_t - A_t b_{t-1})^2/2$.

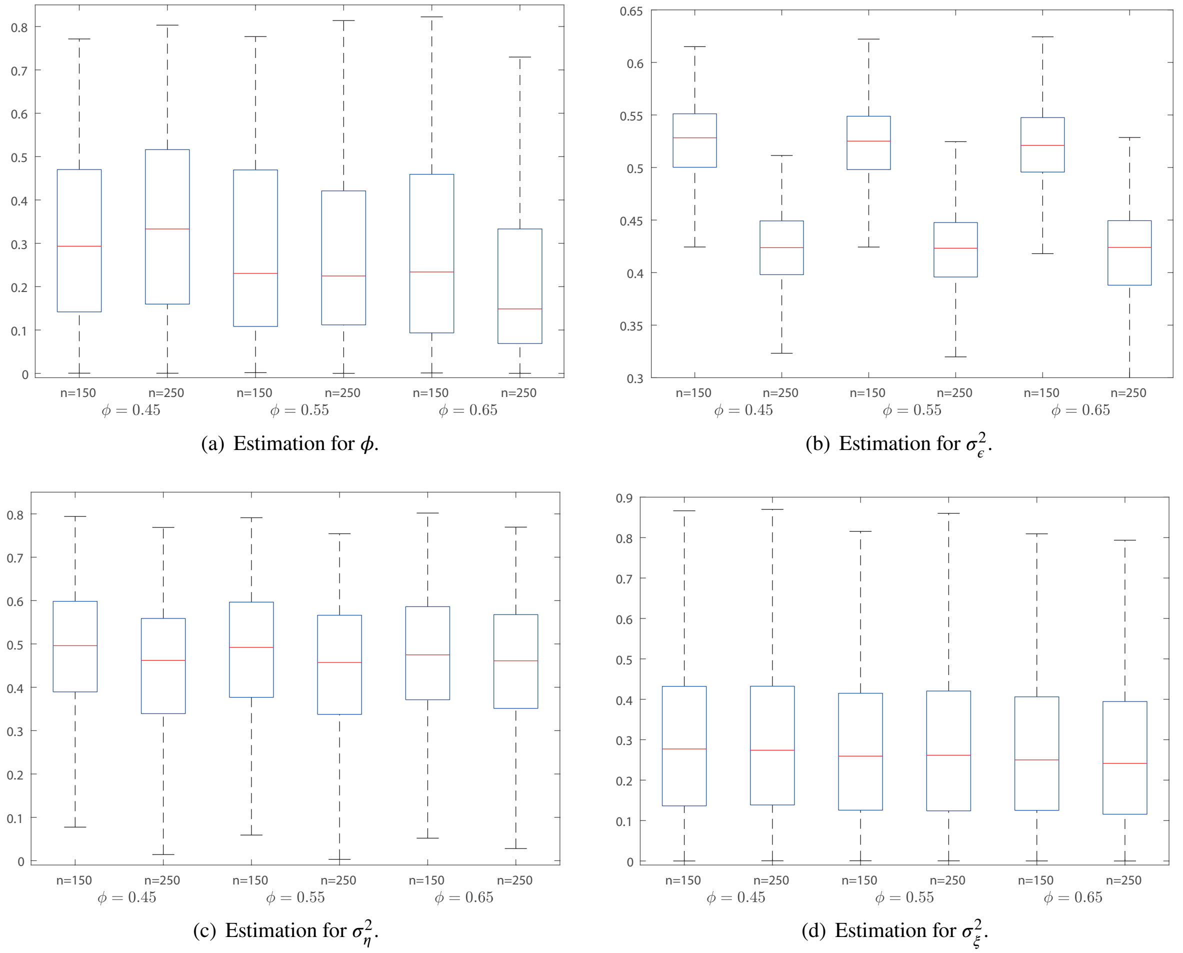
4. Simulation study

Our simulation settings are as follows: To generate the data, we set the number of observations T to either 100 or 250 and set the persistence parameter, ϕ, to 0.95, 0.85, or 0.75, in addition, to compare exponential models with more random coefficient changes, we also set the values to 0.65, 0.55, and 0.45. For each combination of T and ϕ, we simulate time series data of length 1000. Specifically, we draw T observations from the simulate time series data to conduct the simulation experiment. For each series, we first choose σ_ϵ^2 from $Uniform(0.5, 2.5)$. As for the innovation variances σ_η^2 and σ_ξ^2 are determined as σ_ϵ^2/c and σ_ϵ^2/d, where c and d are independently chosen from $Uniform(1, 10)$.

Table 1
Monte Carlo simulations results-APE for ϕ,σ_ϵ^2,σ_η^2,σ_ξ^2. This table reports the mean APE obtained for the estimates of $\phi = 0.75, 0.85, 0.95$.

ϕ		$n = 150$		$n = 250$	
		Mean MCMC	Mean SSPACE	Mean MCMC	Mean SSPACE
$\phi = 0.75$	ϕ	**0.272**	0.382	**0.208**	0.319
	σ_ϵ^2	0.523	**0.080**	0.418	**0.067**
	σ_η^2	**0.470**	0.698	**0.419**	0.663
	σ_ξ^2	**0.280**	11.658	**0.267**	8.751
$\phi = 0.85$	ϕ	0.258	**0.208**	0.180	**0.149**
	σ_ϵ^2	0.520	**0.080**	0.418	**0.066**
	σ_η^2	**0.459**	0.717	**0.408**	0.661
	σ_ξ^2	**0.283**	6.014	**0.261**	3.633
$\phi = 0.95$	ϕ	0.186	**0.062**	0.127	**0.039**
	σ_ϵ^2	0.521	**0.086**	0.419	**0.067**
	σ_η^2	**0.458**	0.659	**0.409**	0.543
	σ_ξ^2	**0.259**	1.456	**0.227**	0.349

We compare the precision of the parameter estimates via Absolute Percentage Errors (APE). The simulation results are in Table 1, Figure 1 and Figure 3, different from Sbrana and Silvestrini (2019), we apply a MCMC algorithm, and all the posterior momments are based on 5,000 draws from the MCMC algorithm introduced in after a burn-in of 1,000. Overall, the estimation of each parameter becomes more accurate as the sample size increases, with RC-MCMC performing slightly better than RC-SSPACE, where RC-SSPACE has consistently lower estimation error for σ_ϵ^2, but the RC-MCMC estimate as a whole has been in a more stable state.



(a) Estimation for ϕ.

(b) Estimation for σ_ϵ^2.

(c) Estimation for σ_η^2.

(d) Estimation for σ_ξ^2.

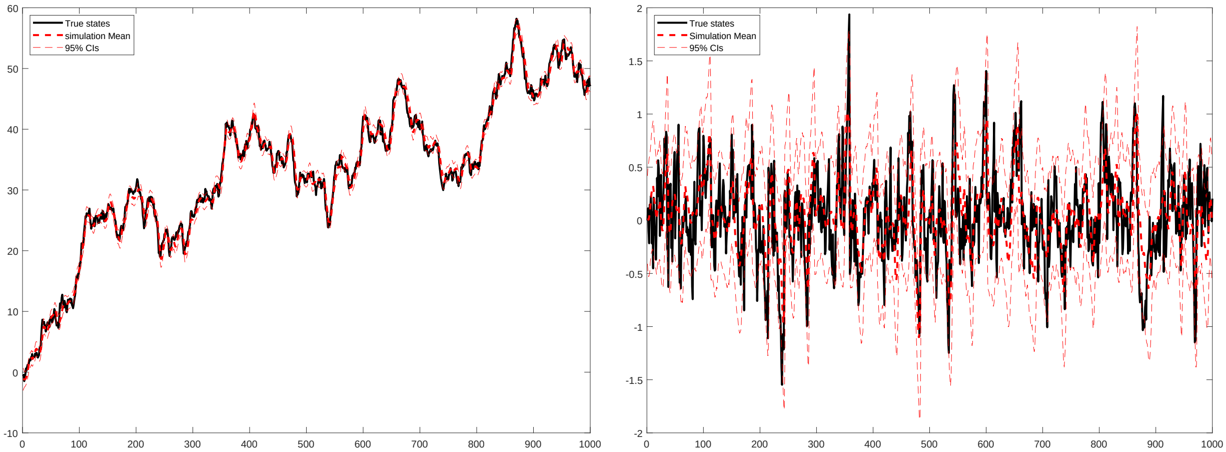
Figure 1: Monte Carlo simulations results-APE for ϕ,σ_ϵ^2,σ_η^2,σ_ξ^2. This figure reports the boxplots of the APE obtained for the estimates of ϕ,σ_ϵ^2,σ_η^2,σ_ξ^2 using RC-MCMC.

Table 2
The estimation accuracy ratio of RC-MCMC based on the 95% confidence interval.

	$\phi = 0.45$	$\phi = 0.55$	$\phi = 0.65$	$\phi = 0.75$	$\phi = 0.85$	$\phi = 0.95$	Mean
l_t	0.806	0.800	0.801	0.800	0.758	0.675	0.773
b_t	0.803	0.804	0.816	0.843	0.847	0.886	0.833

In the group with higher trend persistence, i.e., ϕ set at 0.75, 0.85, and 0.95, RC-MCMC and RC-SSPACE performed relatively similarly, with the former performing better in the parameter σ_ϵ^2, while the latter performed better in estimating σ_η^2 and σ_ξ^2. In the comparison of the more randomly switching coefficients, when ϕ is set to 0.45, 0.55 and 0.65, a comparison of Figure 1 and Figure 3 shows that RC-SSPACE in the set with low trend persistence performs not well, especially for σ_ξ^2, while the estimation of our proposed method is still more reasonable.

To investigate the estimation accuracy of RC-MCMC, we further examine the proportion of true values in the 95% confidence interval through simulations. Table 2 and Figure 2 present the corresponding simulation results for the case where $T = 100$ and $n = 1000$. It can be observed that RC-MCMC exhibits favorable estimation accuracy.



(a) l_t: Smooth States and Simulation Statistics.

(b) b_t: Smooth States and Simulation Statistics.

Figure 2: Smooth States and Simulation Statistics (Results of a single simulation with $\phi = 0.75$).

5. Application to credit forcast

The analysis of credit risk, as well as economic performance across various countries or regions, is critical. These factors play a significant role in the formulation of appropriate macroeconomic policies, and they are indispensable to maintaining financial stability and effective risk management. Research on credit risk enables financial institutins to comprehend the origins and degrees of risk more effectively, allowing for the development of scientifically sound and logical risk management strategies. In the realm of macroeconomic policy formulation, credit risk research plays a pivotal role. The credit-to-GDP ratio is a critical barometer of economic and financial risks. It quantifies the proportion of the size of credit in relation to the Gross Domestic Product (GDP).

We now apply the proposed method to analyze an alternate dataset of quarterly data. The results will be contrasted with those generated using the method developed by Sbrana and Silvestrini (2019). This dataset comprises time-series data related to the total credit, encompassing loans and debt securities, provided to the non-financial sector. These credit figures are expressed as a proportion of the Gross Domestic Product (GDP), forming the credit-to-GDP ratio[1]. The data hails from the Bank for International Settlements[2], covering approximately 50 economies with data spanning an average of over 45 years. The range extends back to the 1940s and 1950s for some countries[3]. Similar to Sbrana and Silvestrini (2020) which utilized this data to compare the MSOE and SSOE frameworks in the context of damped trend models, our study seeks to compare the out-of-sample forecast performance of these two distinct methods.

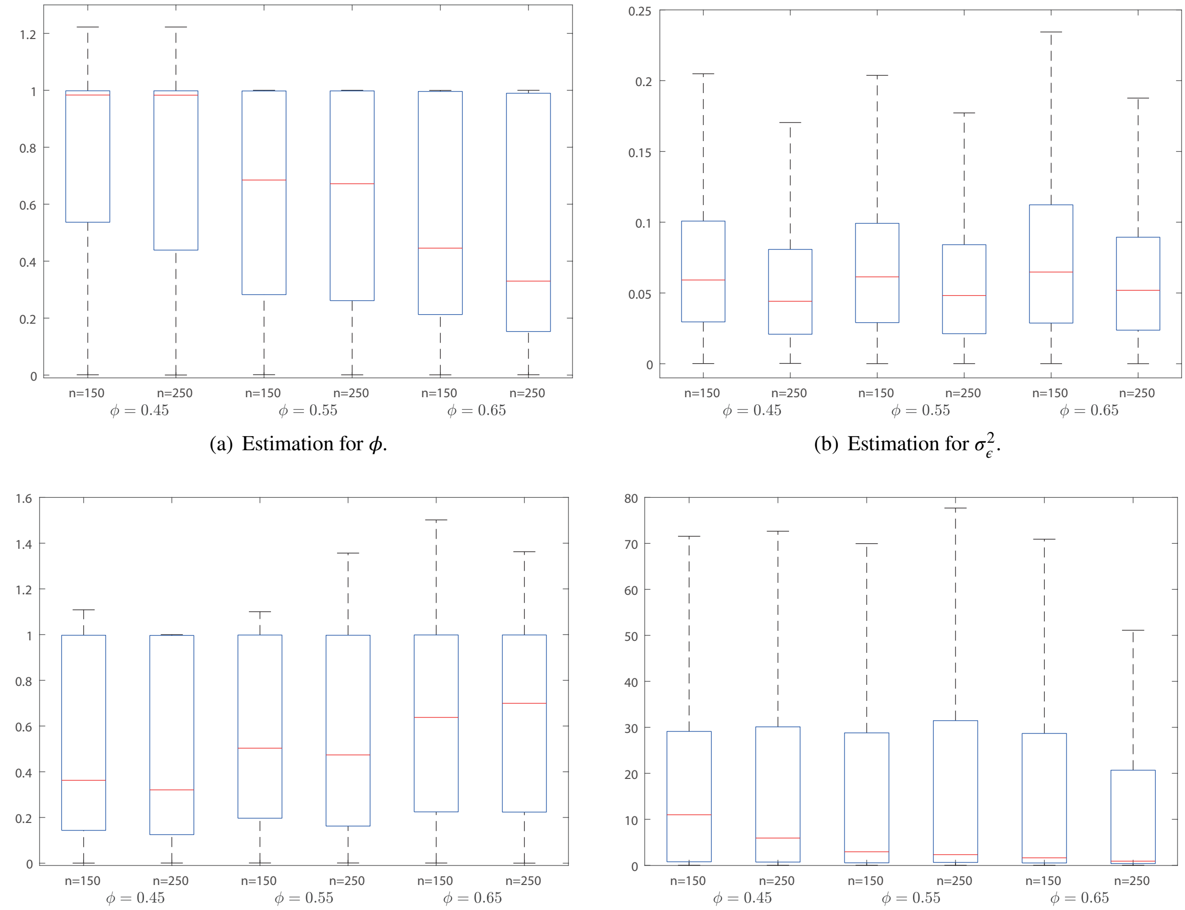
Figure 4 shows a time series chart of the credit-to-GDP ratio for two different groups of market economies. Figure 4(a) shows that the US experienced a huge credit boom in the 1990s, while for the Eurozone and the UK, it stagnated with the global financial crisis, and Japan also experienced a deceleration in credit over time. Figure 4(b) shows the major emerging market economies, which include Brazil, China, India and Russia. Similar to the advanced economies, credit operations in the emerging market economies are expanding rapidly, with momentum evident in China in particular.

The design of the forecasting evaluation work is as follows: each time series (48 in total) is split into training and testing samples, corresponding to the last six observations. The predictive performance of the two models is

[1] In this data, a four-quarter moving sum is used to compute the nominal value of GDP, implying minimal or no residual seasonality. This factor is crucial as the MSOE model utilized in this study does not cater to seasonal time-series.

[2] For more information, refer to `https://www.bis.org/statistics/about_credit_stats.htm`.

[3] The "Non-Financial Sector Credit Long Series" database, accessed on September 3, 2022, is the source of this data. The series used refers to the credit to the non-financial sector (adjusted for breaks), provided by all sectors, valued at market prices, and expressed as a percentage of GDP. It includes data from countries and aggregates: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Czech Republic, Denmark, Euro area, Finland, France, G20, Germany, Greece, Hong Kong SAR, Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Malaysia, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Russia, Saudi Arabia, Singapore, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, United States, as well as aggregate data from all reporting countries, emerging markets, and advanced economies.



(a) Estimation for ϕ.

(b) Estimation for σ_ϵ^2.

(c) Estimation for σ_η^2.

(d) Estimation for σ_ξ^2.

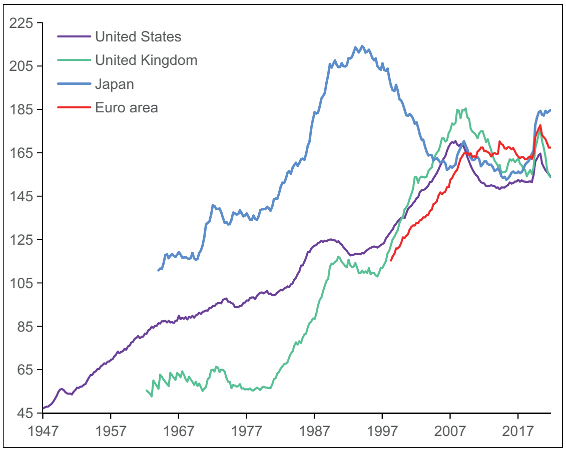
Figure 3: Monte Carlo simulations results-APE for ϕ,σ_ϵ^2,σ_η^2,σ_ξ^2. This figure reports the boxplots of the APE obtained for the estimates of ϕ,σ_ϵ^2,σ_η^2,σ_ξ^2 using RC-SSPACE.

based on 1, 2, 3, 4, 5, and 6-step ahead forecasts, evaluated by the MASE. Table 3 shows the out-of-sample forecast results for the full sample as well as the short-term sample, where Columns "Mean ratio" and "Median ratio" report Mean MCMC/Mean SSPACE and Median MCMC/Median SSPACE, respectively. The column "Ranking" reports the percentage of series for which the MCMC outperforms the SSPACE in terms of MASE.

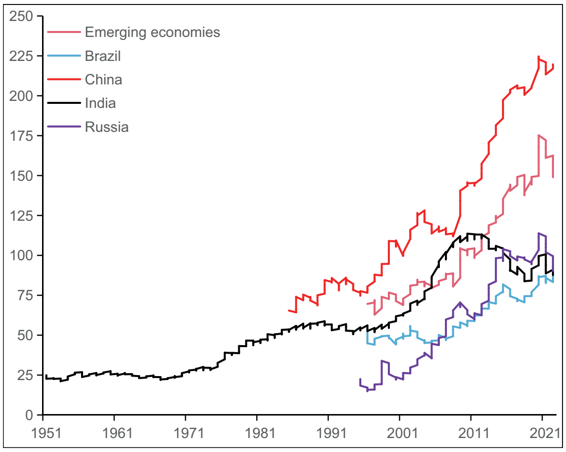
The full sample contains data for all periods of each series, while the short-term sample period spans from 2010:Q2 to 2022:Q3, totaling 50 periods, this ensures that all time series have 45 observations, facilitating a deeper understanding of how sample size influences the predictive performance of each estimation method.

Above all, estimation on the full sample, for each forecast horizon, the mean and median ratios suggest a small advantage of RC-MCMC over RC-SSPACE. This result is supported by the "Ranking" statistics, at least for forecasting horizons 2, 3, 4, and 5 (while the results for forecasting horizons 1 is balanced between the two methods). In the short-term sample, RC-MCMC definitively outperforms RC-SSPACE in terms of forecasting accuracy. This is consistently true no matter which measure ("Mean ratio", "Median ratio", "Ranking") is considered. It also suggests that this forecasting method is not only useful in the medium term but also for short-term horizons.

However, an interesting phenomenon can be observed when examining the parameter estimation results of these two methods. Except for σ_ϵ^2, the other three parameters align with the model parameter settings. Figure 5 illustrates the distribution of the 48 estimated values for σ_ϵ^2. It can be observed that the values estimated using RC-MCMC align with the variance value settings, both in the full sample and short-term sample. But the RC-SSPACE method

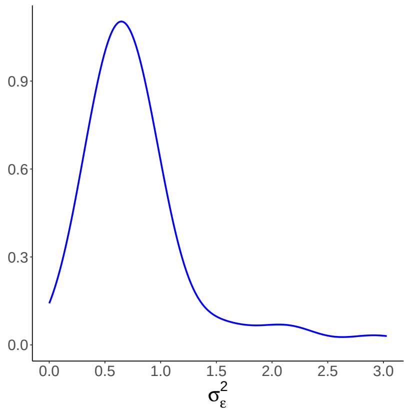


(a) US, UK, Japan and Euro Area.

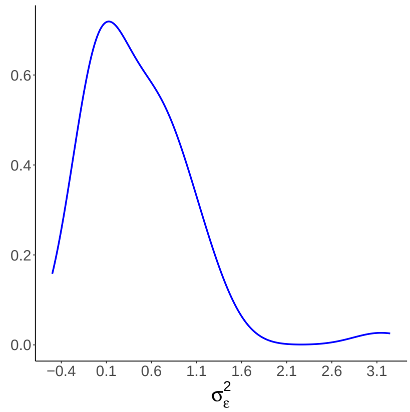


(b) Emerging market economies.

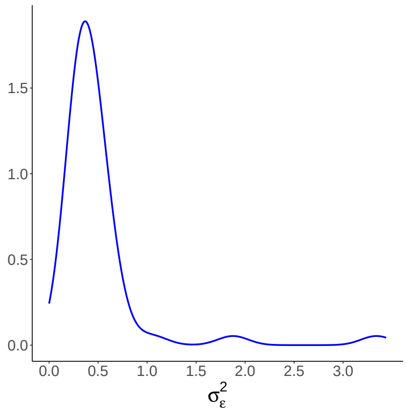
Figure 4: Credit to GDP ratios (total credit to the private non-financial sector, in per cent of GDP).



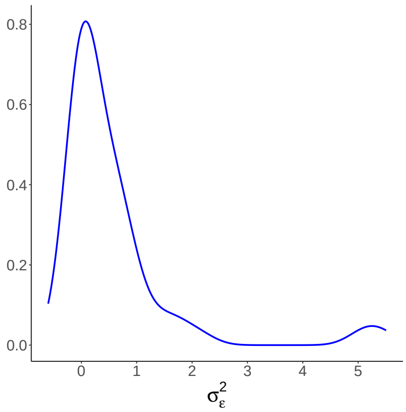
(a) MCMC (Full Sample).



(b) SSPACE (Full Sample).



(c) MCMC (Short-term Sample).



(d) SSPACE (Short-term Sample).

Figure 5: Density for σ_ϵ^2 under the four models.

Table 3
Out-of-sample forecast comparison of RC-MCMC vs. RCSPACE based on quarterly BIS credit data.

	No.ahead step	Mean Ratio	Median Ratio	Ranking
Full Sample	$h = 1$	0.983	1.092	0.5
	$h = 2$	0.899	0.916	0.625
	$h = 3$	0.890	0.837	0.708
	$h = 4$	0.868	0.714	0.708
	$h = 5$	0.936	0.999	0.604
	$h = 6$	1.002	1.066	0.479
	Mean	0.930	0.937	0.604
Short-term Sample	$h = 1$	0.935	1.011	0.617
	$h = 2$	0.891	0.887	0.681
	$h = 3$	0.900	0.820	0.787
	$h = 4$	0.870	0.695	0.723
	$h = 5$	1.002	1.070	0.574
	$h = 6$	0.944	0.973	0.574
	Mean	0.924	0.909	0.660

Notes. In the short-term sample, the data set for Ireland has been removed as the trend in the data is not clear, 47 in total for short-term sample.

estimates for σ_ϵ^2 are all close to zero, with 20 estimates of $\sigma_\epsilon^2 < 0.01$ in the full sample and 24 for the short-term sample–essentially 50% of the number of sequences. This indicates that most of the σ_ϵ^2 values estimated by the RC-SSPACE optimization method do not match the actual situation, whereas the RC-MCMC method does not encounter this problem. Overall, from the out-of-sample forecast results and the parameter estimation results in real data, RC-MCMC has certain advantages over RC-SSPACE for some extent.

6. Conclusions

We introduce a novel precision-based algorithm approach for directly estimating the parameters of the MSOE model-random coefficient state space model. The simulation study involves a comparative analysis of both forecasting performance and parameter estimation accuracy between these two methods. The findings reveal that the RC-MCMC technique outperforms the RC-SSPACE method in terms of estimation accuracy, exhibiting greater stability in the estimated results. And the empirical investigation using total credit-to-GDP data from the BIS further corroborates that the RC-MCMC method generates more reasonable parameter estimates. Furthermore, the Bayesian approach yields superior forecasting outcomes and provides a natural means for quantifying the uncertainties inherent in both the model and parameters during the forecasting process. This evidence suggests that the method is better suited for forecasting and estimating such trend data, particularly in cases involving random switching exponential smoothing with more frequent coefficient changes.

References

Carter, C.K., Kohn, R., 1994. On gibbs sampling for state space models. Biometrika 81, 541–553.

Chan, J., Grant, A.L., 2016. On the observed-data deviance information criterion for volatility modeling. Journal of Financial Econometrics 14, 772–802.

Chan, J.C., 2013. Moving average stochastic volatility models with application to inflation forecast. Journal of Econometrics 176, 162–172.

Chan, J.C., Jeliazkov, I., 2009. Efficient simulation and integrated likelihood estimation in state space models. International Journal of Mathematical Modelling and Numerical Optimisation 1, 101–120.

De Jong, P., Shephard, N., 1995. The simulation smoother for time series models. Biometrika 82, 339–350.

Durbin, J., Koopman, S.J., 2002. A simple and efficient simulation smoother for state space time series analysis. Biometrika 89, 603–616.

Gardner, E.S., 2006. Exponential smoothing: The state of the artpart ii. International Journal of Forecasting 22, 637–666.

Hyndman, R., Koehler, A.B., Ord, J.K., Snyder, R.D., 2008. Forecasting with exponential smoothing: the state space approach. Springer Science & Business Media.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with arch models. The Review of Economic Studies 65, 361–393.

Kourentzes, N., Barrow, D., Petropoulos, F., 2019. Another look at forecast selection and combination: Evidence from forecast pooling. International Journal of Production Economics 209, 226–235.

McKenzie, E., Gardner, E.S., 2010. Damped trend exponential smoothing: A modelling viewpoint. International Journal of Forecasting 26, 661–665.

Monache, D.D., Petrella, I., 2019. Efficient matrix approach for classical inference in state space models. Economics Letters 181, 22–27.

Moura, G.V., Turatti, D.E., 2014. Efficient estimation of conditionally linear and Gaussian state space models. Economics Letters 124, 494–499.

Ord, K., Fildes, R.A., Kourentzes, N., 2017. Principles of business forecasting. 2nd ed., Wessex Press Publishing Co.

Sbrana, G., Silvestrini, A., 2014. Random switching exponential smoothing and inventory forecasting. International Journal of Production Economics 156, 283–294.

Sbrana, G., Silvestrini, A., 2019. Random switching exponential smoothing: A new estimation approach. International Journal of Production Economics 211, 211–220.

Sbrana, G., Silvestrini, A., 2020. Forecasting with the damped trend model using the structural approach. International Journal of Production Economics 226, 107654.