TGU-ECON Discussion Paper Series
#2016-1

Large Shocks and the Business Cycle: The Effect of Outlier Adjustments

Yoshihiro Ohtsuka
Faculty of Economics, Tohoku Gakuin University

April 2016

Large shocks and the Business cycle: The effect of outlier adjustments*

Yoshihiro Ohtsuka†

Abstract

This study examines the impact of outlier-adjusted data on business cycle inferences using coincident indicators of the composite index (CI) in Japan. To estimate the CI and business cycles, this study proposes a Markov switching dynamic factor model incorporating Student's t-distribution in both the idiosyncratic noise and the factor equation. Furthermore, the model includes a stochastic volatility process to identify whether a large shock is associated with a business cycle. From the empirical analysis using unadjusted data, both the factor and the idiosyncratic component have fat-tail error distributions, and the estimated CI and recession probabilities are close to those published by the Economic and Social Research Institute. Compared with the estimated CI using the adjusted data set, the outlier adjustment reduces the depth of the recession. Moreover, the results of the shock decomposition show that the financial crisis in mid-2008 was caused by unexpected shocks. In contrast, the Great East Japan Earthquake in 2011 was derived from idiosyncratic noise and did not cause a recession. When analyzing whether to use a sample that includes outliers associated with the business cycle, it is desirable to use the outlier-adjusted data set.

Key Words: Business cycle inference; Heavy-tailed distribution; Markov chain Monte Carlo (MCMC); Markov switching dynamic factor model; Stochastic volatility.

JEL Classification: C11; C31; C32; R12.

*This research is supported by a grant-in-aid from Zengin Foundation for Studies on Economics and Finance.
†Faculty of Economics, Tohoku Gakuin University, 1-3-1 Tsuchitoi, Aoba-ku, Sendai, Miyagi, 980-8511, Japan ,E-mail: ohtsuka@mail.tohoku-gakuin.ac.jp

1 Introduction

Assessing business fluctuations and cycle phases has attracted much attention among macroeconomists and econometricians, because the magnitudes of economic variables are important to the manufacturing and service industries, central banks, and government. In particular, the severe recession in mid-2008 affected all industrialized countries. The impact of this recession on the business cycles in these countries was so large that it is difficult to measure the business fluctuations using conventional methods. Thus, research on business cycles has intensified, and is required in order to improve business index measures and to extend econometric methodologies.

A composite index (CI) is widely used to measure trends and quantitative business fluctuations in countries such as the United States, United Kingdom, Japan, and so on. This study analyzes Japanese macroeconomic data series and business cycles. The Japanese economy has experienced severe crises in recent decades, for example, the financial crisis in mid-2008 and the major earthquake in March 2011. Although the Japanese economy was damaged by these crises, the shocks themselves vary in their characteristics. Business dropped off in mid-2008, because after the financial crisis, world trade began to shrink rapidly. This shock is one associated with business cycles, particularly recessions. In contrast, the Great East Japan Earthquake is an example of a natural disaster that affects isolated economic activity.

The CI in Japan is an index that replaces outliers based on descriptive statistics. The adjustment identifies outliers as observations in component series that lie outside the normal range of expected observations, and replaces them with estimated values in order to measure robust business cycles. Since this adjustment automatically replaces extreme values, it may be unable to distinguish fully between shocks derived from isolated observations and those from recessions. Therefore, one implication is that using data with outlier replacement discards information on the business cycle. In other words, if the outlier is associated with a recession, then the adjustment reduces the depth of the recession. Surprisingly, despite some studies investigating the link between seasonal adjustments and business cycles, such as Franses and Paap (1999) and Matas-Mir *et al.*(2008), the impact of eliminating outliers has not been well discussed, with the exception of the study of Balke and Fomby (1994). This is an important issue, because the data adjustment may mislead an empirical analysis, resulting in erroneous conclusions. Thus, the first contribution of this study is to examine the impact of outlier adjustments on business cycle inferences.

Large shocks in macro economic data series have been discussed since the 1990s. For example, Christiano and Den Haan (1996) pointed out that macro economic variables exhibit non-Gaussian behavior, which is called the fat-tailed problem. Recently, it has become necessary to consider excess kurtosis when analyzing macro economic time series data, such as industrial production and business cycles (*e.g.*, Fagiolo *et al.*, 2008; Watanabe, 2014; Ascari *et al.*, 2015), and when applying macro economic models such as the dynamic stochastic general equilibrium model (Cúrdia *et al.*, 2014). Therefore, we need to consider the properties of the data distributions in the econometric models.

For business cycle inferences, we employ the Markov switching dynamic factor (MSDF) model, as used by Kim and Yoo (1995), Chauvet (1998), and Kim and Nelson (1998). The MSDF model is widely used to simultaneously track business fluctuations as co-movements among individual economic indicators, and phases such as a recession regime or expansion. Using the MSDF model, numerous papers analyze the properties of business cycles, and these specifications are applied to empirical analyses in macro econometrics (*e.g.*, Kaufman, 2000; Camacho *et al.*, 2014). The MSDF model is extended by incorporating Student's t-distribution and a stochastic volatility (SV) process. The investigation is related to that of Watanabe (2014), who presented evidence supporting the use of heavy-tail error distributions, based on the CI in Japan. He also showed a model with good fit that used an SV and a fat-tail distribution. We assume that the error distributions of both the idiosyncratic equation and the factor equation follow Student's t-distribution. This formation allows us to compute the outliers for the latent factor noise and the idiosyncratic noise. Moreover, including a t-distribution for the error distribution and an SV process in the factor equation makes it possible to decompose shocks to business cycles into unexpected shocks and conditional expected shocks. The second contribution is to extend the MSDF model to include large shocks. Once the SV process and non-Gaussian distribution are introduced, it becomes difficult to evaluate the likelihood. We employ the Markov chain Monte Carlo method in the Bayesian inference to estimate the model.

In our empirical analysis, we apply our MSDF model to 11 coincident economic indicators from February 1985 to December 2014, as in the Economic and Social Research Institute (ESRI) data set. From the results of the estimated parameters, the posterior mean of the degree of freedom in the factor equation is close to 7, denoting that the latent factor follows a heavy-tailed distribution. Compared with the CI of the ESRI, we show that the use of outlier adjusted data leads to poorer performance in estimating business fluctuations and in detecting cyclical characteristics. The lack of information on outliers causes an underestimation of business fluctuations. Moreover, the results of the decomposition of shocks shows that the impact of a financial crisis on the business cycle is large, and is derived from unexpected shocks. On the other hand, the outlier of the Great East Japan Earthquake in 2011 is associated with an idiosyncratic shock, and did not cause a recession. Our model is able to distinguish between shocks associated with a recession and the disaster shocks. The outlier adjustment misleads the inference if the sample includes various types of outliers.

The rest of paper is organized as follows. Section 2 explains the CI in Japan and the outlier adjustment by the ESRI, and shows the effect of the adjustment using descriptive statistics. Section 3 extends the MSDF model by incorporating a heavy-tailed error distribution and an SV process. Section 4 explains the Bayesian method used to analyze this model. Section 5 fits the model to macroeconomic data in Japan and summarizes the results. Finally, Section 6 concludes the paper.

2 Data issues and outlier adjustment

2.1 Composite index and coincident indicators

In Japan, the composite index of coincident indicators (CI) is well known as a quantitative measure of the current business conditions or state of aggregate economic activity. The CI is designed by the ESRI, Cabinet Office, Government of Japan. Recently, two types of CI were announced. The first is an index with outlier replacement. The second is an index without outlier replacement. The ESRI adjust data so that short-term shocks do not overly affect the index. In the past decade, the Japanese economy has experienced large shocks, such as the financial crisis in mid-2008 and the Great East Japan Earthquake in 2011. Owing to these crises, the Japanese economy has struggled. Subsequently, a difference has emerged from the two indices.

Figure 1 plots the CI with and without outlier replacement in the top of the figure, and the growth rates in the bottom of the figure for the period January 1985 to September 2014. The growth rates are computed as the log-difference and multiplied by 100, and include shaded areas to indicate periods of national recession, as determined by the ESRI. The figure shows that the value of the CI with outlier replacement is larger than that without outlier replacement after the financial crisis. These figures show that the two crises had a significant impact on the business cycles. However, we believe that these two events vary in terms of their characteristics. The Japanese economy entered a recession, when world trade began to shrink after the financial crisis. In contrast, the damage to nuclear power stations by the Great East Japan Earthquake reduced electricity supply capacity and caused several indicators of industrial production to decrease sharply. Subsequently, the ESRI did not classify this period as an economic recession, although the decline in the growth rate in 2011 was larger than after the financial crisis. It is natural that the shock from the earthquake not be associated with a recession. Moreover, after the earthquake, the values of the two CIs differ markedly. The differences in the data suggest that the CI with outlier replacement underestimates the business cycle. We need to examine how the adjusted outliers influenced the business condition measure.

Next, we shift the emphasis to the component indicators and the procedure used to exclude outliers. To construct this index, the ESRI selects 11 macroeconomics indicators, related to industrial production, trade sales, and employment[1]. Table 1 summarizes the key macroeconomics indicators and lists the abbreviation for each variable. We use monthly data[2], which is seasonally adjusted by the X-12-ARIMA procedure, from January 1985 to September 2014. Figures 2 and 3 show time series plots of the coincident indicators. Here, the growth rates of the indicators other than the retail sales value (RSV), the wholesale sales value (WSV), and the effective job offer rate (EJOR) are calculated as the log-difference multiplied by 100. The RSV, WSV, and EJOR are calculated as the growth rate from the previous month because these indicators are already published in percentage form, or are zero or negative values. From the figure, we find that almost all indicators fell substantially in the period of the financial crisis. On the other hand, in the period of the Great East Japan Earthquake, although indicators related to the manufacturing industry fell suddenly, those related to sales and operating profits and the job market did not drop by as much.

2.2 Outlier adjustment

Here, we explain the approach used to remove outliers. According to Tsay (1988) and Balke and Fomby (1994), outliers are defined as extraordinary and infrequently occurring shocks that have large and significant effects on time series. These authors show how to identify outliers and their search algorithm using the ARMA model approach. In this study, we illustrate the outlier adjustment utilized by the ESRI. This adjustment is based on descriptive statistics and is also used by the OECD.

Let y_{it} and y_{it}^* denote the growth rate of each indicator and the adjusted growth rate, respectively. Outliers in the growth rate of each indicator are replaced as follows:

$$y_{it}^* = \begin{cases} -\kappa(Q3_i - Q1_i), & \text{if} \quad y_{it} < -\kappa(Q3_i - Q1_i), \\ y_{it}, & \text{if} \quad -\kappa(Q3_i - Q1_i) < y_{it} < \kappa(Q3_i - Q1_i), \\ \kappa(Q3_i - Q1_i), & \text{if} \quad \kappa(Q3_i - Q1_i) < y_{it}, \end{cases} \tag{1}$$

where $Q3_i(Q1_i)$ is the third (first) quartile in the interquartile range of the growth rates of the i th indicator. In this study, the interquartile range is calculated using the sample from January 1985 to December 2010, and κ is set as 2.02 as in the ESRI approach. Figure 4 plots the growth rate of the coincident indicators with outlier replacement. It seems that data series are stationary data after eliminating outliers.

This adjustment defines outliers as observations in component series that lie outside the interquartile range of each growth rate, and are automatically replaced. If all detected outliers are associated with irregular shocks such as war, strikes, or disasters, then this procedure is effective. However, if an outlier is associated with a business cycle, as in the financial crisis in mid-2008, the use of adjusted data may distort the determination of recession. This procedure cannot distinguish between the shocks. Furthermore, this approach is an ad hoc way to exclude

[1]The ESRI does not include data on personal income, which is a major coincident economic indicator in the United States.

[2]Coincident indicators are available at `http://www.esri.cao.go.jp/en/stat/di/di-e.html`.

outliers. We attempt to track the business cycle from the coincident economic indicators in order to examine the influence on estimated business fluctuations and cycles of a lack of information on the outliers. Thus, we use the coincident economic indicators instead of the CI of the ESRI in our empirical analysis.

2.3 Descriptive statistics of coincident indicators

In this subsection, we analyze the statistical properties of the growth rate distributions. Table 2 summarizes the descriptive statistics of the growth rates of the unadjusted data set. The number of time series is 356 for each indicator. The mean of each indicator is not significantly different from 0. However, the standard deviations are high because of the financial crisis and the Great East Japan Earthquake. The skewness of each indicator is negative, and significantly different from 0. The kurtosis of each indicator is significantly over 3, which is larger than that of a normal distribution. The JB (Jarque–Bera) statistics of all the indicators reject the null hypothesis of normality, indicating that the normal distribution does not perform well in describing the growth rate distribution of the indicators, as shown by Christiano and Den Haan (1996). This implies that the growth rates of the coincident indicators have non-normal distributions, or leptokurtic distributions with tails that are much fatter than those of normal distributions. From the values of LB(10)[3] in the table, we confirm that the null hypothesis for the growth rates, other than IIPP, are rejected at the 5% significance level. Furthermore, the results of $\mathrm{LB}_2(10)$, denoting the null hypothesis for the squared growth rates, imply that the volatility in the squared growth rates of all the indicators include autocorrelation.

For reference purposes, Table 3 summarizes the descriptive statistics of the growth rates with outlier replacement. The kurtosis of each indicator is significantly less than 3. According to the value of the JB test, the null hypothesis of normality is not rejected. Other than these points, the results are not significantly different. In summary, the growth rates of non adjusted coincident economic indicators have the following features: negative skewness, excess kurtosis, autocorrelation in the growth rate, and autocorrelation in the squared growth rate. Furthermore, outlier adjustment reduces their kurtosis.

In this study, we track the business fluctuations and cycles from several coincident economic indicators using an econometric model that considers these properties of the sample. In previous analyses on business cycles in Japan, Fukuda and Onodera (2001) utilized the dynamic factor model to estimate the CI. Subsequently, Watanabe (2003) and Hayashida and Hewings (2009) used a Markov switching dynamic factor model proposed by Kim and Nelson (1998). The Markov switching dynamic factor model is widely used to estimate business fluctuations as co-movements among individual economic indicators and cyclical phases (recession or boom). Using the MSDF model, the autocorrelation and negative skewness in the growth rates of coincident economic indicators may partly be explained by the regime switch process of the business cycle in the Markov switching model. Moreover, Watanabe (2014) also showed a model with good fit using an SV and a fat-tail distribution, based on the CI of the ESRI, and did not detect turning points of business cycles with the model based on a normal distribution. From the descriptive statistics, the autocorrelation of the squared growth rate implies that volatility follows an autoregressive process. Therefore, we incorporate the Markov switching factor model with a heavy-tailed error distribution and an SV process for the business cycle inference. In order to divide the outliers into shocks associated with business cycles, especially recessions and shocks that affect isolated business cycles, we assume the error distributions of the idiosyncratic noise and the factor equation both follow Student's t-distribution.

[3] These statistics denote the Ljung-Box statistics adjusted by Diebold (1988) to test the null hypothesis of no autocorrelations up to 10.

3 Econometric model

3.1 Markov switching dynamic factor model

First, we introduce the simplest MSDF model[4] as in Diebold (2003). This study tracks only the business trends from several coincident indicators[5]. We thus fit a single factor model with regime switching. Suppose we have data on n macroeconomic variables. Let y_{it} for $i = 1, \dots, n$ and $t = 1, \dots, T$ denote the growth rate of the i th macroeconomic variable at time t. Then, y_{it} is determined by changing the latent common factor c_t and the idiosyncratic component e_{it}:

$$\begin{aligned} y_{it} &= \gamma_i c_t + e_{it}, && (2) \\ e_{it} &= \psi_i e_{i,t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim \mathcal{N}(0, \sigma_i^2), && (3) \end{aligned}$$

where γ_i denotes a factor loading term, and ψ_i is the autoregressive time dependence parameter. We assume ϵ_{it} is *i.i.d.* across i and t with zero mean and variance σ_i^2. We define a latent variable, c_t $(t = 1, \dots, T)$, assumed as the CI, as the following autoregressive process:

$$\begin{aligned} c_t &= \mu_t + \phi(c_{t-1} - \mu_{t-1}) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, 1),, && (4) \\ \mu_t &= \mu^{(0)}(1 - s_t) + \mu^{(1)} s_t, \quad \mu^{(0)} < \mu^{(1)}, && (5) \\ s_t &= \begin{cases} 1 & \text{boom} \\ 0 & \text{recession} \end{cases}, \end{aligned}$$

where μ_t shifts depending on the state variable s_t, ϕ is an autoregressive time dependence parameter, and η_t is *i.i.d.* across t with zero mean and variance, which is normalized to unity for the identification of the model, and is uncorrelated with ϵ_{it} for $i = 1, \dots, n$ and $t = 1, \dots, T$. We assume state variable s_t takes the value 0 when the economic condition is a recession period, and 1 in a boom. The inequality $\mu^{(0)} < \mu^{(1)}$ is assumed in equation (4) because the growth of business conditions will be greater in a boom regime than in a recession regime. Moreover, based on the conventional MS model, the dynamics of s_t are first-order Markov processes, with the following transition probability matrix:

$$\mathbf{\Pi} = \begin{bmatrix} p_0 & 1 - p_1 \\ 1 - p_0 & p_1 \end{bmatrix}.$$

3.2 Heavy-tailed error distributions and stochastic volatility

In this study, we extend the MSDF model with a heavy-tailed error distribution and SV process. We assume the error terms e_{it} and ϵ_{it} follow Student's t-distribution. This distribution has higher kurtosis than that of the Gaussian distribution, and is widely used in financial econometrics for considering the tail risk. Furthermore, models including an SV process approach have been applied in macro economic data analyses (*e.g.*, Primiceri, 2005; Nakajima *et al.*, 2011; Cúrdia *et al.*, 2014). They showed that it is necessary to incorporate an SV process for macro economic time series data. However, it is difficult to calculate the latent variable for non-Gaussian models using a conventional Kalman filter, because the likelihood with heavy-tailed errors cannot be evaluated analytically. Thus, we add auxiliary variables and utilize the data augmentation approach in a Bayesian inference. Then, the error term in equations (3) and (4) is rewritten, and the SV process is as follows:

$$\begin{aligned} \epsilon_{it} &= \sqrt{\lambda_{it}} u_{it}, \quad u_{it} \sim \mathcal{N}(0, \sigma_i^2), && (6) \\ \eta_t &= \exp\left(\frac{h_t}{2}\right) \sqrt{\omega_t} z_t, \quad z_t \sim \mathcal{N}(0, 1), && (7) \\ h_t &= \beta h_{t-1} + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, \xi^2), && (8) \end{aligned}$$

[4]Strictly speaking, this model is a static factor model.

[5]There has been some recent research using large data sets (Stock and Watson, 2014).

where λ_{it} and ω_t follow the hierarchical prior distribution:

$$\lambda_{it} \sim \mathcal{IG}\left(\frac{\nu_i}{2}, \frac{\nu_i}{2}\right), \quad \frac{\nu_f - 2}{\omega_t} \sim \chi^2(\nu_f),$$

where $\mathcal{IG}$, ν_i, and ν_f denote the inverse gamma distribution and unknown parameters of the degree of freedom, respectively. Then, η_t follows the standardized Student's t-distribution such that the variance is one. We assume that $\nu_i > 2$ and $\nu_f > 2$ in order to satisfy a finite variance. In equation (8), h_t denotes the log volatility at time t, and the parameter β captures the autocorrelation in the volatility. In our empirical analysis, we estimate the MSDF model with both a t-error and SV, and label it as the MSDF-SVt model.

Finally, let $y = \{y_i\}_{i=1}^n$, $y_i = \{y_{it}\}_{t=1}^T$, $\gamma = \{\gamma_i\}_{i=1}^n$, $\psi = \{\psi_i\}_{i=1}^n$, $\sigma^2 = \{\sigma_i^2\}_{i=1}^n$, $\nu = \{\nu_i\}_{i=1}^n$, $\mu = (\mu^{(0)}, \mu^{(1)})'$, and $p = (p_0, p_1)$. To simplify the notation, let $\theta = (\theta_1, \theta_2)$, $\theta_1 = (\gamma, \sigma^2, \psi, \nu, \mu, p, \phi, \nu_f)$, $\theta_2 = (\beta, \xi^2)$, $c = \{c_t\}_{t=1}^T$, $\lambda = \{\lambda_i\}_{i=1}^n$, $\lambda_i = \{\lambda_{it}\}_{t=1}^T$, $\omega = \{\omega_t\}_{t=1}^T$, $s = \{s_t\}_{t=1}^T$, $h = \{h_t\}_{t=1}^T$, and $\vartheta = (c, \lambda, \omega, s, h)$. Given θ and ϑ, the conditional likelihood is as follows:

$$L(y|\theta, \vartheta) = \prod_{i=1}^n \prod_{t=1}^T f(y_{it}|\theta, \vartheta), \tag{9}$$

where

$$f(y_{it}|\theta, \vartheta) = \begin{cases} \sqrt{\frac{1-\psi_i^2}{2\pi\lambda_{it}\sigma_i^2}} \exp\left\{-\frac{(1-\psi_i^2)z_{it}^2}{2\lambda_{it}\sigma_i^2}\right\}, & \text{if } t = 1 \\ \frac{1}{\sqrt{2\pi\lambda_{it}\sigma_i^2}} \exp\left\{-\frac{(z_{it}-\psi_i z_{i,t-1})^2}{2\lambda_{it}\sigma_i^2}\right\}, & \text{if } t > 1 \end{cases}$$

with $z_{it} = y_{it} - \gamma_i c_t$.

4 Posterior analysis

4.1 Joint posterior distribution

Since we adopt a Bayesian approach, we complete the model by specifying the prior distribution over the parameters. Thus, we apply the following prior distribution:

$$\pi(\theta) \quad = \quad \left\{\prod_{i=1}^n \pi(\gamma_i)\pi(\psi_i)\pi(\sigma_i^2)\pi(\nu_i)\right\} \pi(\mu)\pi(\phi)\pi(p)\pi(\nu_f)\pi(\beta)\pi(\xi^2), \tag{10}$$

Given the prior density $\pi(\theta)$ and the likelihood function given in equation (9), the joint posterior distribution can be expressed as

$$\pi(\theta, \vartheta|y) \quad \propto \quad L(y|\theta, \vartheta)\pi(\vartheta|\theta)\pi(\theta). \tag{11}$$

In this study, we assume the following proper prior distributions:

$$\mu \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_{\mu 0})I[\mu^{(0)} < \mu^{(1)}],\ \gamma_i \sim \mathcal{N}(\gamma_0, \sigma_{\gamma 0,i}^2),\ \frac{\psi_i + 1}{2} \sim Beta(a_{\psi,i}, b_{\psi,i}),$$
$$\frac{\phi + 1}{2} \sim Beta(a_\phi, b_\phi),\ p_0 \sim Beta(\iota_{00}, \iota_{01}),\ p_1 \sim Beta(\iota_{11}, \iota_{10}),$$
$$\sigma_i^2 \sim \mathcal{IG}\left(\frac{\tau_{i0}}{2}, \frac{\delta_{i0}}{2}\right),\ \nu_i \sim \mathcal{G}(A_0, B_0)I[\nu_i > 2],\ \nu_f \sim \mathcal{G}(A_0^*, B_0^*)I[\nu_f > 2],$$
$$\frac{\beta + 1}{2} \sim Beta(a_\beta, b_\beta),\ \xi^2 \sim \mathcal{IG}\left(\frac{\tau_h}{2}, \frac{\delta_h}{2}\right),$$

where $\mathcal{G}$ denotes gamma distribution. Then, $I(\cdot)$ is the indicator function that takes one if the condition in parentheses is satisfied, and zero otherwise. The prior distribution of μ is the truncated normal distribution in order to satisfy the condition, $\mu^{(0)} < \mu^{(1)}$. For prior distributions of ψ_i and ϕ, we assume that $\frac{\psi_i+1}{2}$ and $\frac{\phi+1}{2}$ follow the beta distribution, because ψ_i and ϕ satisfy the stationary condition. For the prior distributions of ν_i and ν_f, we assume that truncated gamma distributions (Watanabe, 2001).

4.2 MCMC estimation

This subsection introduces the algorithm to estimate the parameters using the MCMC method. We need to use multiple iterations to evaluate the marginal posterior distribution in the joint posterior distribution (11). It is difficult to solve the marginal posterior distribution if the joint posterior distributions are complicated. Then, we sample the parameters from the full conditional distribution of the parameter using the MCMC method, which is an algorithm that utilizes Markov sampling and Monte Carlo integration to approximate the full conditional distribution. Thus, we draw the random samples from the posterior distributions for the MSDF-SVt model using the MCMC methods, as follows:

1. Initialize θ and ϑ.
2. Draw $\gamma_i \mid \theta_{-\gamma_i}, \vartheta, y$ for $i = 1, \dots, n$.
3. Draw $\psi_i \mid \theta_{-\psi_i}, \vartheta, y$ for $i = 1, \dots, n$.
4. Draw $\sigma_i^2 \mid \theta_{-\sigma_i^2}, \vartheta, y$ for $i = 1, \dots, n$.
5. Draw $\mu \mid \theta_{-\mu}, \vartheta$
6. Draw $\phi \mid \theta_{-\phi}, \vartheta$
7. Draw $c \mid \theta, \vartheta_{-c}, y$
8. Draw $s \mid \theta, \vartheta_{-s}, y$
9. Draw $p \mid \theta_{-p}, \vartheta, y$
10. Draw $\lambda_i \mid \theta, \vartheta_{-\lambda_i}$ for $i = 1, \dots, n$.
11. Draw $\nu_i \mid \theta_{-\nu_i}, \vartheta$ for $i = 1, \dots, n$.
12. Draw $\omega| \theta, \vartheta_{-\omega}$
13. Draw $\nu_f| \theta_{-\nu_f}, \vartheta$
14. Draw $\beta| \theta_{-\beta}, \vartheta$
15. Draw $\xi^2| \theta_{-\xi^2}, \vartheta$
16. Draw $h| \theta, \vartheta_{-h}, y$
17. Go to 2.

Here, $\theta_{-\mu}$ denotes the set of all parameters, excluding μ. We can implement the sampling scheme of Watanabe (2014) for steps 5, 6, 8, and 9. For step 7, we sample the latent variables using an efficient simulation smoother proposed by Durbin and Koopman (2002). We also utilize the acceptance rejection Metropolis Hastings (AR-MH) of Watanabe (2001) for steps 8, 9, 10, and 11. Following Watanabe and Omori (2004), we employ a multi-move sampler in step 16. The procedures are described in detail in the Appendix.

5 Empirical analysis

5.1 Available data and estimation procedure

We apply the MSDF-SVt model to monthly coincident indicators for the period January 1985 to December 2014 obtained from the database of the ESRI, because we limit the CI's selection of time-varying variables to those for which data are available over the entire study period. In related literature on business trends using Japanese data sets, Fukuda and Onodera (2001) estimated the CI using two data sets: (i) index of industrial production (IIPP), large industrial power consumption (LIPC), index of non-scheduled worked hours (NWH), RSV, effective job offer rate (EJOR); and (ii) IIPP, index of producers' shipments (IIPS), index of operating rate[6], NWH, index of sales in small and medium sized enterprises (SME). However, they applied the dynamic factor framework to the data sets. Watanabe (2003) examined data sets with poduction-related variables. Hayashida and Hewings (2009) used IIPP, LIPC, RSV, and EJOR to estimate the regional business cycles with the MSDF framework. Our analysis uses the 11 variables in Table 1 to estimate the model, as in the ESRI approach, since the computational cost of estimating the model is not high. We transform the data into growth rates in the same way as in Section 2. Moreover, we use the growth rate with outlier replacement in (1). We estimate the MSDF-SVt model using the both data sets.

For the prior distributions, we set the hyper-parameters as follows:

$$\mu \sim \mathcal{N}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}\right) I[\mu^{(0)} < \mu^{(1)}],\ \gamma_i \sim \mathcal{N}(0, 10),$$
$$\frac{\psi_i + 1}{2} \sim Beta(1,1),\ \frac{\phi + 1}{2} \sim Beta(1,1),\ p_0 \sim Beta(9,9),\ p_1 \sim Beta(9,9),$$
$$\sigma_i^2 \sim \mathcal{IG}\left(\frac{6}{2}, \frac{4}{2}\right),\ \nu_i \sim \mathcal{G}(1.2, 0.03) I[\nu_i > 2],\ \nu_f \sim \mathcal{G}(1.2, 0.03) I[\nu_f > 2],$$
$$\frac{\beta + 1}{2} \sim Beta(2,1),\ \xi^2 \sim \mathcal{IG}\left(\frac{6}{2}, \frac{4}{2}\right),$$

The beta prior distributions for $(\psi_i + 1)/2$ and $(\beta + 1)/2$ imply that the mean and standard deviation are $(0.5, 0.29)$ and $(0.67, 0.06)$, respectively. The mean and standard deviations of the gamma and inverse gamma priors are $(40, 36.51)$ and $(1, 1)$, respectively. We perform the MCMC procedure by generating $30,000$ draws in a single sample path and discard the first $15,000$ draws as the initial burn-in. All the results in this study are calculated using Ox version 6.2 (Doornik, 2006).

5.2 Parameter estimates

The MSDF-SVt models are estimated for both the unadjusted and outlier-adjusted data. First, we report the results of the parameter estimation for the MSDF-SVt model using the unadjusted data. Tables 4 and 5 summarize the estimates for the MSDF-SVt model, where Mean, Stdev, 2.5% (97.5%) CI, CD, and IF represent the posterior mean, the standard deviation, 2.5% (97.5%) credible intervals, Geweke's convergence diagnostics[7], and inefficiency factor[8], respectively.

[6]The index of producers' shipment of durable consumer goods is used instead of the index of operating rates from November 2011.

[7]CD represents the p-value based on the test statistic on the difference between two sample means (i.e., dividing all the generated random draws into three parts, we compute two sample means from the first 10% and last 50% of the random draws), where the test statistics are asymptotically distributed as standard normal random variables. We confirm that the random draws generated by MCMC do not converge to the random draws generated from the target distribution when CE is less than 0.01 (see Geweke, 1992 for a detailed discussion of CD).

[8]The inefficiency factor, which is an index that measures how well the chain mixes, as proposed by Chib (2001), is defined as

$$\mathrm{IF} = 1 + 2\sum_{l=1}^{\infty} \hat{\rho}_l,$$

where $\hat{\rho}_l$ denotes the sample autocorrelation at lag l. It is the ratio of the numerical variance of the sample posterior mean to the variance of the sample mean from the hypothetical uncorrelated draws.

We focus on the estimated results of the factor equation in equations (4), (7), and (8) in Table 4. The posterior means of $\mu^{(0)}$ and $\mu^{(1)}$ are -0.962 and 0.322, respectively. The 95% credible intervals of those parameters do not include zero, providing evidence of the division of the business cycle into two separate phases. The posterior mean of ϕ is -0.256 and its 95% credible interval includes 0. This implies that the autocorrelation in c_t can be explained by both the switch in its mean and the past value, c_{t-1}. The result that the posterior mean of p_0 is less than that of p_1 implies the average duration of a recession is shorter than that of a boom. The posterior mean of ν_f is 7.111, which implies that the estimated business conditions followed heavy-tailed distribution, which have higher kurtosis than that of a normal distribution. The posterior mean of β is 0.742, indicating relatively high volatility clustering that shocks to volatility are persistent.

Table 5 summarizes the estimated parameters of the idiosyncratic component. The 95% credible intervals of ψ other than the LIPC and NSWH do not include zero. These idiosyncratic components are explained by the past $e_{i,t-1}$. In particular, the posterior mean of ψ in operating profits (OP) and EJOR is over 0.7, denoting that these indicators have high positive autocorrelation. The posterior means of ν other than NWH are less than 10. Since the degrees of freedom in the IIPP, IIPS, and OP are close to 2, we confirm that these idiosyncratic components follow a leptokurtic distribution.

Tables 6 and 7 summarize the estimated results using the adjusted data set. Because of the elimination of outliers, the posterior mean of ν_f is 31.418, indicating that the factor equation follows a normal distribution. In the same way, most degrees of freedom in the idiosyncratic equation increase by more than those using the unadjusted data set. Consequently, the influence of the outlier adjustment is revealed in the estimated degrees of freedom.

5.3 Estimated composite index and cycle turning points

Figure 5 depicts the posterior mean and 95% credible interval of the growth rate of the CI, estimated using the MSDF-SVt model and the unadjusted data set jointly with the growth rate of the CI without outlier replacement announced by the ESRI. Other than the period of the Great East Japan Earthquake in 2011, the 95% credible interval include the CI of the ESRI.

Figure 6 depicts the posterior mean based on unadjusted data and adjusted data, and the CI of the ESRI. There almost is no difference between the CI estimated by both data sets. However, the estimated CI using the adjusted data underestimates the sharp decline after the financial crisis when compared with the unadjusted data. With regard to estimating business trends, the influence of the lack of information on outliers is evident.

Figure 7 depicts the estimated recession probabilities using both data sets. For reference purposes, the figures include shaded areas to indicate periods of a national recession regime from peak to trough, as determined by the ESRI. The posterior probability of a recession $1 - s_t^*$ is calculated using the posterior mean of the state variable, s_t^* for $t = 1, \ldots, T$. The figure shows that the recession probabilities based on unadjusted data are in close agreement with the ESRI reference cycle after 2000, although the MSDF-SVt model does not estimate those appropriately during the period from the 1980s through the 1990s. In this paper, we estimate the business cycle turning points as follows. We define t as a peak when $1 - s_{t-1}^* < 0.5$ and $1 - s_t^* > 0.5$, and as a trough when $1 - s_{t-1}^* > 0.5$ and $1 - s_t^* < 0.5$. Table 8 shows the results. After 2000, the differences between the data on turning points of the ESRI and those using the unadjusted data set become less than three months. Compared with the results of the adjusted data, it seems that the estimated turning points based on the unadjusted data are more favorable. In summary, in terms of estimating business conditions and cycle turning points, since the adjustment reduces the depth of a recession, using data without outlier replacement is preferable to using adjusted data.

5.4 Shock decomposition

Figure 8 depicts the time series of estimated ω_t (top) and $\exp\left(\frac{h_t}{2}\right)$ (bottom) for $t = 1, \ldots, T$. Our model is able to estimate the shocks to business conditions, and decomposes those into unexpected shocks and conditional expected shocks. During the period of the financial crisis, the posterior means of *i.i.d.* shocks are larger than the others. In contrast, we confirm that the clustering shocks tend to be low in a recession. The Great East Japan Earthquake in 2011 was not a shock derived from an economic recession and did not cause a recession. Therefore, it is shown that the impact of the financial crisis in mid-2008 was substantial and caused an economic crisis associated with a depression. Moreover, these results imply that economic recessions are caused by unexpected shocks, and that it is difficult to predict recessions, as Hamilton (2011) concluded.

Figure 9 depicts the time series of estimated λ_{it}. The shocks of IIPP, IIPS, LIPC, and PSDC related to industrial productivity are substantial during the Great East Japan Earthquake in 2011. The impact of the disaster on business conditions and cycles is slight, and is a temporal idiosyncratic shock. This must be why the posterior means of the CI estimated using our model are not close to those of the CI of the ESRI during this period. The shocks of the NWH, IPSI, and OP are largest in terms of turning points from economic recession to recovery in 2009, indicating the immediate recovery of the economy from the recession caused investment and working hours to increase in terms of uncertainty. Moreover, the shocks of the RSV and WSV related to commerce tend to be large when a consumption tax started and increased[9]. Thus, it may be difficult to forecast the growth of retail sales and wholesale sales data.

6 Concluding remarks

This study tracks the growth rates of business fluctuations and cycles in Japan by applying the MSDF model framework and coincident economic indicators from February 1985 to December 2014. Moreover, we use two data sets, namely outlier-adjusted data and raw data, and analyze how omitting the outlier affects estimating the business conditions and cycles. We extend the MSDF model by incorporating with a heavy-tailed distribution in the idiosyncratic equations and the factor equation, along with an SV process in the factor equation. From the empirical results of the Bayesian Markov chain Monte Carlo method, the main findings are as follows. From the estimated degrees of freedom in the factor and idiosyncratic equations, both the factor and idiosyncratic components have fat-tail error distributions. Compared with the results using the data with outlier replacement, the business conditions and recession probabilities estimated without replacing the outliers are close to those published by the ESRI. Moreover, the results of the shock decomposition show that the financial crisis is caused by unexpected shocks. It is also shown that the Great East Japan Earthquake was derive from idiosyncratic noise and did not cause a recession. When samples include large shocks, outlier replacement for each macro economic variable is not desirable, and we should use econometric models with heavy-tailed error and stochastic volatility processes.

Finally, we state some remaining issues. First, we do not consider the correlation among the idiosyncratic components. Here, we should use a collapsed dynamic factor analysis applied by Bräusing and Koopman (2014). Second, real GDP also is a key indicator to measure business conditions and cycles. Although real GDP is a quarterly data set, Mariano and Murasawa (2003) and Aruoba *et al.*(2009) estimated the coincident index using a dynamic factor framework by considering the data at mixed frequencies, such as monthly and quarterly data series. Third, we should also use skewed distributions, such as the generalized hyperbolic skew Student's *t*-

[9] The Japanese Government approved consumption tax law in 1988 and carried it out from April 1989. The consumption tax was increased to 5% from 3% in April 1997 and increased again to 8% in April 2014.

distribution and other error distributions (Aas, 2005; Ascari *et al.*, 2015; Nakajima, 2015). Fourth, our model is capable of capturing the common trend, shock, and idiosyncratic shock. Thus, it would be interesting to estimate the dynamics of a global recession and country-specific noise, for example EuroCoin (Altissimo *et al.*, 2001). These topics will be discussed in our future research. However, it is important to know the properties of the business conditions and cycles if a sample includes large shocks, and we think that, in this respect, our empirical results represent an interesting step.

Acknowledgements

We gratefully acknowledge helpful the discussions and suggestions of Yasutomo Murasawa and Toshiaki Watanabe on several points in the paper. This research is supported by a grant-in-aid from Zengin Foundation for Studies on Economics and Finance.

References

[1] Aas, K. (2005), "The generalized hyperbolic skew Student's t-distribution," *Journal of Financial Econometrics*, **4**, pp.275–309.

[2] Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Lippi, L. Reichlin, and G. Veronese (2001), "EuroCoin: A real-time coincident indicator of the Euro area business cycle," CEPR working paper, No. 3108.

[3] Aruoba, S B., F. X. Diebold, and C. Scotti (2009), "Real-time measurement of business conditions," *Journal of Business & Economics Statistics*, **27**, pp.417–427.

[4] Ascari, G., G. Fagiolo, and A. Roventini (2015), "Fat-tail distributions and business-cycle models," *Macroeconomic Dynamics*, **19**, pp.465–476.

[5] Balke, N. S. and T. B. Fomby (1994), "Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series," *Journal of Applied Econometrics*, **9**, pp.181–200.

[6] Bräuning, F., and S. J. Koopman (2014), "Forecasting macroeconomic variables using collapsed dynamic factor analysis," *International Journal of Forecasting*, **30**, pp.572–584.

[7] Camacho, M., G. P. Quiros, and P. Poncela (2014), "Green shoots and double dips in the euro area: A real-time measure," *International Journal of Forecasting*, **30**, pp.52,–535.

[8] Chauvet, M. (1998), "An econometric characterization of business cycle dynamics with factor structure and regime switching," *International Economic Review*, **39**(4), pp.969–996.

[9] Chib, S. (2001), "Markov chain Monte Carlo methods: Computation and inference," *Handbook of Econometrics*, (Heckman, J. J. and E. Leaper. eds), Elsevier, pp.3569–3649.

[10] Chib, S. and E. Greenberg (1995), "Understanding the Metropolis-Hastings algorithm," *American Statistician*, **49**, pp.327–335.

[11] Christiano, L. J. and W. Den Haan (1996), "Small sample properties of GMM for business cycle analysis," *Journal of Business & Economic Statistics*, **14**, 309–327.

[12] Cúrdia, V., M. Del Negro, and D. Greenwald, (2014), "Rare shocks, great recessions," *Journal of Applied Econometrics*, **29**, pp.1031–1052.

[13] Diebold, F. (1988), *Empirical modeling of exchange rate dynamics*, Spriger-Verlag.

[14] Diebold, F. X. (2003), " 'Big data' Dynamic factor models for macroeconomic measurement and forecasting," (discussion of Feichlin and Watson papers), in *Advances in Economics and Econometrics, Eighth World Congress of the Econometric Society*, eds. M. Dewatripoint, L. P. Hansen, and S. Turnovsky, Cambridge: Cambridge University Press, pp.115–122.

[15] Doornik, J. A. (2006), *Ox: Object Oriented Matrix Programming Language*, London: Timberlake Consultants Press.

[16] Durbin, J. and S. J. Koopman (2002), "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, **89**(3), pp.603–616.

[17] Fagiolo, G., M. Napoletano, and A. Roventini (2008), "Are output growth-rate distributions fat-tailed?" *Journal of Applied Econometrics*, **23**, pp.639–669.

[18] Franses, P H. and R. Paap (1999), "Does seasonality influence the dating of business cycle turning points?" *Journal of Macroeconomics*, **21**, pp.79–92.

[19] Fukuda, S. and T. Onodera (2001), "A new composite index of coincident economic indicators in Japan: How can we improve the forecast performance?" *International Journal of Forecasting*, **17**, pp.483–498.

[20] Geweke, J. (1992), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in J. M. Bernardo, J. O. Berger, A. P. David, and A. F. M. Smith, eds. *Bayesian Statistics 4*, Oxford: Oxford University Press, pp.169–193.

[21] Hamilton, J. D. (2011), "Calling recessions in real time," *International Journal of Forecasting*, **27**, pp.1006–1026.

[22] Hayashida, M. and G. J. D. Hewings (2009), "Regional business cycles in Japan," *International Regional Science Review*, **32**, pp.119–147.

[23] Kaufman, S. (2000), "Measuring business cycle with a dynamic Markov switching factor model: An assessment using Bayesian simulation methods," *Econometrics Journal*, **3**, pp.39–65.

[24] Kim, M. and J. Yoo (1995), "New index of coincident indicators: A multivariate Markov switching factor model approach," *Journal of Monetary Economics*, **36**(3), pp.607–630.

[25] Kim, C. J. and C. R. Nelson (1998), "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching," *The Review of Economics and Statistics*, **80**, pp.188–201.

[26] Kim, C. J. and C. R. Nelson (1999), *State-space models with regime switching*, Cambridge, Massachusetts: MIT Press.

[27] Mariano, R. S. and Y. Murasawa (2003), "A new coincident index of business cycles based on monthly and quarterly series," *Journal of Applied Econometrics*, **18**, pp.427–443.

[28] Matas-Mir, A., D. R. Osborn, and M. J. Lombardi (2008), "The effect of seasonal adjustment on the properties of business cycle regimes," *Journal of Applied Econometrics*, **23**, pp.257–278.

[29] Nakajima, J., M. Kasuya, and T. Watanabe (2011), "Bayesian analysis of time-varying parameter vector autoregressive model for the Japanese economy and monetary policy," *Journal of the Japanese and International Economics*, **25**, pp.225–245.

[30] Nakajima, J. (2015), "Bayesian analysis of multivariate stochastic volatility with skew distribution," *Econometric Review*, in press.

[31] Pitt, M. K. and N. Shephard (1999), "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, **94**, pp.590–599.

[32] Primiceri, G. E. (2005), "Time-varying structural vector autoregressions and monetary policy," *Review of Economic Studies*, **72**(3), pp.821–852.

[33] Stock, J. H. and M. W. Watson (2014), "Estimating turning points using large data sets," *Journal of Econometrics*, **178**, pp.368–381.

[34] Tsay, R. S. (1988), "Outliers, level shifts, and variance changes in time series," *Journal of Forecasting*, **7**, pp.1–20.

[35] Tierney, L. (1994), "Markov chains for exploring posterior distributions(with discussion)," *Annals of Statistics*, **23**, pp.1701–1762.

[36] Watanabe, T. (2001), "On sampling the degree-of-freedom of Student's-t disturbances," *Statistics & Probability Letters*, **52**, pp.177–181.

[37] Watanabe, T. (2003), "Measuring business cycle turning points in Japan with a dynamic Markov switching factor model," *Monetary and Economic Studies*, pp.35–68.

[38] Watanabe, T. and Y. Omori (2004), "A multi-move sampler for estimating non-Gausian time series model: Comments on Shephard & Pitt (1997)," *Biometrika*, **91**, pp.246–248.

[39] Watanabe, T. (2014), "Bayesian analysis of business cycle in Japan using Markov switching model with stochastic volatility and fat-tail distribution," *Economic research at Hitotsubashi University*, **65**(2), pp.156–167.

Table 1: Variables used by the ESRI to construct the composite index, and abbreviations

	Abbreviation	Official name
1	IIPP	Index of Industrial Production (Mining and Manufacturing)
2	IIPS	Index of Producers' Shipments (Producer Goods for Mining and Manufacturing)
3	LIPC	Large Industrial Power Consumption
4	PSDC	Index of Producers' Shipment of Durable Consumer Goods
5	NWH	Index of Non-Scheduled Worked Hours (Industries Covered)
6	IPSI	Index of Producers' Shipment (Investment Goods Excluding Transport Equipments)
7	RSV	Retail Sales Value (Change From Previous Year)
8	WSV	Wholesale Sales Value (Change From Previous Year)
9	OP	Operating Profits (All Industries)
10	SME	Index of Sales in Small and Medium-Sized Enterprises (Manufacturing)
11	EJOR	Effective Job Offer Rate (Excluding New School Graduates)

Table 2: Descriptive statistics of growth rates of coincident indicators

	Mean	SD	Skew	Kurt	JB	LB(10)	$LB_2(10)$
IIPP	0.043	1.986	−2.801	24.151	7161.25	13.20	19.93
	(0.105)						
IIPS	0.122	2.082	−2.336	17.153	3322.67	47.97	87.40
	(0.110)						
LIPC	0.074	1.346	−2.751	21.230	5424.25	21.73	45.75
	(0.071)						
PSDC	0.028	4.550	−0.896	17.987	3407.85	15.41	171.13
	(0.240)						
NWH	−0.033	1.272	−0.856	6.713	250.12	75.79	172.99
	(0.067)						
IPSI	0.016	2.464	−0.330	5.204	79.16	45.02	21.23
	(0.130)						
RSV	−0.015	2.551	−0.715	13.569	1701.59	84.20	36.38
	(0.135)						
WSV	−0.021	2.909	−0.236	5.238	78.26	62.15	32.04
	(0.154)						
OP	0.181	4.967	−0.937	24.225	6791.09	231.47	816.03
	(0.262)						
SME	−0.021	1.610	−0.381	4.676	50.66	28.48	57.65
	(0.085)						
EJOR	0.128	1.818	−0.705	4.243	52.87	1084.95	283.05
	(0.096)						

Note: The numbers in parentheses are standard errors. SD, Skew, Kurt, and JB mean standard deviation, skewness, kurtosis, and the Jarque–Bera statistics, which test the null hypothesis of normality, respectively. LB(10) is the Ljung–Box statistics proposed by Diebold (1988) to test the null hypothesis of no autocorrelation up to 10 lags. $LB_2(10)$ is LB statistics of squared growth rate.

Table 3: Descriptive statistics of growth rates of coincident indicators with outlier replacement

	Mean	SD	Skew	Kurt	JB	LB(10)	$LB_2(10)$
IIPP	0.124	1.518	−0.199	3.330	4.01	43.15	90.82
	(0.080)						
IIPS	0.201	1.548	−0.207	3.061	2.61	41.44	162.05
	(0.082)						
LIPC	0.136	0.983	−0.275	3.325	6.12	21.41	57.51
	(0.052)						
PSDC	0.065	3.161	−0.090	2.771	1.27	28.75	54.92
	(0.167)						
NWH	−0.007	1.174	−0.107	3.184	1.19	73.10	74.34
	(0.062)						
IPSI	0.032	2.251	−0.157	2.925	1.55	45.67	36.48
	(0.119)						
RSV	0.006	1.894	−0.029	3.072	0.13	70.85	52.29
	(0.100)						
WSV	−0.006	2.703	−0.078	3.175	0.83	65.76	39.44
	(0.143)						
OP	0.270	2.457	−0.317	3.203	6.62	297.45	631.73
	(0.130)						
SME	0.004	1.452	−0.091	3.181	0.99	39.78	50.51
	(0.077)						
EJOR	0.150	1.720	−0.477	2.960	13.63	1123.89	395.81
	(0.091)						

Note: The numbers in parentheses are standard errors. SD, Skew, Kurt, and JB mean standard deviation, skewness, kurtosis, and the Jarque–Bera statistics, which test the null hypothesis of normality, respectively. LB(10) is the Ljung–Box statistics proposed by Diebold (1988) to test the null hypothesis of no autocorrelation up to 10 lags. $LB_2(10)$ is LB statistics of squared growth rate.

Table 4: Estimation results for the factor equation using unadjusted data

Parameter	Mean	SD	95%CI	CD	IF
$\mu^{(0)}$	−0.962	0.667	[−1.939, −0.445]	0.31	224.15
$\mu^{(1)}$	0.322	0.076	[0.169, 0.473]	0.23	33.05
ϕ	−0.256	0.080	[−0.403, −0.087]	0.78	27.91
ν_f	7.111	2.583	[4.394, 11.925]	0.44	223.03
p_0	0.825	0.077	[0.636, 0.934]	0.37	23.84
p_1	0.952	0.023	[0.897, 0.987]	0.56	13.00
β	0.742	0.037	[0.670, 0.815]	0.71	0.68
ξ^2	0.223	0.016	[0.193, 0.256]	0.13	1.11

Note: Mean, SD, 95%CI, CD, and IF represent the posterior mean, the standard deviation, 95% credible interval, p-value of Geweke's convergence diagnostics, and inefficiency factor, respectively.

Table 5: Estimation results for the idiosyncratic equation using unadjusted data

Variable	Parameter	Mean	SD	95%CI	CD	IF
IIPP	γ	1.233	0.088	[1.067, 1.416]	0.61	67.61
	ψ	-0.207	0.069	[-0.359, -0.089]	0.79	26.74
	σ^2	0.117	0.027	[0.072, 0.176]	0.91	19.18
	ν	2.474	0.388	[2.020, 3.477]	0.29	12.81
IIPS	γ	1.262	0.095	[1.091, 1.460]	0.70	57.01
	ψ	0.206	0.069	[0.062, 0.332]	0.47	18.64
	σ^2	0.209	0.044	[0.135, 0.305]	0.48	20.33
	ν	2.695	0.498	[2.043, 3.945]	0.46	27.22
LIPC	γ	0.408	0.062	[0.291, 0.536]	0.87	16.77
	ψ	-0.109	0.063	[-0.226, 0.018]	0.36	5.34
	σ^2	0.545	0.073	[0.418, 0.700]	0.35	13.35
	ν	3.755	0.813	[2.544, 5.674]	0.61	18.85
PSDC	γ	1.704	0.171	[1.395, 2.064]	0.62	26.52
	ψ	-0.306	0.057	[-0.419, -0.194]	0.41	2.61
	σ^2	3.885	0.519	[2.958, 4.995]	0.81	12.39
	ν	3.12	0.524	[2.259, 4.311]	0.69	17.62
NWH	γ	0.537	0.079	[0.380, 0.688]	0.56	55.63
	ψ	-0.121	0.077	[-0.266, 0.033]	0.33	24.46
	σ^2	1.023	0.124	[0.779, 1.267]	0.94	107.65
	ν	21.178	22.877	[5.158, 84.915]	0.58	201.99
IPSI	γ	1.078	0.107	[0.879, 1.300]	0.88	34.43
	ψ	-0.341	0.05	[-0.438, -0.244]	0.11	5.46
	σ^2	1.947	0.289	[1.436, 2.571]	0.07	22.69
	ν	4.268	1.138	[2.677, 7.107]	0.03	30.16
RSV	γ	0.260	0.068	[0.128, 0.397]	0.86	8.27
	ψ	-0.367	0.045	[-0.456, -0.282]	0.12	5.11
	σ^2	1.703	0.219	[1.309, 2.167]	0.98	8.71
	ν	2.994	0.486	[2.205, 4.091]	0.25	10.46
WSV	γ	0.870	0.112	[0.663, 1.101]	0.82	23.52
	ψ	-0.298	0.05	[-0.396, -0.200]	0.20	1.98
	σ^2	4.560	0.58	[3.499, 5.777]	0.75	36.32
	ν	8.769	4.574	[4.056, 20.972]	0.14	80.23
OP	γ	0.000	0.011	[-0.022, 0.021]	0.51	1.25
	ψ	0.979	0.010	[0.960, 0.997]	0.42	7.52
	σ^2	0.090	0.019	[0.059, 0.133]	0.51	42.52
	ν	2.093	0.019	[2.020, 2.099]	0.04	762.51
SME	γ	0.878	0.073	[0.739, 1.027]	0.80	46.76
	ψ	-0.367	0.055	[-0.473, -0.258]	0.37	5.44
	σ^2	0.365	0.065	[0.255, 0.504]	0.90	23.68
	ν	2.888	0.591	[2.078, 4.294]	0.95	28.46
EJOR	γ	0.056	0.042	[-0.025, 0.140]	0.91	2.60
	ψ	0.761	0.036	[0.691, 0.831]	0.01	1.80
	σ^2	0.953	0.122	[0.739, 1.21]	0.01	18.00
	ν	5.692	2.222	[3.348, 10.099]	0.08	40.92

Note: Mean, SD, 95%CI, CD, and IF represent the posterior mean, the standard deviation, 95% credible interval, p-value of Geweke's convergence diagnostics, and inefficiency factor, respectively.

Table 6: Estimation results for the factor equation using adjusted data

Parameter	Mean	SD	95%CI	CD	IF
$\mu^{(0)}$	-2.379	0.360	$[-3.084, -1.670]$	0.99	37.27
$\mu^{(1)}$	0.350	0.081	[0.203, 0.521]	0.36	21.66
ϕ	-0.133	0.084	$[-0.289, 0.040]$	0.99	13.69
ν_f	31.418	5.622	[21.501, 43.427]	0.58	23.29
p_0	0.608	0.119	[0.367, 0.820]	0.89	18.82
p_1	0.933	0.027	[0.871, 0.972]	0.80	23.03
β	0.742	0.036	[0.671, 0.812]	0.52	0.57
ξ^2	0.223	0.017	[0.192, 0.258]	0.03	1.15

Note: Mean, SD, 95%CI, CD, and IF represent the posterior mean, the standard deviation, 95% credible interval, p-value of Geweke's convergence diagnostics, and inefficiency factor, respectively.

Table 7: Estimation results for the idiosyncratic equation using adjusted data

Variable	Parameter	Mean	SD	95%CI	CD	IF
IIPP	γ	0.898	0.058	[0.784, 1.014]	0.48	44.78
	ψ	-0.234	0.087	[-0.396, -0.053]	0.53	15.35
	σ^2	0.114	0.027	[0.069, 0.175]	0.46	22.25
	ν	2.447	0.393	[2.018, 3.420]	0.44	18.65
IIPS	γ	0.868	0.058	[0.760, 0.984]	0.53	39.65
	ψ	0.101	0.062	[-0.021, 0.221]	0.57	6.43
	σ^2	0.232	0.044	[0.157, 0.330]	0.32	15.59
	ν	2.580	0.434	[2.033, 3.665]	0.32	19.94
LIPC	γ	0.285	0.039	[0.210, 0.364]	0.16	11.34
	ψ	-0.048	0.058	[-0.162, 0.066]	0.28	3.27
	σ^2	0.704	0.071	[0.566, 0.845]	0.10	61.45
	ν	32.590	25.460	[7.191, 100.051]	0.23	129.97
PSDC	γ	1.182	0.117	[0.960, 1.418]	0.48	22.39
	ψ	-0.135	0.061	[-0.252, -0.018]	0.47	8.39
	σ^2	4.588	0.644	[3.473, 5.983]	0.88	51.75
	ν	6.466	2.771	[3.435, 13.851]	0.89	70.65
NWH	γ	0.334	0.058	[0.221, 0.452]	0.50	11.23
	ψ	-0.041	0.071	[-0.179, 0.100]	0.03	4.15
	σ^2	1.018	0.107	[0.813, 1.231]	0.00	45.21
	ν	24.365	22.377	[6.672, 87.060]	0.06	140.98
IPSI	γ	0.937	0.082	[0.778, 1.102]	0.44	19.40
	ψ	-0.267	0.056	[-0.372, -0.154]	0.47	4.03
	σ^2	1.948	0.290	[1.436, 2.575]	0.63	16.90
	ν	5.056	1.635	[2.950, 9.325]	0.96	37.44
RSV	γ	0.248	0.066	[0.120, 0.379]	0.19	4.13
	ψ	-0.264	0.058	[-0.377, -0.150]	0.63	19.21
	σ^2	2.471	0.319	[1.861, 3.102]	0.32	62.45
	ν	15.377	13.131	[4.763, 52.987]	0.03	137.74
WSV	γ	0.736	0.098	[0.549, 0.934]	0.72	11.49
	ψ	-0.268	0.052	[-0.370, -0.165]	0.07	1.47
	σ^2	4.952	0.489	[4.007, 5.935]	0.04	32.61
	ν	30.840	23.537	[7.533, 96.359]	0.00	156.17
OP	γ	0.002	0.005	[-0.007, 0.011]	0.92	1.30
	ψ	0.995	0.003	[0.988, 1.000]	0.72	1.25
	σ^2	0.023	0.005	[0.015, 0.033]	0.59	8.21
	ν	2.140	0.000	[2.140, 2.140]	0.01	1098.75
SME	γ	0.717	0.054	[0.614, 0.823]	0.32	34.65
	ψ	-0.287	0.059	[-0.403, -0.172]	0.07	4.94
	σ^2	0.290	0.051	[0.202, 0.399]	0.91	10.95
	ν	2.666	0.459	[2.041, 3.761]	0.87	14.95
EJOR	γ	0.046	0.034	[-0.020, 0.114]	0.61	2.69
	ψ	0.783	0.036	[0.712, 0.853]	0.28	3.40
	σ^2	0.920	0.110	[0.718, 1.151]	0.82	17.09
	ν	7.032	2.442	[3.853, 13.284]	0.88	36.00

Note: Mean, SD, 95%CI, CD, and IF represent the posterior mean, the standard deviation, 95% credible interval, p-value of Geweke's convergence diagnostics, and inefficiency factor, respectively.

Table 8: Business cycle turning points

ESRI	Unadjusted data	Adjusted data
	Peak	
85/06	–	–
91/02	91/10	92/03
97/05	97/09	98/02
00/10	01/01	01/01
08/02	08/03	08/05
12/04	12/04	12/05
	Trough	
86/11	–	–
93/10	93/1	92/12
99/01	98/09	98/04
02/01	01/12	01/10
09/03	09/04	09/04
–	12/12	12/10

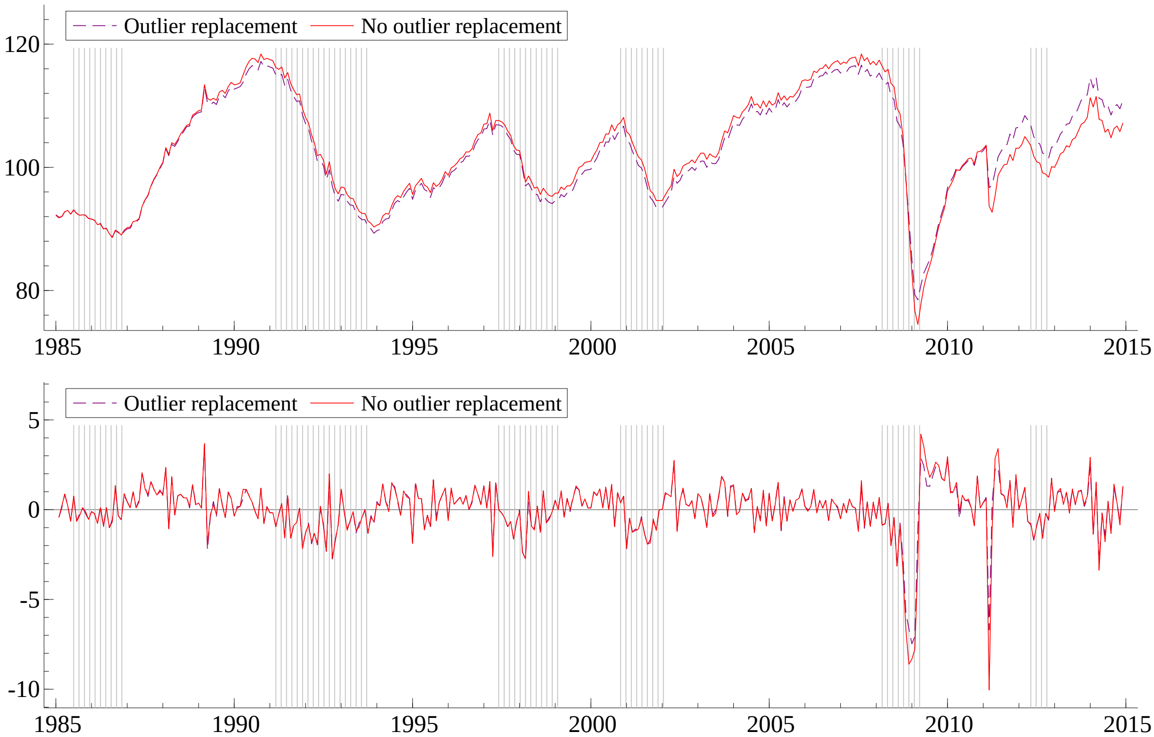

Figure 1: Time series plots of CI (1985/1 – 2014/12)

Note: Raw data (top) and growth rate by log-difference multiplied by 100 (bottom). Dotted and solid lines mean CI with and without outlier replacement, respectively. The shaded bars indicate recessions reported by the ESRI.

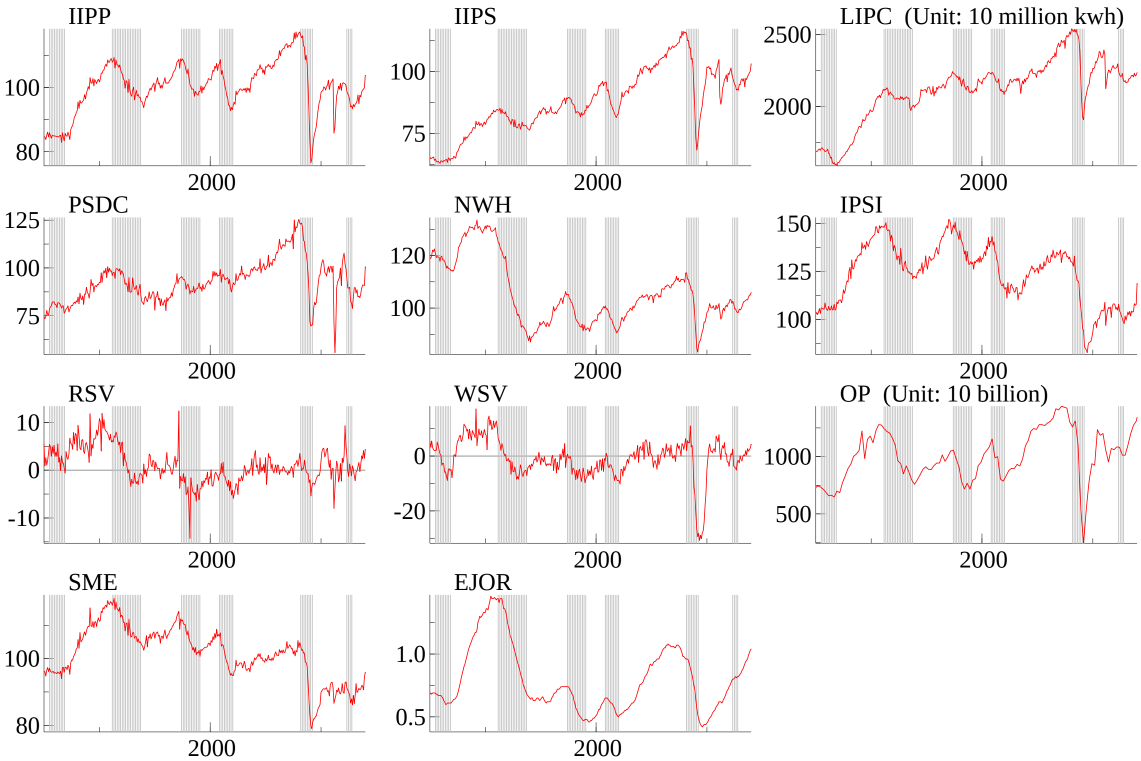

Figure 2: Time series plots of coincident economic indicators (1985/1– 2014/12)

Note: The shaded bars indicate recessions at the national level reported by the ESRI.

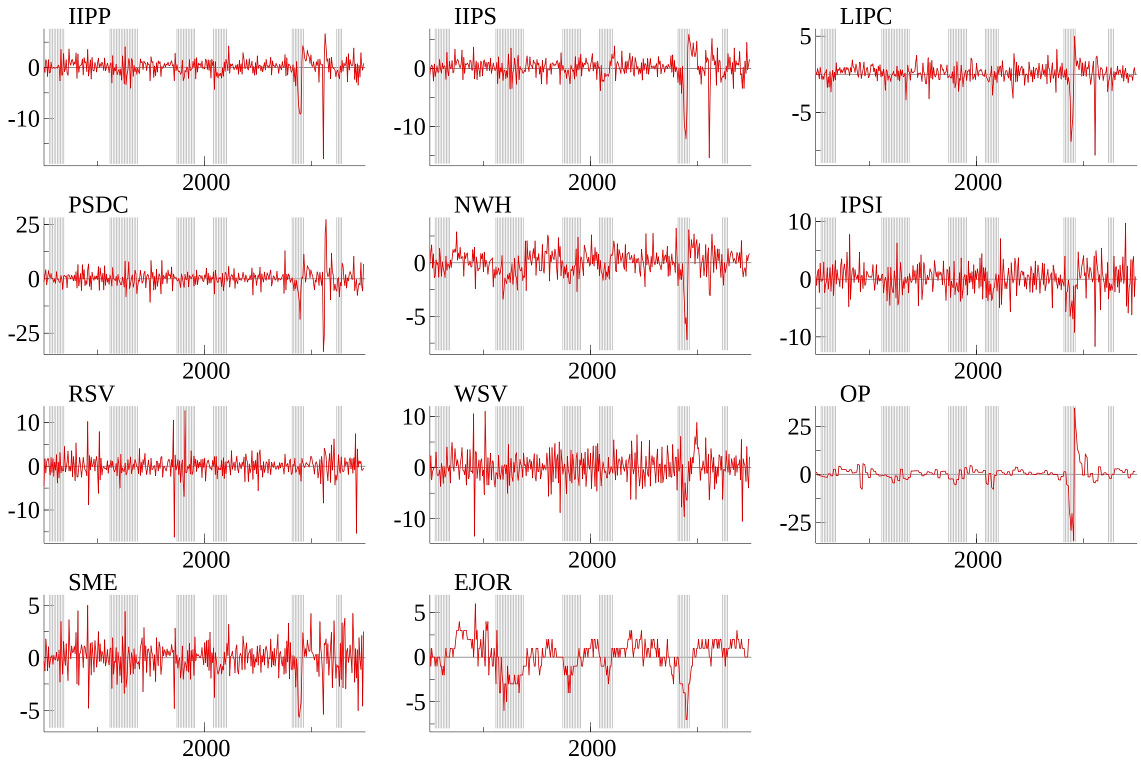

Figure 3: Growth rate of coincident indicators (1985/02– 2014/12)

Note: The shaded bars indicate recessions at the national level reported by the ESRI.

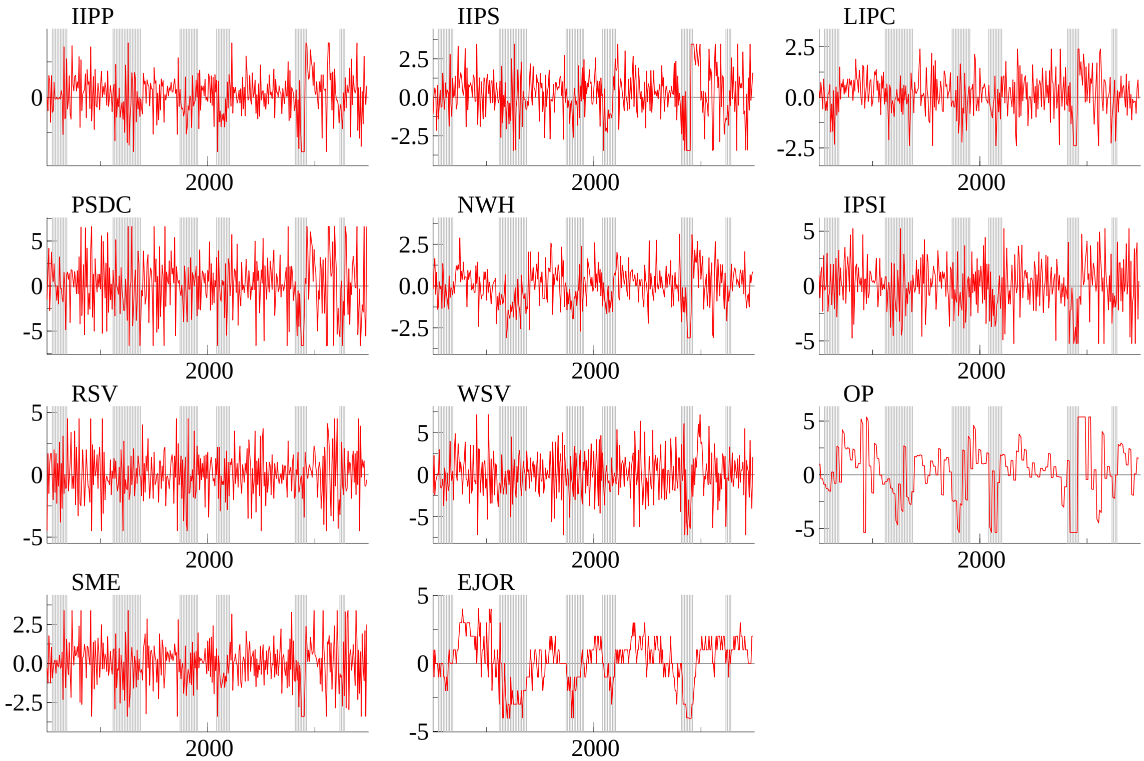

Figure 4: Growth rate of coincident indicators with outlier replacement (1985/02– 2014/12)

Note: The shaded bars indicate recessions at the national level reported by the ESRI.

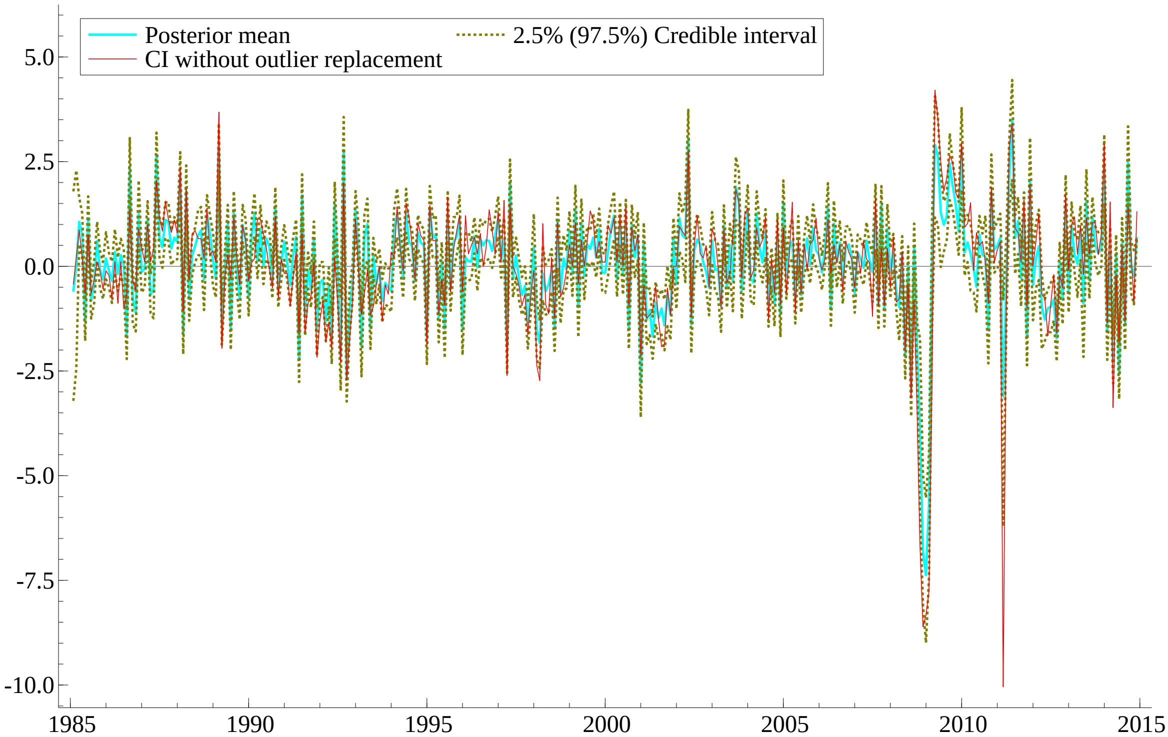

Figure 5: Estimated business conditions using unadjusted data

Note: Dotted, long dash, and solid line denote the posterior means, 2.5% (97.5%) credible interval, and CI without outlier replacement, respectively.

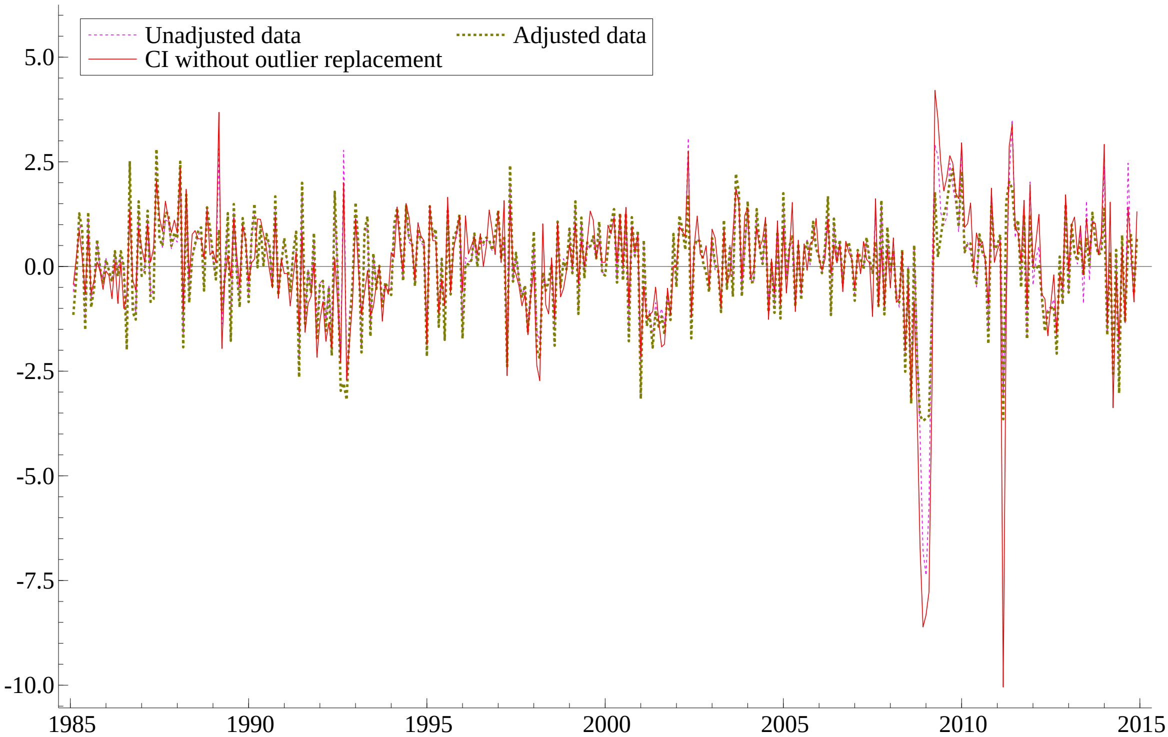

Figure 6: Comparison of posterior means using unadjusted data and adjusted data

Note: Dotted, long dash, and solid line denote the posterior means of the unadjusted data and adjusted data, and CI without outlier replacement, respectively.

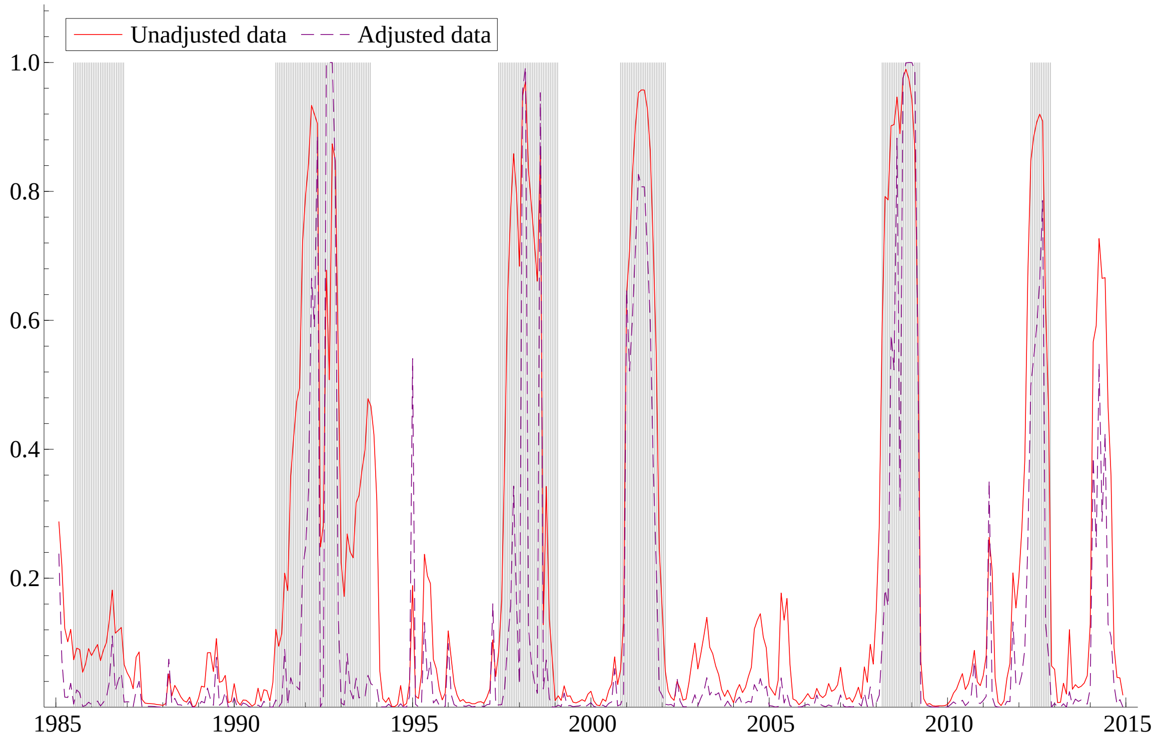

Figure 7: Recession probabilities

Note: Solid line and long dash line denote the posterior recession probabilities using unadjusted data and adjusted data, respectively. The shaded bars indicate recessions at the national level reported by the ESRI.

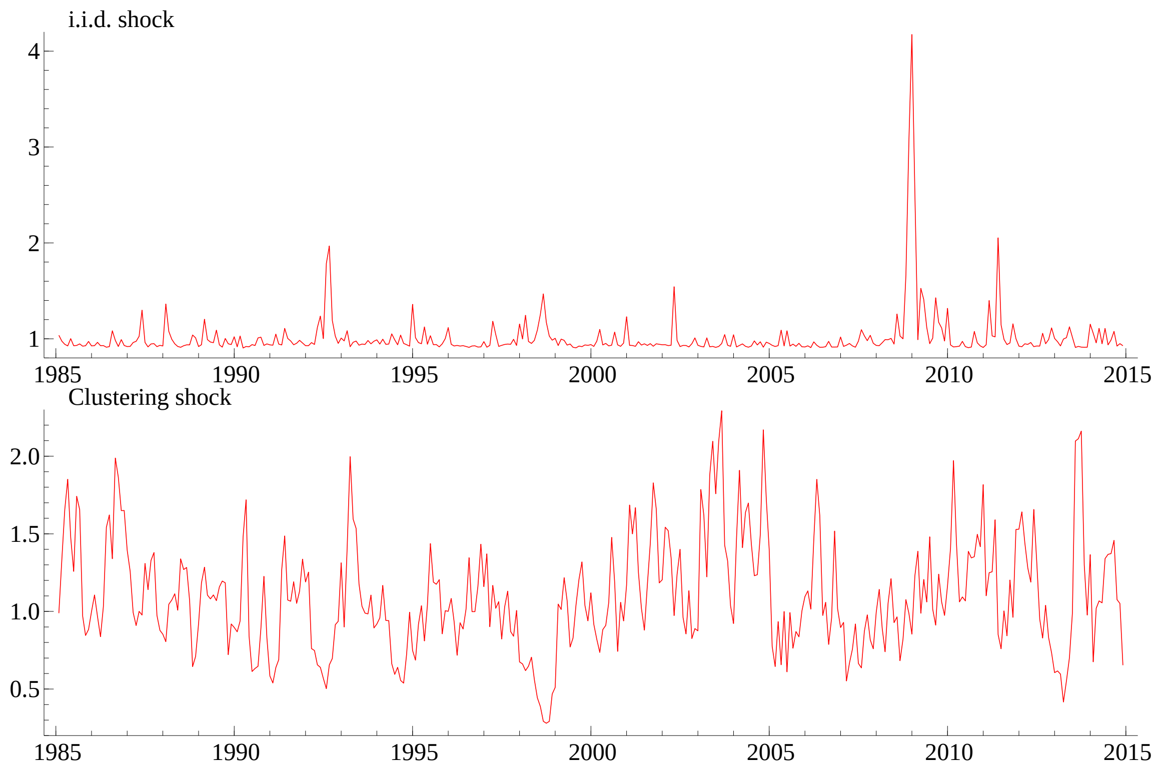

Figure 8: Estimated i.i.d. shocks and clustering shocks

Note: i.i.d. shocks (top) and clustering shocks (bottom) means estimated as $\sqrt{\omega_t}$ and $\exp\left(\frac{h_t}{2}\right)$, respectively.

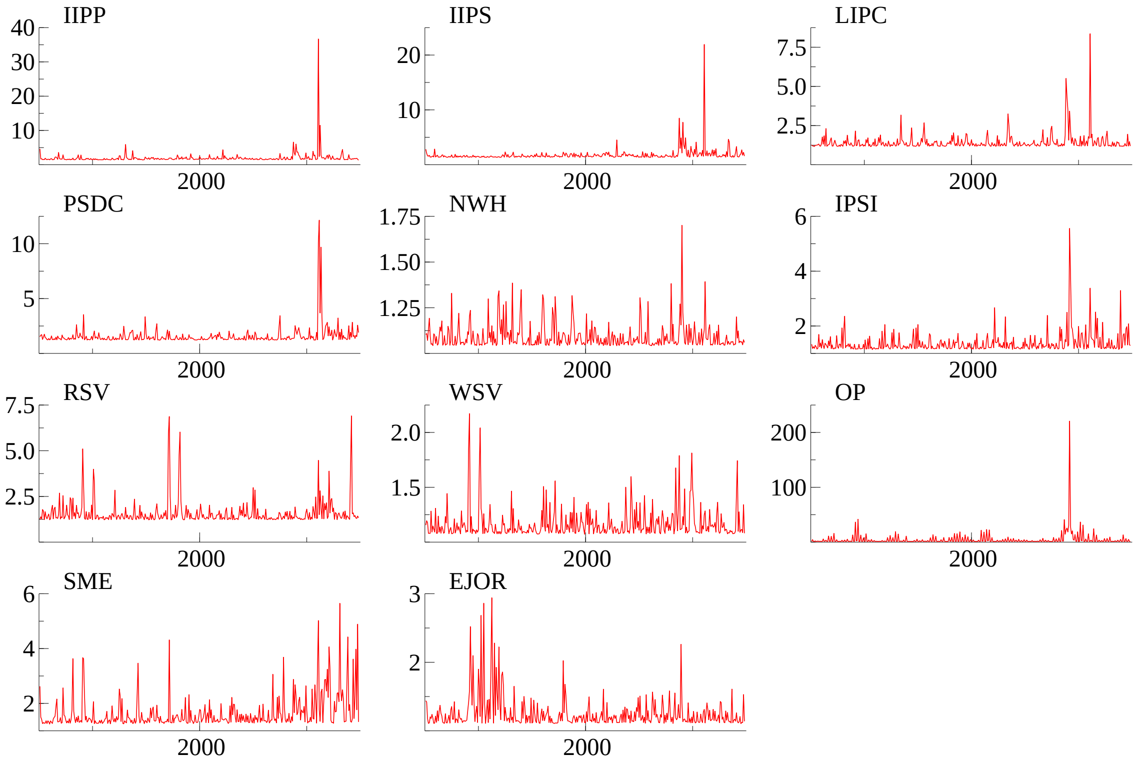

Figure 9: Estimated idiosyncratic shock

Note: Idiosyncratic shock mean estimated $\sqrt{\lambda_{it}}$.

A Sampling algorithm for parameters

A.1 Sampling γ_i

Let $\Delta(\psi_i) = 1 - \psi_i L$, where L denotes a lag operator, and $\psi_i^* = (1 - \psi_i^2)^{\frac{1}{2}}$. Given c and s, the equation (3) can be rewritten as

$$\bar{\mathbf{y}}_i = \bar{\mathbf{x}}_i \gamma_i + \sqrt{\boldsymbol{\lambda}_i} \boldsymbol{\epsilon}_i,$$

where $\bar{\mathbf{y}}_i = (\psi_i^* y_{i1}, \Delta(\psi_i) y_{i2}, \ldots, \Delta(\psi_i) y_{iT})'$, $\bar{\mathbf{x}}_i = (\psi_i^* c_1, \Delta(\psi_i) c_2, \ldots, \Delta(\psi_i) c_T)'$, $\boldsymbol{\lambda}_i = diag(\lambda_{i1}, \ldots, \lambda_T)$, $\boldsymbol{\epsilon}_i = (\epsilon_{1i}, \ldots, \epsilon_{Ti})'$. Thus, the full conditional distribution of γ_i is as follows:

$$\gamma_i | \theta_{-\gamma_i}, \vartheta, y \sim \mathcal{N}(\hat{\mu}_{\gamma_i}, \hat{\sigma}^2_{\gamma_i}), \tag{12}$$

where $\hat{\mu}_{\gamma_i} = \hat{\sigma}^2_{\gamma_i} (\sigma_i^{-2} \bar{\mathbf{x}}_i' \boldsymbol{\lambda}_i^{-1} \bar{\mathbf{y}} + \sigma_{\gamma 0}^{-2} \gamma_0)$ and $\hat{\sigma}^2_{\gamma_i} = (\sigma_i^{-2} \bar{\mathbf{x}}_i' \boldsymbol{\lambda}_i^{-1} \bar{\mathbf{x}}_i + \sigma_{\gamma 0}^{-2})^{-1}$.

A.2 Sampling ψ_i and σ_i^2

For sampling parameter ψ_i, we employ the Metropolis-Hastings algorithm proposed by Chib and Greenberg (1995). The full conditional distribution of ψ_i is given by

$$\pi(\psi_i | \theta_{-\psi_i}, \vartheta, y) \propto g_\psi(\psi_i) \prod_{t=2}^{T} \exp\left[-\frac{(z_{it} - \psi_i z_{i,t-1})^2}{2\lambda_{it}\sigma_i^2} \right],$$

where

$$g_\psi(\psi_i) = (1 + \psi_i)^{a_{\psi,i}-1} (1 - \psi_i)^{b_{\psi,i}-1} \psi_i^* \exp\left[-\frac{(\psi_i^* z_{i1})^2}{2\lambda_{i1}\sigma_i^2} \right].$$

It is difficult to directly draw the parameter. We generate the value from the following candidate distribution:

$$\psi_i | \theta_{-\psi_i}, \vartheta, y \sim \mathcal{TN}_{|\psi_i|<1}(\hat{\mu}_{\psi_i}, \hat{\sigma}^2_{\psi_i}),$$

where

$$\hat{\mu}_{\psi_i} = \hat{\sigma}^{-2}_{\psi_i} \sum_{t=2}^{T} \frac{z_{it} z_{i,t-1}}{\lambda_{it}}, \text{ and } \hat{\sigma}^2_{\psi_i} = \frac{\sigma_i^2}{\sum_{t=2}^{T} \frac{z_{i,t-1}^2}{\lambda_{it}}}.$$

Let ψ_i^{old} be the previous value. Then, we draw a candidate ψ_i^{new} from $\mathcal{N}(\hat{\mu}_{\psi_i}, \hat{\sigma}^2_{\psi_i})$, truncated on $(-1, 1)$, in order to satisfy the stationary condition, and accept it with probability

$$\alpha(\psi_i^{old}, \psi_i^{new}) = \min\left[\frac{g_\psi(\psi_i^{new})}{g_\psi(\psi_i^{old})}, 1 \right].$$

Next, the full conditional distribution of σ_i^2 is as follows:

$$\sigma_i^2 | \theta_{-\sigma_i^2}, \vartheta, y \sim \mathcal{IG}\left(\frac{\hat{\tau}_i}{2}, \frac{\hat{\delta}_i}{2} \right),$$

where $\hat{\tau}_i = \tau_{0i} + T$ and $\hat{\lambda}_i = \sum_{t=1}^{T} \frac{e_{it}^2}{\lambda_{it}} + \delta_{0i}$, with

$$e_{it} = \begin{cases} \psi_i^* z_{i1} & (t = 1) \\ z_{it} - \psi_i z_{i,t-1} & (t > 1) \end{cases}.$$

A.3 Sampling c

We show the state space representation of the model for drawing the latent variable c. Let $\Delta y_{it} = y_{it} - \psi_i y_{i,t-1}$, for $i = 1, \ldots, n$, and $\Delta \mathbf{y}_t = (\Delta y_{1t}, \ldots, \Delta y_{nt})'$ and $\sigma^2_{ht} = \exp(h_t)$. Then, the model can be represented as

$$\begin{aligned} \Delta \mathbf{y}_t &= \mathbf{\Gamma} \boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{H}_t), & (13)\\ \boldsymbol{\alpha}_t &= \mathbf{m}_t + \mathbf{T}\boldsymbol{\alpha}_{t-1} + \mathbf{G}_t \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2), & (14) \end{aligned}$$

where $\mathbf{I}_2$ denotes 2×2 unit matrix, and $\mathbf{\Gamma}$, $\boldsymbol{\alpha}_t$, $\mathbf{H}_t$, $\mathbf{m}_t$, $\mathbf{T}$, and $\mathbf{G}_t$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1 & -\gamma_1\psi_1 \\ \vdots & \vdots \\ \gamma_n & -\gamma_n\psi_n \end{bmatrix}, \ \boldsymbol{\alpha}_t = \begin{bmatrix} c_t \\ c_{t-1} \end{bmatrix}, \ \mathbf{H}_t = diag(\lambda_{1t}\sigma_1^2, \ldots, \lambda_{nt}\sigma_n^2),$$

$$\mathbf{m}_t = \begin{bmatrix} \mu_t - \phi\mu_{t-1} \\ 0 \end{bmatrix}, \ \mathbf{T} = \begin{bmatrix} \phi & 0 \\ 1 & 0 \end{bmatrix}, \ \mathbf{G}_t = \begin{bmatrix} \sqrt{\omega_t \sigma^2_{ht}} & 0 \\ 0 & 0 \end{bmatrix}.$$

Since equations (13) and (14) constitute the linear Gaussian state space mode, we can sample c using the efficient simulation smoother (Durbin and Koopman, 2002).

A.4 Sampling μ

Let $\Delta(\phi) = 1 - \phi L$, $\phi^* = (1-\phi^2)^{\frac{1}{2}}$, and $\tilde{\mathbf{x}} = (1 - s_t, s_t)$. Given c and s, the equation is rewritten as:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{x}}\mu + \boldsymbol{\omega}^{\frac{1}{2}}\boldsymbol{\eta},$$

where $\tilde{\mathbf{y}} = (\phi^* c_1, \Delta(\phi)c_2, \ldots, \Delta(\phi)c_T)'$, $\tilde{\mathbf{x}} = (\phi^*\mathbf{x}_1', \Delta(\phi)\mathbf{x}_2', \ldots, \Delta(\phi)\mathbf{x}_T')'$, $\boldsymbol{\omega} = diag(\omega_1, \ldots, \omega_T)$, and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_T)'$. The full conditional distribution of μ can be obtained as

$$\mu|\theta_{-\mu}, \vartheta, y \sim \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\mathbf{\Sigma}}_{\boldsymbol{\mu}}) I[\mu^{(0)} < \mu^{(1)}],$$

where $\hat{\boldsymbol{\mu}} = \hat{\mathbf{\Sigma}}_{\boldsymbol{\mu}}(\tilde{\mathbf{x}}'\boldsymbol{\omega}^{-1}\tilde{\mathbf{y}} + \mathbf{\Sigma}_{\mu 0}^{-1}\boldsymbol{\mu}_0)$ and $\hat{\mathbf{\Sigma}}_{\boldsymbol{\mu}} = (\tilde{\mathbf{x}}'\boldsymbol{\omega}^{-1}\tilde{\mathbf{x}} + \mathbf{\Sigma}_{\mu 0}^{-1})^{-1}$. For sampling $\boldsymbol{\mu}$, if it does not satisfy the inequality $\mu^{(0)} < \mu^{(1)}$, the generated values are rejected and then sampled again.

A.5 Sampling ϕ

We employ the MH algorithm as in sampling ψ_i. Let $\dot{c}_t = c_t - \mu_t$. The full conditional distribution of ϕ is given by

$$\pi(\phi|\theta_{-\phi}, \vartheta, y) \propto g_\phi(\phi) \prod_{t=2}^{T} \exp\left[-\frac{(\dot{c}_t - \phi\dot{c}_{t-1})^2}{2\omega_t\sigma^2_{ht}}\right],$$

where

$$g_\phi(\phi) = (1+\phi)^{a_\phi - 1}(1-\phi)^{b_\phi - 1}\sqrt{1-\phi^2}\exp\left[-\frac{(1-\phi^2)\dot{c}_1^2}{2\omega_1\sigma^2_{h1}}\right].$$

We draw the proposal from the following candidate distribution

$$\phi|\theta_{-\phi}, \vartheta, y \sim \mathcal{TN}_{|\phi|<1}(\hat{\mu}_\phi, \hat{\sigma}^2_\phi),$$

where

$$\hat{\mu}_\phi = \frac{\sum_{t=2}^T \dot{c}_t \dot{c}_{t-1}}{\sum_{t=2}^T \frac{\dot{c}_{t-1}^2}{\omega_t \sigma_{ht}^2}}, \text{ and } \hat{\sigma}_\phi^2 = \frac{1}{\sum_{t=2}^T \frac{\dot{c}_{t-1}^2}{\omega_t \sigma_{ht}^2}}.$$

Let ϕ^{old} be the previous value. Then, we draw a candidate ϕ_i^{new} from $\mathcal{N}(\hat{\mu}_\phi, \hat{\sigma}_\phi^2)$, truncated on $(-1, 1)$, in order to satisfy the stationary condition, and accept it with probability

$$\alpha(\phi^{old}, \phi^{new}) = \min\left[\frac{g_\phi(\phi^{new})}{g_\phi(\phi^{old})},\ 1\right].$$

A.6 Sampling p

Following Watanabe (2014), we employ the acceptance rejection algorithm for sampling p_0 and p_1. The full conditional distribution of p is following as:

$$\begin{aligned}
\pi(p|\theta_{-p}, \vartheta, y) &\propto \frac{(1-p_0)^{s_1}(1-p_1)^{1-s_1}}{2-p_0-p_1} \\
&\quad \times p_0^{\iota_{00}+n_{00}}(1-p_0)^{\iota_{01}+n_{01}} p_1^{\iota_{11}+n_{11}}(1-p_1)^{\iota_{10}+n_{10}} \\
&= g_p(p) \times p_0^{\iota_{00}+n_{00}}(1-p_0)^{\iota_{01}+n_{01}} p_1^{\iota_{11}+n_{11}}(1-p_1)^{\iota_{10}+n_{10}},
\end{aligned}$$

where n_{ij} means the number of transitions from state i to j. We sample a proposed value p_0^{new} and p_1^{new} from the following independent beta distribution:

$$p_0^{new} \sim Beta(\iota_{00}+n_{00}, \iota_{01}+n_{01}), \quad p_1^{new} \sim Beta(\iota_{11}+n_{11}, \iota_{10}+n_{10}).$$

Since $0 < g_p(p^{new}) < 1$, we employ the MH step. Finally, we accept the proposed values with probability

$$\alpha(p^{old}, p^{new}) = \min\left[\frac{g_p(p^{new})}{g_p(p^{old})}, 1\right].$$

A.7 Sampling s

For sampling s, we employ the multi-move sampler (Kim and Nelson, 1998, 1999). The joint conditional distribution of s is as follows:

$$f(s|\theta, \vartheta_{-s}, y) = f(s_T|\theta, \vartheta_{-s_T}, y) \prod_{t=1}^{T-1} f(s_t|s_{t+1}, \theta, \vartheta_{-s}, y), \tag{15}$$

First, we sample s_T, which is the first term on the right-hand side of equation (15). Given s_T sampled from s_{T-1}, we can proceed backwards in time. Then, $f(s_t|s_{t+1}, \theta, \vartheta_{-s}, y)$ includes the following:

$$f(s_t|s_{t+1}, \theta, \vartheta_{-s}, y) \propto f(s_{t+1}|s_t) f(s_t|\theta, \vartheta_{-s}, y),$$

where $f(s_{t+1}|s_t)$ means the transition probability. Next, $f(s_t|\theta, \vartheta_{-s}, y)$, for $t = 1, \ldots, T$, is calculated using Hamilton's (1989) filter. s_T is sampled from $f(s_T|\theta, \vartheta_{-s}, y)$, and s_t is sampled using s_T and equation (16) backward in time.

A.8 Sampling λ_{it} and ν_i

Given c, s, and ω, the full conditional distribution of λ_{it} is as follows:

$$\lambda_{it}|\theta, \vartheta_{-\lambda_{it}}, y \sim \mathcal{IG}\left(\frac{\hat{a}_{it}}{2}, \frac{\hat{b}_{it}}{2}\right).$$

where $\hat{a}_{it} = \nu_i + 1$ and $\hat{b}_{it} = \nu_i + \sigma_i^{-2}(z_{it} - \psi_i z_{i,t-1})^2$.

Finally, the full conditional distribution of ν_i is given by

$$\pi(\nu_i|\lambda_i) \quad \propto \quad \nu^{A_0-1}\exp(-B_0\nu_i)\prod_{t=1}^{T}\frac{\frac{\nu_i}{2}^{\frac{\nu_i}{2}}}{\Gamma\left(\frac{\nu_i}{2}\right)}\lambda_{it}^{-\frac{\nu_i}{2}+1}\exp\left(-\frac{\nu_i}{2\lambda_{it}}\right), \tag{16}$$

To sampling the degrees of freedom parameter, we employ the AR-MH algorithm extended by Watanabe (2001). The AR-MH algorithm was proposed by Tierney (1994) (see also Chib and Greenberg (1995) for details). This algorithm samples the parameter using the AR and MH step. Suppose there is a candidate function $h(\nu_i)$ which can be directly sampled, and $f(\nu_i)$, defined as the target distribution. Then, the AR step proceeds as follows:

1. Sample the candidate ν_i from $h(\nu_i)$ and u from the uniform distribution $\mathcal{U}[0,1]$.
2. If $u \leq \frac{f(\nu_i)}{ch(\nu_i)}$, return $\nu_i^{new} = \nu_i$. Else, go to 1.

This step is repeated until the candidate draw is accepted. In this study, we utilize a normal distribution as the candidate function. Let $p^*(\nu_i)$ denote $p(\nu_i|\omega)$ with the constant subtracted, and the log of $p^*(\nu_i)$ is given by

$$\ln p^*(\nu_i) = \frac{T}{2}\nu_i \ln\left(\frac{\nu_i}{2}\right) - T\ln\Gamma\left(\frac{\nu_i}{2}\right) - J_i\nu_i + (A_0 - 1)\ln(\nu_i), \tag{17}$$

where

$$J_i = \frac{1}{2}\sum_{t=1}^{T}\left\{\ln(\lambda_{it}) + \frac{1}{\lambda_{it}}\right\} + B_0.$$

We apply the second-order Taylor expansion around $\nu_i = \nu_i^*$ to (17), which yields

$$\begin{aligned}\ln p^*(\nu_i) \quad &\approx \quad \ln p^*(\nu_i^*) + C_i'(\nu_i - \nu_i^*) + \frac{C_i''}{2}(\nu_i - \nu_i^*)^2 \\ &= \quad h(\nu_i),\end{aligned}$$

where

$$\begin{aligned}C_i' \quad &= \quad \left.\frac{d\ln p^*(\nu_i)}{d\nu_i}\right|_{\nu_i=\nu_i^*} = \frac{T}{2}\left\{\ln\left(\frac{\nu_i}{2}\right) + 1 - \psi\left(\frac{\nu_i^*}{2}\right)\right\} - J_i + \frac{A_0 - 1}{\nu_i^*}, \\ C_i'' \quad &= \quad \left.\frac{d^2\ln p^*(\nu_i)}{d^2\nu_f}\right|_{\nu_i=\nu_i^*} = \frac{T}{2}\left\{\frac{1}{\nu_i} - \frac{1}{2}\psi'\left(\frac{\nu_i^*}{2}\right)\right\} - \frac{A_0 - 1}{\nu_i^{*2}},\end{aligned}$$

with $\psi(\nu_i)$ and $\psi'(\nu_i)$ denoting a digamma function defined by $\psi(\nu_i) = \frac{d\ln\Gamma(\nu_i)}{d\nu_i}$, and a trigamma function defined by $\psi'(\nu_i) = \frac{d\psi(\nu_i)}{d\nu_i}$. Then, the normalized version of $h(\nu_i)$ has a normal density with mean $\nu_i^* - \frac{C_i'}{C_i''}$ and variance $-\frac{1}{C_i''}$.

Next, let the previous sampled value of ν_i be $\bar{\nu_i}$. Then, the MH step proceeds as follows:

1. Calculate the acceptance probability q

   - If $p^*(\bar{\nu_i}) < \kappa h(\bar{\nu_i})$, then set $q = 1$;
   - If $p^*(\bar{\nu_i}) \geq \kappa h(\bar{\nu_i})$ and $p^*(\nu_i^{new}) < \kappa h(\nu_i^{new})$, then set $q = \dfrac{\kappa h(\bar{\nu_i})}{p^*(\bar{\nu_i})}$;
   - If $p^*(\bar{\nu_i}) \geq \kappa h(\bar{\nu_i})$ and $p^*(\nu_i^{new}) \geq \kappa h(\nu_i^{new})$, then set $q = \min\left[\dfrac{p^*(\nu_i^{new})h(\bar{\nu_i})}{p^*(\bar{\nu_i})h(\nu_i^{new})}, 1\right]$;

2. Sample a value u from the uniform distribution $\mathcal{U}[0,1]$.

3. If $u \leq q$, return $\nu_i = \nu_i^{new}$. Else, return $\nu_i = \bar{\nu_i}$.

In this step, the candidate value is accepted with probability q, and otherwise rejected. If a draw is rejected, the previously sampled value is sampled again. In the empirical analysis, we set $\kappa = 1$.

A.9 Sampling ω_t and ν_f

Since the full conditional distribution of ω_t are mutually independent, it is straightforward to sample ω_t. Thus, the full conditional distribution of ω_t are given as

$$\left.\frac{\eta_t^2 + \nu_f - 2}{\left(\frac{h_t}{2}\right)\omega_t}\right| \theta, \vartheta_{-\omega_t}, y \sim \chi^2(\nu_f + 1), \ t = 1, \ldots, T.$$

Finally, we sample ν_f using the AR-MH algorithm, as in sampling ν_i.

A.10 Sampling β and ξ^2

Given h, the full conditional distribution of β is given by

$$\pi(\beta|\theta, \vartheta_{-\beta}, y) \propto g_\beta(\beta) \prod_{t=2}^{T} \exp\left[-\frac{(h_t - \beta h_{t-1})^2}{2\xi^2}\right],$$

where

$$g_\beta(\beta) = (1+\beta)^{a_\beta - 1}(1-\beta^2)^{b_\beta - 1}\sqrt{1-\beta^2}\exp\left[-\frac{\{(1-\beta^2)h_1\}^2}{2\xi^2}\right].$$

It is difficult to directly draw the parameter. We generate the value from the following candidate distribution

$$\beta|\theta_{-\beta}, \vartheta, y \sim \mathcal{TN}_{|\beta|<1}(\hat{\mu}_\beta, \hat{\sigma}_\beta^2),$$

where

$$\hat{\mu}_\beta = \hat{\sigma}_\beta^{-2}\sum_{t=2}^{T} h_t h_{t-1}, \text{ and } \hat{\sigma}_\beta^2 = \frac{\xi^2}{\sum_{t=2}^{T} h_{t-1}^2}.$$

Let β^{old} be the previous value. Then, we draw a candidate β^{new} from $\mathcal{N}(\hat{\mu}_\beta, \hat{\sigma}_\beta^2)$, truncated on $(-1,1)$, in order to satisfy the stationary condition, and accept it with probability

$$\alpha(\beta^{old}, \beta^{new}) = \min\left[\frac{g_\beta(\beta^{new})}{g_\beta(\beta^{old})},\ 1\right].$$

Next, the full conditional distribution of ξ^2 is as follows:

$$\xi^2|\theta_{-\xi^2}, \vartheta, y \sim \mathcal{IG}\left(\frac{\hat{\tau}_h}{2}, \frac{\hat{\delta}_h}{2}\right),$$

where $\hat{\tau}_h = \tau_{0h} + T$ and $\hat{\delta}_h = \sum_{t=1}^T e_{ht}^2 + \delta_{0h}$, with

$$e_{ht} = \begin{cases} \sqrt{1-\beta^2}h_1 & (t = 1) \\ h_t - \psi_i h_{t-1} & (t > 1) \end{cases}.$$

A.11 Sampling h

For sampling the latent variable h, we employ the multi-move sampler extended by Watanabe and Omori (2004). First, we divide h into $K+1$ blocks, $(h_{k_{l-1}}, \ldots, h_{k_l})$ for $l = 1, \ldots, K+1$ with $k_0 = 0$ and $k_{K+1} = T$. The K knots $(k_1, \ldots, k_K)$ are randomly drawn from

$$k_j = int\left[T \times \frac{j \times U_j}{K+2}\right],$$

where U_l are independent uniforms in $[0, 1]$ and "*int*" means the integer part. Following Pitt and Shephard (1997), we draw the error term $(\zeta_{k_{j-1}}, \ldots, \zeta_{k_j - 1})$ instead of $(h_{k_{i-1}+1}, \ldots, h_{k_i})$ from their full conditional distributions,

$$\pi(\eta_{t-1}, \ldots, \eta_{t+k-1}|h_{t-1}, h_{t+k+1}, \theta, \vartheta_{-h_{k_{i-1}+1},\ldots,h_{k_i}}). \tag{18}$$

Next, let $k_{l-1} = t-1$, $k_l = t+k$, and $h(k) = \{h_j\}_{j=t-1}^{t+k}$. Then, we construct a candidate distribution in order to sampling the error vectors. The log of the posterior density (18) is described as follows:

$$\begin{aligned}
&\log \pi(\zeta_{t-1}, \ldots, \zeta_{t+k}|h_{t-1}, h_{t+k+1}, \theta, \vartheta_{-h(k)}, y) \\
&= \text{const.} + \frac{1}{2\xi^2}\sum_{j=t}^{t+k}\zeta_j^2 - \frac{1}{2}\sum_{j=t}^{t+k}\left\{h_j + \frac{e_j^{*2}}{\omega_j}\exp(-h_j)\right\} - \frac{1}{2\xi^2}(h_{t+k+1} - \beta h_{t+k})^2,
\end{aligned}$$

where e_j^* means the residual of equation (4). Then, we evaluate this logarithm of the posterior density using the Taylor expansion of the log-likelihood,

$$l(h_j) = -\frac{1}{2}h_j - \frac{1}{2}\frac{e_j^{*2}}{\omega_j}\exp(-h_j),$$

around the mode $\hat{h}_j$, as follows:

$$\begin{aligned}
&\log \pi(\zeta_{t-1}, \ldots, \zeta_{t+k-1}|\theta, \vartheta_{-h(k)}, y) \\
&\approx \text{const.} + \sum_{j=t}^{t+k}\left\{l(\hat{h}_j) + l'(\hat{h}_j)(h_j - \hat{h}_j) + \frac{1}{2}l''(\hat{h}_j)(h_j - \hat{h}_j)^2\right\} \\
&+ -\frac{1}{2}\sum_{j=t}^{t+k}\left\{h_j + \frac{e_j^{*2}}{\omega_j}\exp(-h_j)\right\} - \frac{1}{2\xi^2}(h_{t+k+1} - \beta h_{t+k})^2 \\
&\equiv \log g_h(\zeta_{t-1}, \ldots, \zeta_{t+k-1}),
\end{aligned}$$

where

$$
\begin{aligned}
l'(\hat{h}_j) &\equiv \frac{\partial l(\hat{h}_j)}{\partial h_j} = -\frac{1}{2} + \frac{1}{2}\frac{e_j^{*2}}{\omega_j}\exp(-\hat{h}_j), \\
l''(\hat{h}_j) &\equiv \frac{\partial l(\hat{h}_j)}{\partial h_j} = -\frac{1}{2}\frac{e_j^{*2}}{\omega_j}\exp(-\hat{h}_j).
\end{aligned}
$$

We sample the error term from the posterior distribution with the simulation smoother. Moreover, we employ the AR-MH algorithm. Finally, in order to select the posterior mode $\hat{h}_j$, we apply the Kalman filter and disturbance smoother (Watanabe and Omori, 2004).