



Estimating Cambodia's Economic Conditions by Dynamic Factor Model

Kimleng Sa 



Graduate School of Public Policy, The University of Tokyo, Japan.
Email: sakimleng@yahoo.com

Abstract

This study constructed a coincident indicator (CI) as the unobserved state of the economy in Cambodia by combining principal components and a dynamic factor model (DFM). In the first step, it estimated the factor loadings (coefficients of the unobserved state variables) by ordinary least squares (OLS) and feasible generalized least squares (FGLS) methods using the state variable produced by the first principal component. In the second step, it estimated the unobserved state variables through the DFM by replacing the coefficients with their estimators in the first step. Doz, Giannone, and Reichlin (2011) introduced this hybrid approach for stationary data. The coincident indicator showed that Cambodia's economy fell below its potential level between 2016 and 2017 and started recovering after mid-2017. By exploiting the coincident index, the study examined comovement between the foreign direct investment (FDI) inflow and the state of the economy by using the autoregressive distributed lags (ARDL) model. The result showed that an acceleration of the economic condition contributed to an increase in FDI inflow in the short-term for all models; the long-term coefficient became negative. One reason for this could be the diminishing marginal product of capital that made foreign capital investment less attractive.

Keywords: Coincident indicator, Dynamic factor model, Principal components, Kalman filter, State of the economy, Foreign direct investment.

JEL Classification: E1, E3.

Citation | Kimleng Sa (2020). Estimating Cambodia's Economic Conditions by Dynamic Factor Model. Asian Journal of Economics and Empirical Research, 7(2): 268-281.

History:


Received: 30 September 2020

Revised: 19 September 2020

Accepted: 16 November 2020

Published: 7 December 2020

Licensed: This work is licensed under a [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

Attribution 3.0 License 

Publisher: Asian Online Journal Publishing Group

Funding: This study received no specific financial support.

Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

Transparency: The author confirm that the manuscript is an honest, accurate, and transparent account of the study was reported; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Ethical: This study followed all ethical practices during writing.

Contents

1. Introduction	269
2. Literature Review	269
3. Methodology	271
4. Result of the Study	273
5. Discussion and Limitation of the Study	274
6. Conclusion	274
References	275

Contribution of this paper to the literature

This study contributes to existing literature by constructing a coincident indicator (CI) as the unobserved state of the economy in Cambodia.

1. Introduction

The early warning system plays a crucial role in monitoring economic activity in Cambodia. Despite high economic growth rate, the economy was vulnerable to shocks (World Bank, 2018).¹ The IMF (2018) identified export and fiscal shocks, and contingency liability, as major risks that could slow down Cambodia's economy. Without an early warning system the countercyclical policy may be plausible and sometimes even push the economy into a worse condition. Wyplosz (2005) emphasized the use of fiscal policy as a countercyclical policy could be inefficient and possibly do more harm than good. For these reasons, a well-designed early warning system could lower the magnitude of shocks, improve the surveillance capacity, and serve as an alert when the economy moves into a recession.

Monitoring economic activity, and mainly observing the business cycle, requires a high-frequency indicator that represents the cyclical movement of the economy, and gross domestic product (GDP) would be a suitable candidate (OECD, 2010). However, monthly and quarterly GDP is not available for Cambodia mainly due to resource and time constraints for compiling such data. For these reasons, there is a need to create a similar indicator that could be closely correlated to GDP to monitor economic conditions, and to observe the business cycle.

Similar to GDP, the coincident indicator could be a useful tool in the study of the business cycle. The CI, for the unobserved state of the economy, is constructed by exploiting its relationship with high-frequency observable variables, which are available on a monthly basis. It compiles high-frequency data into a single index that captures the current state of the economy and real-time economic performance. The OECD (2010) used macroeconomic variables that hold both economic and statistical relevance to quarterly real GDP in establishing a coincident indicator. By proper design, it could help policymakers in surveillance activity and monitoring the real sector to signal when the economy moves into a recession. Additionally, it could be a useful tool for policymakers dealing with unemployment and inflation issues during the peaks and troughs (Zarnowitz & Moore, 1983). It was also a useful indicator in the analysis of the short-term macroeconomic dynamic (Guo, Ozyildirim, & Zarnowitz, 2009).

Although the coincident index helps to monitor economic performance, there are some challenges to be considered. One issue concerns the number of variables used for its establishment. Caggiano, Kapetanios, and Labhard (2011) suggested that using many variables is not always the best solution. In addition, how well this index can signal a recession is unknown. The problem concerns when the real economy will respond or how long it will take to come into effect. The OECD (2010) suggested that the coincident indicator signaled five to six months in advance of the economy going into recession. The coincident indicator should be used cautiously and will require adjustment when additional information is available. Despite its limitations, there is no doubt that the coincident index plays a crucial role for policymakers in monitoring economic activities.

Exploiting the CI, the study examined the movement of foreign direct investment inflow and the state of the economy. Large capital mobility could put the economy at risk due to the lack of capital control mechanisms. Although it may be difficult to adjust FDI within a short period, a more liquid FDI tends to flow out of a country when the economy moves into a bad situation, such as a recession. Large capital mobility could put the balance of payment at risk due to exchange rate depreciation, especially for net debtor countries.

This paper has the following structure: section two reviews the empirical literature of the coincident index and the dynamic factor model, section three identifies the methodologies used to establish the coincident indicator, section four shows the results, section five discusses the limitations of the study, and the final part comprises the concluding remarks.

2. Literature Review

The coincident index for the unobserved state of the economy has often been used in the study of the business cycle, for example, by Stock and Watson (1989); Kim and Nelson (1998); Altissimo, et al. (2001); Aruoba, Diebold, and Scotti (2009); OECD (2010) and Bakarić, Tkalec, and Vizek (2016). The idea behind the coincident indicator is that many macroeconomic variables comove with a single unobserved variable called the state of the economy. This index became a useful tool in observing real-time economic performance and movement of output growth (Mariano & Murasawa, 2010). Stock and Watson (1989) examined the business cycle as a comovement of aggregated time series data that coincided with the latent variable—the state of the economy. Stating the business cycle as a latent variable, Aruoba et al. (2009) followed a similar method in constructing a coincident indicator using high-frequency data via the dynamic factor model.

The choice of indicators used for creating coincident indicators remains controversial. The selected variables should have significant relevance, as stated by the OECD (2010). The National Bureau of Economic Research (NBER) constructed a coincident indicator using four leading indicators: industrial production, real personal income less transfer, real manufacturing and trade sales, and employment in non-agriculture (Stock & Watson, 1989). However, the selected variables vary from study to study, for example, in those by Nilsson and Brunet (2006); Albu and Dinu (2009); Guo et al. (2009) and Bakarić et al. (2016), especially in developing countries where data availability is limited. Some studies used a large number of variables, such as those by Altissimo, Cristadoro, Forni, Lippi, and Veronese (2010) and Gupta and Kabundi (2011).

The methodologies used to establish coincident indexes also vary across studies. These include ad hoc procedures and weighted average of aggregate time series data to more complicated methods. One issue in the ad hoc approach is weighting; more important variables should be given higher weights (Freudenberg, 2003). Various approaches can be used for weighting, for example, the regression analysis, correlation coefficient, and dimensional

¹According to the World Bank, the GDP growth rate was around 7.68% on average between 1995-2019.

reduction approaches, such as principal component and factor analysis. On the other hand, the standard methods for producing coincident indicator receive a lot of attention. For example, [Stock and Watson \(1989\)](#) followed the dynamic factor model to integrate aggregated time series data and used the Kalman filter algorithm to estimate the unobserved state, while the parameters were estimated by the maximum likelihood. [Rua and Nunes \(2005\)](#) implemented the band-pass filter and principal component to combine multivariate variables into a single index. The principal component works on the linear combination of multiple variables into a new set of variables that are linear independent. The first principal component captures the largest variation of the original data with the largest eigenvalue, followed by the second and third components, and so on. [Levanon \(2010\)](#) studied the business cycle by using Markov switching to estimate recession probability. [Bujosa, Garcia-Ferrer, Juan, and Martin-Arroyo \(2018\)](#) used linear dynamic harmonic regression (LDHR) based on a spectral approach.

The dynamic factor model receives a lot of attention in modern econometrics, especially in estimating the unobserved variables. The process involves using the Kalman algorithm to estimate the unobserved factors. The Kalman algorithm was first designed for tracking objects indirectly in spacecraft to improve the accuracy of position for navigation purposes. Later, it became popular in the study of economic time series as an application to estimate the unobserved variables, for example, the state of the economy, asset pricing, and permanent income. By expressing the observable variables as a linear function of the unobserved variables and unobservable errors, and the movement of the unobserved variables across time, mainly in autoregressive structures, the Kalman filter algorithm could estimate the unobserved variables with minimum mean squared errors (MSEs), while the unknown parameters were estimated by the maximum likelihood estimation ([Durbin & Koopman, 2012](#); [Shumway & Stoffer, 2017b](#); [Stock & Watson, 1989](#)).

For asymptotic properties, the estimated factors calculated by the dynamic factor model and principal components are consistent with an increase in cross-sectional (i) and time dimensions (T).² [Forni, Hallin, Lippi, and Reichlin \(2000\)](#) showed the asymptotic consistency of the estimated common factors for the dynamic factors model. [Stock and Watson \(2002\)](#) and [Bai \(2003\)](#) derived the asymptotic consistency and normality of estimated common factors and factor loadings using the principal components with serial and cross-sectional correlations in the idiosyncratic noises.³ With an increase in cross-sectional dimensions, [Doz et al. \(2011\)](#) showed the consistency of the estimation of unobserved common factors and factor loadings by a two-step procedure of combining principal components with the Kalman smoother. As an increase in cross-sectional dimension leads to the consistent estimation of the common factors it is common to include as many variables as possible. However, later studies indicated that the cross-sectional dimension does not necessarily have to be large for a consistent estimation. For example, [Caggiano et al. \(2011\)](#) showed that 12 to 22 variables could achieve the best result in extracting common factors, and [Poncela and Ruiz \(2012\)](#) showed that variables did not have to be large to achieve consistency under the Kalman filter. Under the Gaussian assumption, parameters are estimated by the maximum likelihood estimation method. The asymptotic consistency and normality of the estimated parameters in the dynamic factor model were shown by [Caines \(1988\)](#) and [Durbin and Koopman \(2012\)](#). The expectation-maximization (EM) algorithm became a common tool to estimate the parameters in the maximum likelihood estimation. An alternative algorithm, Newton–Raphson, showed a faster convergence rate ([Lindstrom & Bates, 1988](#)). However, in estimating many parameters, the Newton–Raphson algorithm could be unstable in the iteration process unless the initial guesses were close to the true values ([Wilks, 2019](#)).

Although one method is not necessarily superior to another, the dynamic factor model receives more attention for many reasons in the study of common factors. [Rodríguez and Ruiz \(2012\)](#) pointed out that under a Gaussian assumption on idiosyncratic noises with known parameters, the Kalman filter provided the best linear unbiased predictions of the common factors in the context of the linear state-space model. Additionally, the dynamic factor model provided flexible specifications compared to the principal components, such as working with non-stationary datasets and strong correlation of idiosyncratic noises, imposing restrictions, and handling irregular elements and missing datasets ([Poncela & Ruiz, 2012](#)). The Kalman filter could produce the mean squared error in the finite sample, while only asymptotic MSE is available for the principal component. Moreover, the Kalman filter performs better for correlated idiosyncratic noises. Although moderate serial and cross-sectional correlations (0.5) of errors produced a marginal impact on the estimators and forecasting quality, [Stock and Watson \(2002\)](#) showed that strong serial and cross-sectional correlations (0.9) caused a deterioration of the estimators and forecasting quality in the case of the principal component. [Poncela and Ruiz \(2012\)](#) showed that regardless of weak or strong correlations in errors, the Kalman filter could produce the efficient minimum mean squared error when the number of variables was around 30.⁴ [Doz et al. \(2011\)](#) combined the principal components and Kalman filter. This hybrid approach could substantially improve the estimation of the common factors if the common factors were small and persistent ([Giannone, Reichlin, & Small, 2008](#); [Stock & Watson, 2011](#)). Another method was to use the bootstrap procedure proposed by [Rodríguez and Ruiz \(2012\)](#) to improve the predicting MSE of the unobserved variables, which gave a better finite sample property.

Critics stated that both time and cross-sectional domains lack satisfied properties in the finite sample, which lead to a more rigorous study of the dynamic factor model using a small sample. It is worth mentioning that the MSE under the Kalman filter has two sources of uncertainty: one comes from the stochastic process of the filtering, and the other one is from the estimation of the unknown parameters.⁵ This second source came from substituting the consistent parameters when the true values were unknown.⁶ With known parameters and non-persistent serial

² Both i and $T \rightarrow \infty$. In asymptotic property, when the time dimension approaches infinity ($T \rightarrow \infty$), the estimated parameters converges to the population parameters. On the other hand, the cross-sectional dimension ($i \rightarrow \infty$) approaches infinity, so the uncertainty in the extraction procedure will approach zero ([Poncela & Ruiz, 2012](#)).

³ [Choi \(2012\)](#) derived a smaller variance using the generalized principal component estimator without normality assumption (first derived in 2007). However, there is a challenge in finding a well-behaved idiosyncratic error variance matrix that made generalized principal component estimator infeasible, as pointed in [Stock. and Watson \(2011\)](#).

⁴ They found out that as the mean squared error approaches zero in cross-sectional dimensions, the total uncertainty has a U shape because as more variables are included, the number of estimated parameters increase and induces uncertainty.

⁵ As an increase in the cross-sectional dimensions will increase the numbers of parameters to be estimated, the DFM estimation deteriorates.

⁶ This uncertainty accounted for about 5% ($T=100$) of the total uncertainty in the univariate non-stationary one factor model ([Rodríguez & Ruiz, 2012](#)).

correlation in idiosyncratic noises, the filter uncertainty was a non-increasing function in the cross-sectional dimension regardless of weak or strong contemporaneous correlation of noises (Poncela & Ruiz, 2012). For small cross-sectional dimensions, they showed that the uncertainty only slightly increased for a practical purpose, while the ratio of parameter uncertainty to total uncertainty was at a minimum when the number of variables was around ten.⁷

The estimation of the coincident indicator has some challenges and shortcomings. Munda and Nardo (2005) explained the weighting issue of the linear aggregation rule, a weakness that data normalization did not capture in building the coincident index. Another issue involves the stationary assumption in constructing the coincident indicator. With this assumption, it throws away some important information if cointegration exists⁸, and ignoring the long-term relationship has a detrimental effect on forecasting quality (Smeekes & Wijler, 2019). It is worth mentioning that the unobserved state variable can be stationary or non-stationary in the context of the dynamic factor model. For this reason, using the hybrid approach (combining principal components and the Kalman filter or smoother) for non-stationary data may improve the estimation of common factors in the finite sample, as shown by Corona, Poncela, and Ruiz (2020). Peña and Poncela (2004) and Moon and Perron (2007) also worked on non-stationary series in estimating common factors using the dynamic factor model. An additional issue concerns the measurement unit. The process involved using data normalization, or standardization, in combining a group of variables into a single index (Altissimo. et al., 2001; Freudenberg, 2003), and Freudenberg (2003) mentioned various normalization methods. As the coincident indicator is unit-free by its construction, it causes a problem for interpretation. (Mariano & Murasawa, 2003) pointed out a shortcoming of an economic interpretation of the standard coincident index. Lastly, the study implements the Kalman filter using linearity and normality assumptions. Many studies assumed the linear projection of the coincident indicator for simplicity; however, if the function is non-linear, this creates the misspecification of the functional form, so weights would not only be inconsistent, but also biased. With unknown parameters, the model would become non-linear when expressing in the state-space form (Murphy, 2012). Other versions of Kalman filters have been initiated to deal with the non-linear system, for example, the extended Kalman filter (EKF) and unscented Kalman filter (UKF). The intuition of the EKF is linearization of the function using the Taylor series and applies the standard Kalman filter to solve the systems. The performance of the EKF could improve through the iteration process. However, it performed poorly for prior large covariance and function that was highly non-linear near the current mean (Murphy, 2012). A better version is the unscented Kalman filter, proposed by Julier and Uhlmann (1997). The UKF approximates the Gaussian distribution using the unscented transformation by creating several sample points called sigma points.⁹ The UKF became more accurate than the EKF in capturing mean and covariance at least to the second order of any non-linear function (Murphy, 2012). Durbin and Koopman (2012) showed that the normality assumption is not necessary from the minimum variance linear unbiased estimation perspective.

3. Methodology

3.1. Description of Data

With economic relevance to the state of the economy, this study used monthly macroeconomic and banking data from 2010 to mid-2019, to construct the coincident indicator.¹⁰ These data are available on the official website of the Ministry of Economy and Finance and the National Bank of Cambodia. These include total bank credits (Y_1), banks lending to the service-related sectors (Y_2), banks lending to the manufacturing sector (Y_3), banks lending to the retail trade sector (Y_4), banks lending to the wholesale sector (Y_5), electricity supply (Y_6), export value (Y_7), import value (Y_8), corporate income or profit tax (Y_9), domestic value-added tax (Y_{10}), and import value-added tax (Y_{11}). Additionally, the study includes exogenous variables, such as the official exchange rate (Z_1), money supply (Z_2), and interest rate of bank lending (Z_3).

The study treats data in the following manner; first, to reduce the issue of irregular elements, it uses seasonally adjusted data¹¹, next, the study proceeds with data normalization.¹² A standardized, or unit-free, dataset plays a crucial role in combining multiple variables into a single index; otherwise, the weights will be biased (Altissimo. et al., 2001; Freudenberg, 2003). Lastly, it uses non-stationary data to extract the common factor.¹³ Since many macroeconomic variables comove with the state of the economy in the long term, allowing for the common trend is better than ignoring the cointegration. Corona et al. (2020) showed that combining principal components and the Kalman filter to extract the common factors, using the original series could improve the estimation than differencing the series when cointegration exists in the finite sample. In the first step, it used the principal components to estimate the initial unobserved state. The first principal component captured the highest proportion of the variation of the series (about 94% of the total proportion). The use of principal components improved the estimation of common factors substantially in the dynamic factor model, especially for the extraction of small common factors (Giannone et al., 2008; Stock & Watson, 2011).

3.2. Model Specification

The study follows the state-space or dynamic factor model. Estimating the unobserved state follows the two-step procedure proposed by Doz et al. (2011). They showed that this hybrid approach yields consistent estimators

⁷ Poncela and Ruiz (2012) used sample sizes (T) between 100 and 200 in their simulation.

⁸ If cointegration exists between the state and observed variables, the error is stationary. In this case, both the state of the economy (S_t) and observed variables (Y_{it}) are $I(1)$, while the error (u_{it}) is $I(0)$.

⁹ See: Wan and Van Der Merwe (2000): The unscented Kalman filter for nonlinear estimation.

¹⁰ Data are available from 2007 to the second quarter of 2019; however, many missing values may affect the result, so the study only selected the period between 2010 and 2019.

¹¹ For seasonal adjustment, the data used the ARIMA (X-13) method, which is available in E-views.

¹² Data normalization formula: $Y_{i,n} = \frac{Y_{i,o} - \bar{Y}_i}{\delta_i}$; where $Y_{i,n}$ is a new transform variable, $Y_{i,o}$ is the original variable, \bar{Y}_i is the mean, and δ_i is the standard deviation of variable i .

¹³ Trend and seasonality are the main issues of non-stationarity. Series are stationary if their mean, variance, and covariance are constant over time.

of common factors in the dynamic factor model when cross-sectional and time domains approach infinity ($i, T \rightarrow \infty$). With known parameters and mutual independence of the idiosyncratic noises, the estimated factors were unbiased regardless of the number of variables used (Poncela & Ruiz, 2012).

The ARMAX linear state-space model can be written as:

$$Y_t = \rho_t S_t + \alpha_t Z_t + u_t; \tag{1}$$

$$S_{t+1} = \theta_t S_t + v_t; \tag{2}$$

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim (\text{i.i.d}) N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}; \begin{bmatrix} R_t & 0 \\ 0 & Q_t \end{bmatrix} \right)$$

- Y_t is $m \times 1$ the vector of the observed variable.
- S_t is $n \times 1$ the unobserved state variable.
- Z_t is $p \times 1$ the vector of the exogenous variable.
- u_t, v_t are idiosyncratic noises with serially and contemporaneous independences.
- ρ_t is $m \times n$, α_t is $m \times p$, and θ_t is $n \times n$ matrices (where only $\alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{23}, \alpha_{33}, \alpha_{43}, \alpha_{53}, \alpha_{71}$, and α_{81} are non-zero, while the other coefficients are restricted as zero).

The state-space model consists of two types of equations: the observation or signal equation and the transition or state equation. Equation 1 is the observation equation, which explains the relationship of the observed variables as a linear function of the unobserved state. Equation 2 is the transition equation, which captures the movement of the state variable over time. Although the most common form of the transition equation is in the autoregressive (AR) structure, it could also include the autoregressive moving average (ARMA) form.¹⁴ The exogenous variables can enter either the observation or the transition equations without losing any interpretation. Including the exogenous variables improves the model goodness of fit. The study refers to the Akaike Information Criteria (AIC) for model selection of the ARMAX linear state-space structure.¹⁵ Shumway and Stoffer (2017b) showed the consistency of the common factors of the ARMAX linear state-space model.

The study imposed some assumptions. First, S_t and Y_{it} cointegrated, so parameters are estimated by ordinary least squares.¹⁶ Second, u_t and v_t are serial and contemporaneous uncorrelated (mutually independence).¹⁷ Third, for simplicity, the study examines the dynamic factor model in the context of the linear system; for a non-linear system, other versions of the Kalman filter could be implemented, for example, the EKF and UKF. Fourth, the model also assumes the initial mean and variance of the state variable to be Gaussian $S_0 \sim N(S_0^0, P_0^0)$. Without knowledge of the initial value, the study sets the diffuse initial state condition.

Any system of equations that can be expressed in the state-space form can be solved using the Kalman filter. The study uses the Kalman filter to estimate the unobserved state of the economy (S_t).¹⁸ The intuition of the Kalman filter is to update the state from S_t^t to S_{t+1}^t when the observation Y_{t+1} is available. It involves a two steps process of predicting and updating. With the above initial state value, the Kalman filter algorithm for the ARMAX linear state-space model in this study is:¹⁹

$$S_{t+1}^t = \theta_t S_t^t; \tag{3}$$

$$P_{t+1}^t = \theta_t P_t^t \theta_t' + Q_t; \tag{4}$$

$$S_{t+1}^{t+1} = S_{t+1}^t + K_{t+1} \epsilon_{t+1}; \text{ where } [\epsilon_{t+1} = Y_{t+1} - (\rho_{t+1} S_{t+1}^t + \alpha_{t+1} Z_{t+1})]; \tag{5}$$

$$P_{t+1}^{t+1} = (I - K_{t+1} \rho_{t+1}') P_{t+1}^t; \text{ where } I \text{ is the identity matrix}; \tag{6}$$

$$\text{The Kalman gain } (K_{t+1}): K_{t+1} = P_{t+1}^t \rho_{t+1}' (\rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1})^{-1} \tag{7}$$

Alternatively, we can use Kalman smoother to estimate the state variable. The Kalman smoother uses all the observations for updating. For the dynamic factor model in (1) and (2), the process of updating the state variable via the Kalman smoother is:

$$S_t^n = S_t^t + J_t (S_{t+1}^n - S_{t+1}^t); \tag{8}$$

$$P_t^n = P_t^t + J_t (P_{t+1}^n - P_{t+1}^t) J_t'; \tag{9}$$

$$\text{Where } J_t = P_t^t \theta_t' (P_{t+1}^t)^{-1}; \tag{10}$$

Deriving these equations is based on a Shumway and Stoffer (2017b) textbook on the state-space model and is shown in Appendix A. This study assumes the independence of idiosyncratic noises. In the case of correlated noises, it generates a quite different result, but it does not affect the updating of the Kalman filter and smoother in equations 5, 6, 8, and 9.²⁰

To implement the Kalman filter it replaces parameters ρ_t , in the system with their consistent estimators $\hat{\rho}_t$. Using the state variable generated by the first principal component, the study estimated the initial weights by the ordinary least squares and feasible generalized least squares. For the cointegrated series, the OLS gave the consistently estimated parameters, but the inference did not hold (Stock, 1987). For non-cointegrated series, the study used the FGLS to estimate the parameters, as suggested by Wu, You, and Zou (2016). The remaining parameters were estimated by maximum likelihood estimation in the dynamic factor model. Let $\varphi = \{Q_t, \alpha_t, R_t\}$ refer to the vector of the parameters to be estimated with a known initial state $S_0 \sim N(S_0^0, P_0^0)$, where idiosyncratic noises, u_t and v_t , are serially and contemporaneous independents. The likelihood is calculated from the innovations $\epsilon_1, \epsilon_2, \dots, \epsilon_t$.

$$\epsilon_t = Y_t - (\rho_t S_t^{t-1} + \alpha_t Z_t); \epsilon_t \sim N(0, \Sigma_t); \tag{11}$$

¹⁴ The ARMA linear state-space models vary across studies depending on the interest of authors.

¹⁵ The MARSS package in R program allow for a flexibility of adding the exogenous vector into the state space model.

¹⁶ This is for cointegrated series. If series are not cointegrated with the state variable, their residuals, u_{it} , are not $I(0)$. In this case, instead of using OLS, the study uses FGLS to estimate the parameters.

¹⁷ This is for simplicity. The $\text{Cov}(u_t, v_t) \neq 0$ could be the case, but it did not affect the updating process (Shumway & Stoffer, 2017b).

¹⁸ Kalman smoother can be implemented to estimate the state variable as well. To estimate the unobserved state, S_t , using data $Y_{1:s} = \{Y_1, Y_2, \dots, Y_s\}$, the process is called filtering when $s = t$ while it is called smoothing for $s > t$ (Shumway & Stoffer, 2017b).

¹⁹ The notation of $S_t^s = E(S_t | Y_s)$; $P_{t_1, t_2}^s = E\{(S_{t_1} - S_{t_1}^s)(S_{t_2} - S_{t_2}^s)' | Y_s\}$. For $t_1 = t_2$, it uses the notation P_t^s .

²⁰ See: Shumway and Stoffer (2017b) for the case of correlated noises.

Ignoring the constant, the log-likelihood of $\log L(\varphi)$ is:

$$\log L(\varphi) = -\frac{1}{2} \sum_{t=1}^n \log |\Sigma_t(\varphi)| - \frac{1}{2} \sum_{t=1}^n \epsilon_t(\varphi)' \Sigma_t(\varphi)^{-1} \epsilon_t(\varphi); \quad (12)$$

Asymptotic properties of consistency and normality of estimators hold in general (Shumway & Stoffer, 2017b).

4. Result of the Study

Table 1 summarizes the statistical properties of the series. Additionally, it uses the Augmented Dickey–Fuller (ADF) test to check the stationarity of individual series. All series are $I(1)$. In the case of cointegration between the state and observed variables, using the level series may improve the estimation of common factors, as pointed out by Corona et al. (2020).

4.1. State Estimation

4.1.1. Initial Coefficient Estimation

The increase in cross-sectional dimension induces the number of parameters to be estimated in the system that causes the dynamic factor model to be less feasible in practice, especially for the finite sample. Doz et al. (2011) came up with the idea of replacing the parameters with their consistent estimators. With the stationary assumption, they estimated the parameters by the OLS method using the state variable produced by the principal component. This method improved the estimation of the common factor in the dynamic factor model.

Empirical studies revealed that many macroeconomic variables cointegrate with the state of the economy. Ignoring the cointegration will throw away a large amount of information. For this reason, this study used the non-stationary series. Table 2 shows the result of the OLS of each series on the state variable and their residual tests. Figure 3 plots the residuals of this result (see Appendix B). The OLS results show that some series cointegrated with the state variable. For the cointegrated series, the OLS parameter estimation holds, although its inference is not valid. For series that are not cointegrated, the OLS estimation is spurious, so the study refers to the FGLS to estimate the parameters.

4.1.2. Dynamic Factor Model (ARMAX)

So far, the study has not indicated a specific form of ARMAX linear state-space model. Using the state variable generated by the principal components, it constructed the state equation of the dynamic factor model in autoregressive (AR) form. Table 3 shows various lag selection criteria. The Akaike information criterion (AIC) suggested the AR (4) model for the state equation. For the non-stationary autoregressive model, the asymptotic distribution of AIC held, while the Bayesian information criterion (BIC) was weakly consistent (Tsay, 1984). For the observation equation, the study introduced two lags of the state variable. The study controlled for exogenous variables to improve the model fitness.²¹ Moreover, it replaced the parameters ρ_t with $\hat{\rho}_t$ estimated by the OLS and FGLS.

The study used a diffuse initial state condition. Additionally, it restricts the variances of the idiosyncratic noises of both observed and state equations to be non-negative. Table 4 shows the results of the state-space model. Figure 4 shows the movement of the state estimated by the Kalman filter and smoother together with confidence intervals as well as its residual movement. Figure 5 shows the distribution of the estimated residuals of both state and observation equations. The disturbance of the state equation behaves like a normal distribution. On the other hand, some disturbances of the observation equations fail to meet the Gaussian assumption. Durbin and Koopman (2012) showed that even without the normality assumption, from the minimum variance linear unbiased estimation approach, the estimation of state variables (S_{t+1} and S_t) and their variances (P_{t+1} and P_t) were the same as the estimates from the classical and Bayesian viewpoints.

Figure 1 summarized the result of the estimation of the Cambodian economic condition. Panel (a) shows the estimation of the state variable by various methods (principal component, Kalman filter, and Kalman smoother), and Panel (b) shows their comovement with GDP growth rate.²² All methods tended to capture well when the economy performs below its potential level. All methods revealed a similar pattern that the economy performed below the average level during 2016 and 2017. The economy somehow recovered after mid-2017.

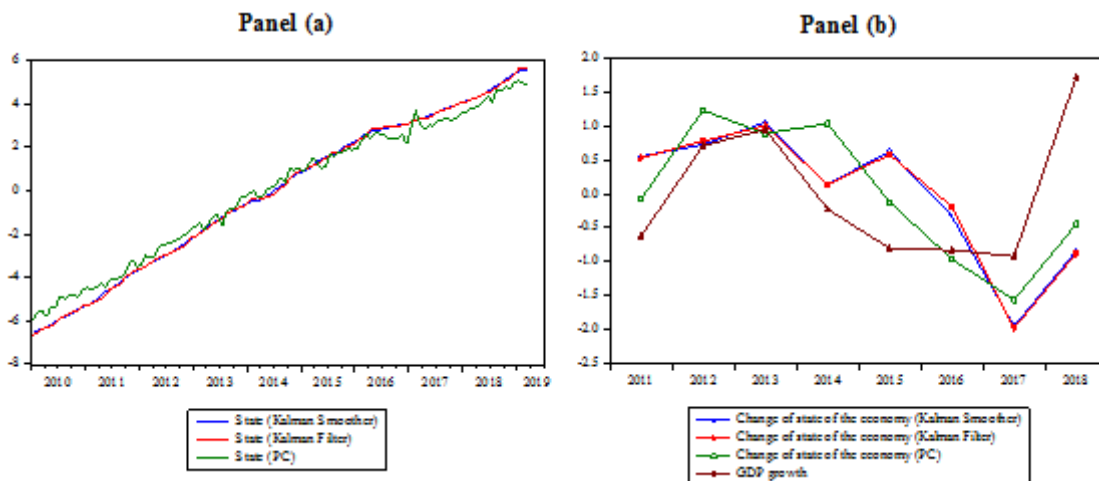


Figure-1. State estimation by Kalman smoother, Kalman filter, and principal components and their comovement with GDP.

²¹ The AIC value of the model with exogenous variables is -1.940, while it is 3.129 for the model without controlling exogenous variables.

²² The monthly movement of coincident indicators produced by each methodology are aggregated into annual data.

4.2. Foreign Direct Investment Inflow and the State of the Economy

Empirical studies on the relations between foreign direct investment and economic growth have been long discussed. Some studies found the impact of FDI on economic growth, e.g. De Mello (1999) and Devajit (2012), while another revealed economic growth as a factor of FDI inflow, e.g. Roy and Mandal (2012). Additionally, Srinivasan, Kalaivani, and Ibrahim (2010) and Hossain and Hossain (2012) discovered cointegration between the two variables. Türkcan, Duman, and Yetkiner (2008) pointed out the simultaneous causation between FDI and economic growth, and Alfaro (2003) studied the heterogeneity across sectors on the relations between FDI and growth. This study skipped this discussion.

The study explored the impact of the economic condition on the FDI inflow by exploiting the coincident index using the ARDL model.²³ Bevan and Estrin (2000) indicated macroeconomic variables, such as growth, inflation, and exchange rate risk, as determinants of FDI inflow to a transitional economy. Pan (2003) examined the determinants of FDI inflow for a country-specific study.

Table 5 shows the dynamic relationship between the state of the economy and FDI inflow. It revealed that the state of the economy had a positive impact on FDI inflow in the short-term. All of the models using the state variables produced by the Kalman smoother, Kalman filter, and principal component show a similar tendency. The coefficient of the output gap shows the same pattern.²⁴ Additionally, the study revealed a cointegration between the state of the economy and FDI inflow; the long-term coefficient turns negative. One explanation of this negative impact is the diminishing marginal product of capital. From the supply side, capital investment becomes less attractive, which deters investors. Additionally, the coefficients of exogenous variables are well-behaved even though most of them are not significant, and other variables are significant. For example, inflation shows a negative effect on FDI inflow. High inflation indicates a higher cost of investment that is often associated with a country's risks. Interest rate shows a positive impact on FDI inflow. From the supply side, an increase in interest rate attracts capital inflow as a return on lending. However, interpretation of the impact of interest rate on FDI inflow is still uncertain. One reason is that the high rate of return often relates to a country's high risks of investment, especially for a small open economy. The effect of trade is diverse across the models, which could be due to the inclusion of the short lags.

5. Discussion and Limitation of the Study

The construction of the coincident index in this study had some limitations. First, it estimated the unobserved state in the linear context. Many macroeconomic variables comove in a non-linear form. Murphy (2012) indicated that by putting the system into a state-space form, parameters were no longer linear, even though the true model was. Other extensions of the Kalman filter dealing with the non-linear structure are the extended Kalman filter (EKF) and the unscented Kalman filter (UKF). György, Kelemen, and Dávid (2014) discussed how each algorithm works. Julier and Uhlmann (1997) introduced the UKF as a superior version to the EKF. The UKF is a derivative-free filter that does not need to calculate the Jacobian compared to the EKF. Both the EKF and UKF approximate the distribution with the Gaussian assumption. Another type of application, called a particle filter, could also be used for the non-linear system. Unlike the EKF and UKF, the particle filter does not require the Gaussian assumption.

The second limitation of this study involves the implementation of time-invariant parameters in the dynamic factor model. Stability of the factor loadings (coefficients of the state variables) may not appropriate if the economy goes through a structural change. The structural break will cause time-variate parameters. Bates, Plagborg-Møller, Stock, and Watson (2013) categorized conditions that the standard estimation of factors could tolerate temporal parameter instability. Stock and Watson (2002) showed that the estimated factors under principal components were consistent with small time-variate parameters. The structural break may be less of an issue because the period in this study is relatively short.

Another issue relates to the quality of the data and how the real economy performs. The theory depends mainly on generating data to verify the results. However, real-time data is subject to measurement error. One potential challenge in this study is the availability and quality of data. As a developing country, Cambodia faces a constraint in collecting data from the informal sectors. Looking at this issue, how well a coincident indicator could capture the state of the economy is unknown. Although the quality of data in constructing the CI remains a topic for discussion, its creation plays an important role to monitor economic activity. The CI is a useful tool for early warning and signaling when the economy moves into a recession. The effect may come with long or short lags depending on the characteristics of the economy. Policymakers should use this index with caution.

Lastly, there are two main problems regarding the ARDL model in this study. One of them involves the cointegration issue. Pesaran and Shin (1998) mentioned that the cointegration in the ARDL model must be unique; in other words, there should not be a cointegration among regressors. In restricting cointegration among regressors, this study introduced the first differencing method to transform all I(1) regressors, except the state variable. Another issue relates to the serial correlation and endogeneity problems. Türkcan et al. (2008) suggested an endogenous relationship between FDI inflow and economic growth. It transpired that the ARDL model can address serial correlation and endogeneity problems, and can resolve these issues by adding appropriate lag regressors, for example, ARDL (p, q) to ARDL (p, m), for $m \geq q$ Pesaran and Shin (1998). The study conducted a residual diagnostic by checking serial correlation using the Lagrange Multiplier test. It does not find serial correlation problems in this study.

6. Conclusion

This study constructed a coincident indicator as an unobserved state of the economy using the two-step procedure proposed by Doz et al. (2011). In the first step, it estimated the parameters by the OLS and FGLS

²³ The study uses the quarterly data because only quarterly FDI is available.

²⁴ Output gap is calculated as the deviation between the state of the economy and its trend, where the Hodrick–Prescott (HP) filter is used to decompose the trend.

methods using the state variable generated by the principal component. In the second step, it estimated the unobserved state via the dynamic factor model (DFM) by substituting parameters with its estimators in the first step. This approach could improve the estimation of the common factors substantially as emphasized by Giannone et al. (2008) and Stock and Watson (2011). The study used non-stationary data, as it discards a large amount of information if cointegration exists. Corona et al. (2020) showed that in the case of cointegration, using the original series could improve the estimation of state variables compared to differencing the series. The coincident indicator showed a slower economy performance (below its potential level) during 2016 and 2017. The economy somehow rebounded after mid-2017. Additionally, the study observed the comovement between FDI inflow and the state of the economy. It discovered a positive impact of FDI inflow on the state of the economy in the short-term. Models using the state variable produced by various methods show a similar result, while the long-term coefficient becomes negative. One reason for this is due to a diminishing marginal rate of capital.

References

- Albu, L., & Dinu, V. (2009). how deep and how long could be the recession in romania. *The Amfiteatru Economic Journal*, 11(Special 3), 675-683.
- Alfaro, L. (2003). Foreign direct investment and growth: Does the sector matter. *Harvard Business School*, 2003, 1-31.
- Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., & Veronese, G. (2010). New eurocoin: Tracking economic growth in real time. *The review of Economics and Statistics*, 92(4), 1024-1034. Available at: https://doi.org/10.1162/rest_a_00045.
- Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Lippi, M., Reichlin, L., & Veronese, G. (2001). EuroCOIN: A real time coincident indicator of the euro area business cycle. CEPR Discussion Papers No. 3108.
- Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, 27(4), 417-427. Available at: <https://doi.org/10.1198/jbes.2009.07205>.
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1), 135-171. Available at: <https://doi.org/10.1111/1468-0262.00392>.
- Bakarić, I. R., Tkalec, M., & Vizek, M. (2016). Constructing a composite coincident indicator for a post-transition country. *Economic RESEARCH*, 29(1), 434-445. Available at: <https://doi.org/10.1080/1331677x.2016.1174388>.
- Bates, B. J., Plagborg-Møller, M., Stock, J. H., & Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177(2), 289-304.
- Bevan, A. A., & Estrin, S. (2000). The determinants of foreign direct investment in transition economies. William Davidson Institute Working Paper 342.
- Bujosa, M., Garcia-Ferrer, A., Juan, A. D., & Martin-Arroyo, A. (2018). Evaluating early warning and coincident indicators of business cycles using smooth trends. *Journal of Forecasting*, 39(1), 1-17.
- Caggiano, G., Kapetanios, G., & Labhard, V. (2011). Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK. *Journal of Forecasting*, 30(8), 736-752. Available at: <https://doi.org/10.1002/for.1208>.
- Caines, P. E. (1988). Maximum likelihood estimation of Guassian ARMAX and state-space system. In P. E. Caines, Linear stochastic systems (pp. 411-480). New York: Society for Industrial and Applied Mathematics (SIAM).
- Choi, I. (2012). Efficient estimation of factor models. *Econometric Theory*, 28(2), 274-308. Available at: <https://doi.org/10.1017/s0266466611000338>.
- Corona, F., Poncela, P., & Ruiz, E. (2020). Estimating non-stationary common factors: Implications for risk sharing. *Computational Economics*, 55(1), 37-60. Available at: <https://doi.org/10.1007/s10614-018-9875-9>.
- De Mello, L. R. (1999). Foreign direct investment-led growth: Evidence from time series and panel data. *Oxford Economic Papers*, 51(1), 133-151. Available at: <https://doi.org/10.1093/oep/51.1.133>.
- Devajit, M. (2012). Impact of foreign direct investment on Indian economy. *Research Journal of Management Sciences*, 1(2), 29-31.
- Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164(1), 188-205. Available at: <https://doi.org/10.1016/j.jeconom.2011.02.012>.
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (2nd ed.). United Kingdom: Oxford University Press.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82(4), 540-554. Available at: <https://doi.org/10.1162/003465300559037>.
- Freudenberg, M. (2003). *Composite indicators of country performance: A critical assessment*. Paris: OECD Publishing.
- Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665-676. Available at: <https://doi.org/10.1016/j.jmoneco.2008.05.010>.
- Guo, F., Ozyildirim, A., & Zarnowitz, V. (2009). On the measurement and analysis of aggregate economic activity for China: The coincident economic indicators approach. *China Economic Journal*, 2(2), 159-186. Available at: <https://doi.org/10.1080/17538960903083475>.
- Gupta, R., & Kabundi, A. (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting*, 27(4), 1076-1088. Available at: <https://doi.org/10.1016/j.ijforecast.2010.10.001>.
- György, K., Kelemen, A., & Dávid, L. (2014). Unscented Kalman filters and Particle Filter methods for nonlinear state estimation. *Procedia Technology*, 12(1), 65-74. Available at: <https://doi.org/10.1016/j.protcy.2013.12.457>.
- Hossain, A., & Hossain, M. K. (2012). Empirical relationship between foreign direct investment and economic output in South Asian countries: A study on Bangladesh, Pakistan and India. *International Business Research*, 5(1), 9-21.
- IMF. (2018). Cambodia: 2018 Article IV consultation—press release; staff report; staff statement; and statement by the executive director for Cambodia. International Monetary Fund. Retrieved from: <https://www.imf.org/en/Publications/CR/Issues/2018/12/17/Cambodia-2018-Article-IV-Consultation-Press-Release-Staff-Report-Staff-Statement-and-46478>.
- Julier, S. J., & Uhlmann, J. K. (1997). New extension of the Kalman filter to nonlinear systems. In Signal processing, sensor fusion, and target recognition. *International Society for Optics and Photonics*, 3068, 182-193. Available at: <https://doi.org/10.1117/12.280797>.
- Kim, C.-J., & Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics*, 80(2), 188-201. Available at: <https://doi.org/10.1162/003465398557447>.
- Levanon, G. (2010). Evaluating and comparing leading and coincident economic indicators. *Business Economics*, 45(1), 16-27. Available at: <https://doi.org/10.1057/be.2009.29>.
- Lindstrom, M. J., & Bates, D. M. (1988). Newton—Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, 83(404), 1014-1022. Available at: <https://doi.org/10.2307/2290128>.
- Mariano, R. S., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4), 427-443. Available at: <https://doi.org/10.1002/jae.695>.
- Mariano, R. S., & Murasawa, Y. (2010). A coincident index, common factors, and monthly real GDP. *Oxford Bulletin of Economics and Statistics*, 72(1), 27-46.
- Moon, H. R., & Perron, B. (2007). An empirical analysis of nonstationarity in a panel of interest rates with factors. *Journal of Applied Econometrics*, 22(2), 383-400. Available at: <https://doi.org/10.1002/jae.931>.
- Munda, G., & Nardo, M. (2005). *Constructing consistent composite indicators: The issue of weights*. Ispra, Italy: Institute for the Protection and Security of the Citizen, Joint Research Centre.
- Murphy, K. P. (2012). State space models. In K. P. Murphy, Machine learning: A probabilistic perspective (Vol. The MIT Press, pp. 631-660): Cambridge.
- Nakata, T., & Tonetti, C. (2010). *Kalman filter and kalman smoother*: New York University.

- Nilsson, R., & Brunet, O. (2006). Composite leading indicators for major OECD non-member economies: Brazil, China, India, Indonesia, Russian Federation, South Africa. OECD. OECD Statistics Working Papers. Retrieved from: <http://dx.doi.org/10.1787/834716666802>.
- OECD. (2010). Asian business cycles quarterly. Retrieved from: <http://www.oecd.org/dev/asia-pacific/45623832.pdf>. 1-5.
- Pan, Y. (2003). The inflow of foreign direct investment to China: The impact of country-specific factors. *Journal of Business Research*, 56(10), 829-833.
- Peña, D., & Poncela, P. (2004). Forecasting with nonstationary dynamic factor models. *Journal of Econometrics*, 119(2), 291-321.
- Pesaran, M. H., & Shin, Y. (1998). An autoregressive distributed-lag modelling approach to cointegration analysis. In S. Strom (Ed.), *Econometrics and Economic Theory in the 20th Century* (pp. 371-413). Cambridge: Cambridge University Press.
- Poncela, P., & Ruiz, E. (2012). More is not always better: Back to the Kalman filter in dynamic factor models. Universidad Carlos III de Madrid. Madrid: UC3M Working Paper.
- Rodríguez, A., & Ruiz, E. (2012). Bootstrap prediction mean squared errors of unobserved states based on the Kalman filter with estimated parameters. *Computational Statistics & Data Analysis*, 56(1), 62-74. Available at: <https://doi.org/10.1016/j.csda.2011.07.010>.
- Roy, S., & Mandal, K. K. (2012). Foreign direct investment and economic growth: An analysis for selected Asian countries. *Journal of Business Studies Quarterly*, 4(1), 15-24.
- Rua, A., & Nunes, L. C. (2005). Coincident and leading indicators for the euro area: A frequency band approach. *International Journal of Forecasting*, 21(3), 503-523. Available at: <https://doi.org/10.1016/j.ijforecast.2005.02.001>.
- Shumway, R. H., & Stoffer, D. S. (2017b). State space models. In *Time series analysis and its applications: With R examples* (4th ed., pp. 324-411). New York: Springer.
- Smeeke, S., & Wijler, E. (2019). High-dimensional forecasting in the presence of unit roots and cointegration. Maastricht University, Department of Quantitative Economics Working Paper.
- Srinivasan, P., Kalaivani, M., & Ibrahim, P. (2010). FDI and economic growth in the ASEAN countries: Evidence from cointegration approach and causality test. *IUP Journal of Management Research*, 9(1), 38-63.
- Stock, J. H. (1987). Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica: Journal of the Econometric Society*, 55(5), 1035-1056.
- Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. NBER Macroeconomics Annual (Vol. 4). Cambridge: MIT Press.
- Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460), 1167-1179.
- Stock, J. H., & Watson, M. (2011). *Dynamic factor models*. In M. P. Clements, & D. F. Hendry (Eds.), *Oxford handbook on economic forecasting*. Oxford: Oxford University Press.
- Tsay, R. S. (1984). Order selection in nonstationary autoregressive models. *The Annals of Statistics*, 12(4), 1425-1433. Available at: <https://doi.org/10.1214/aos/1176346801>.
- Türkcan, B., Duman, A., & Yetkiner, I. H. (2008). *How does FDI and economic growth affect each other? The OECD case*. Paper presented at the In International Conference on Emerging Economic Issues in a Globalizing World.
- Wan, E. A., & Van Der Merwe, R. (2000). *The unscented kalman filter for nonlinear estimation*. Paper presented at the In Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373).
- Wilks, D. S. (2019). Parametric probability distributions. In D. S. Wilks, *Statistical methods in the atmospheric sciences* (4th ed., pp. 77-141). Massachusetts: Elsevier Inc.
- World Bank. (2018). The World Bank in Cambodia. Retrieved from: <https://www.worldbank.org/en/country/cambodia/overview>.
- Wu, M., You, P., & Zou, N. (2016). On spurious regressions with trending variables. *Cornell University Library arXiv:1606.05049*.
- Wyplosz, C. (2005). Fiscal policy: Institution versus rules. *National Institute Economic Review*, 191(1), 64-78.
- Zarnowitz, V., & Moore, G. H. (1983). Sequential signals of recession and recovery. In G. H. Moore, *Business cycles, inflation, and forecasting* (2nd ed., pp. 23 - 60). Ballinger.

Appendix A

Given data $Y_t = \{y_1, \dots, y_t\}$, estimation of the unobserved state variables (S_t) by DFM can be done using the Kalman filter or smoother. Shumway and Stoffer (2017b) showed that variance produced by the Kalman smoother is lower than the Kalman filter.

ARMAX linear state-space model with initial condition $S_0 \sim N(S_0^0, P_0^0)$ can be expressed by Equations 1 and 2.²⁵ Additionally, for simplicity, it assumes the errors, u_t and v_t to be independent from each other.

$$Y_t = \rho_t S_t + \alpha_t Z_t + u_t; \quad (1)$$

$$S_{t+1} = \theta_t S_t + v_t; \quad (2)$$

where $\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim (\text{i. i. d}) N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_t & 0 \\ 0 & Q_t \end{bmatrix} \right)$

Because the sum of Gaussian distributions is a Gaussian, it follows that S_{t+1} and Y_t are also Gaussians. Besides, for two Gaussian distributions with mean, variance, and covariance specified below, the conditional expectation is:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

$$x_1 | x_2 \sim N(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (x_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$$

1. Kalman filter

Notation $S_{t+1}^t = E(S_{t+1} | Y_t)$; $P_{t+1}^t = E\{(S_{t+1} - S_{t+1}^t)(S_{t+1} - S_{t+1}^t)' | Y_t\}$

From (2): $S_{t+1}^t = E(\theta_t S_t + v_t | Y_t)$

$$= \theta_t S_t^t; \quad (3)$$

$$\begin{aligned} P_{t+1}^t &= E\{(S_{t+1} - S_{t+1}^t)(S_{t+1} - S_{t+1}^t)' | Y_t\} \\ &= E\{(\theta_t(S_t - S_t^t) + v_t)(\theta_t(S_t - S_t^t) + v_t)' | Y_t\} \\ &= \theta_t P_t^t \theta_t' + Q_t; \end{aligned} \quad (4)$$

Let $\epsilon_{t+1} = Y_{t+1} - E(Y_{t+1} | Y_t) = Y_{t+1} - (\rho_{t+1} S_{t+1}^t + \alpha_{t+1} Z_{t+1})$; Where $E(\epsilon_{t+1}) = 0$

$$\text{Var}(\epsilon_{t+1}) = \text{Var}[\rho_{t+1}(S_{t+1} - S_{t+1}^t) + u_{t+1}] = \rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1}$$

Under the Gaussian assumption above, $E(\epsilon_t Y_s') = 0$ for $s < t$

$$\begin{aligned} \text{Cov}(S_{t+1}, \epsilon_{t+1} | Y_t) &= \text{Cov}(S_{t+1}, Y_{t+1} - (\rho_{t+1} S_{t+1}^t + \alpha_{t+1} Z_{t+1}) | Y_t) \\ &= \text{Cov}(S_{t+1} - S_{t+1}^t, Y_{t+1} - (\rho_{t+1} S_{t+1}^t + \alpha_{t+1} Z_{t+1}) | Y_t) \end{aligned}$$

²⁵ A similar version of this derivation can be found in the literature such as in Nakata and Tonetti (2010).

$$\begin{aligned}
 &= \text{Cov}[S_{t+1} - S_{t+1}^t, \rho_{t+1}(S_{t+1} - S_{t+1}^t) + u_{t+1}] \\
 &= P_{t+1}^t \rho'_{t+1} \\
 &\text{So, the joint distribution between } S_{t+1} \text{ and } \epsilon_{t+1} \text{ conditions on } Y_t. \\
 &\begin{pmatrix} S_{t+1} \\ \epsilon_{t+1} \end{pmatrix} | Y_t \sim N \left(\begin{bmatrix} S_{t+1}^t \\ 0 \end{bmatrix}, \begin{bmatrix} P_{t+1}^t & P_{t+1}^t \rho'_{t+1} \\ \rho_{t+1} P_{t+1}^t & \rho_{t+1} P_{t+1}^t \rho'_{t+1} + R_{t+1} \end{bmatrix} \right) \\
 &S_{t+1}^{t+1} = E(S_{t+1} | Y_1, \dots, Y_t, Y_{t+1}) = E(S_{t+1} | Y_t, \epsilon_{t+1}) \\
 &S_{t+1}^{t+1} = S_{t+1}^t + P_{t+1}^t \rho'_{t+1} (\rho_{t+1} P_{t+1}^t \rho'_{t+1} + R_{t+1})^{-1} \epsilon_{t+1} \\
 &\text{Let } K_{t+1} \text{ (Kalman gain)} = P_{t+1}^t \rho'_{t+1} (\rho_{t+1} P_{t+1}^t \rho'_{t+1} + R_{t+1})^{-1}; \tag{7} \\
 &S_{t+1}^{t+1} = S_{t+1}^t + K_{t+1} \epsilon_{t+1}; \tag{5} \\
 &P_{t+1}^{t+1} = \text{Cov}(S_{t+1} | Y_t, \epsilon_{t+1}) = P_{t+1}^t - P_{t+1}^t \rho'_{t+1} (\rho_{t+1} P_{t+1}^t \rho'_{t+1} + R_{t+1})^{-1} \rho_{t+1} P_{t+1}^t \\
 &= [I - P_{t+1}^t \rho'_{t+1} (\rho_{t+1} P_{t+1}^t \rho'_{t+1} + R_{t+1})^{-1} \rho_{t+1}] P_{t+1}^t \\
 &P_{t+1}^{t+1} = (I - K_{t+1} \rho_{t+1}) P_{t+1}^t; \tag{6}
 \end{aligned}$$

2. Kalman Smoother

The Kalman smoother used all the observations (n) to update the state variable. The joint distribution of S_t and S_{t+1} conditions on Y_t is:

$$\begin{aligned}
 &\begin{pmatrix} S_t \\ S_{t+1} \end{pmatrix} | Y_t \sim N \left(\begin{bmatrix} S_t^t \\ S_{t+1}^t \end{bmatrix}, \begin{bmatrix} P_t^t & P_t^t \theta_t' \\ \theta_t P_t^t & P_{t+1}^t \end{bmatrix} \right) \\
 &E(S_t | S_{t+1}, Y_t) = S_t^t + P_t^t \theta_t' (P_{t+1}^t)^{-1} (S_{t+1} - S_{t+1}^t) \\
 &\text{Var}(S_t | S_{t+1}, Y_t) = P_t^t - P_t^t \theta_t' (P_{t+1}^t)^{-1} \theta_t P_t^t \\
 &\text{Let } J_t = P_t^t \theta_t' (P_{t+1}^t)^{-1}; \tag{10}
 \end{aligned}$$

For (n) total samples, by the law of iterated expectation:

$$\begin{aligned}
 &S_t^n = E(S_t | Y_n) = E(E(S_t | S_{t+1}, Y_n) | Y_n) = E(E(S_t | S_{t+1}, Y_t) | Y_n); \text{ For } n > t \\
 &= E(S_t^t + J_t(S_{t+1} - S_{t+1}^t) | Y_n) \\
 &S_t^n = S_t^t + J_t(S_{t+1}^n - S_{t+1}^t); \tag{8}
 \end{aligned}$$

$$\begin{aligned}
 &P_t^n = E(S_t - S_t^n)(S_t - S_t^n)' \\
 &\text{From (8): } S_t - S_t^n = S_t - S_t^t - J_t(S_{t+1}^n - S_{t+1}^t) \\
 &S_t - S_t^n + J_t S_{t+1}^n = S_t - S_t^t + J_t S_{t+1}^t \\
 &\text{Multiply both sides by its transpose and take expectation, we get} \\
 &E[(S_t - S_t^n + J_t S_{t+1}^n)(S_t - S_t^n + J_t S_{t+1}^n)'] = E[(S_t - S_t^t + J_t S_{t+1}^t)(S_t - S_t^t + J_t S_{t+1}^t)'] \\
 &\text{Because cross-product terms are zero, so} \\
 &P_t^n + J_t E(S_{t+1}^n S_{t+1}^n)' J_t' = P_t^t + J_t \theta_t E(S_t^t S_t^t)' \theta_t' J_t' \\
 &E(S_{t+1}^n S_{t+1}^n)' = E(S_{t+1} S_{t+1}') - P_{t+1}^n = \theta_t E(S_t S_t)' \theta_t' + Q_t - P_{t+1}^n \\
 &E(S_t^t S_t^t)' = E(S_t S_t)' - P_t^t \\
 &P_t^n + J_t [\theta_t E(S_t S_t)' \theta_t' + Q_t - P_{t+1}^n] J_t' = P_t^t + J_t \theta_t [E(S_t S_t)' - P_t^t] \theta_t' J_t' \\
 &P_t^n = P_t^t + J_t P_{t+1}^n J_t' - [J_t (\theta_t P_t^t \theta_t' + Q_t) J_t'] \\
 &P_t^n = P_t^t + J_t P_{t+1}^n J_t' - J_t P_{t+1}^t J_t' \\
 &P_t^n = P_t^t + J_t (P_{t+1}^n - P_{t+1}^t) J_t'; \tag{9}
 \end{aligned}$$

Appendix B:

Tables and Figures

Table 1. Summary statistics.

Variables	Mean	SD	Unit Root test (Level)		Unit Root test (1 st difference)	
			ADF	P-Value	ADF	P-Value
Total bank credits (Y₁)	17.246	0.717	-0.410	0.903	-14.049	0.000
Banks lending to the service-related sectors (Y₂)	7.652	0.396	0.763	0.993	-12.542	0.000
Banks lending to the manufacturing sector (Y₃)	7.885	0.503	-3.017	0.057	-13.758	0.000
Banks lending to the retail trade sector (Y₄)	8.609	0.640	-0.633	0.858	-13.001	0.000
Banks lending to the wholesale sector (Y₅)	8.539	0.608	-2.312	0.170	-14.933	0.000
Total electricity supply (Y₆)	5.796	0.588	0.350	0.980	-12.247	0.000
Export Value (Y₇)	19.611	0.798	-0.781	0.822	-21.410	0.000
Import Value (Y₈)	14.252	0.961	-1.130	0.704	-9.796	0.000
Corporate income or profit tax (Y₉)	11.746	0.588	-0.194	0.935	-8.347	0.000
Domestic VAT (Y₁₀)	11.521	0.432	-0.412	0.902	-8.870	0.000
Import VAT (Y₁₁)	12.029	0.452	-0.299	0.920	-9.615	0.000

Table 2. Cointegration results (OLS result of individual series on state variables).

Variables	Coefficient	Residual's unit root test	
		t-statistics	P-value
Total bank credits (Y ₁)	0.302*** (0.002)	-7.808	0.000
Banks lending to the service-related sectors (Y ₂)	0.292*** (0.006)	-1.812	0.373
Banks lending to the manufacturing sector (Y ₃)	0.298*** (0.008)	-1.204	0.671
Banks lending to the retail trade sector (Y ₄)	0.300*** (0.004)	-1.810	0.374
Banks lending to the wholesale sector (Y ₅)	0.297*** (0.008)	-1.240	0.655
Total electricity supply (Y ₆)	0.307*** (0.004)	-2.489	0.121
Export Value (Y ₇)	0.297*** (0.007)	-4.715	0.000
Import Value (Y ₈)	0.309*** (0.006)	-2.697	0.078
Corporate income or profit tax (Y ₉)	0.289*** (0.008)	-10.034	0.000
Domestic VAT (Y ₁₀)	0.276*** (0.013)	-5.155	0.000
Import VAT (Y ₁₁)	0.287*** (0.008)	-4.715	0.000

Note: Standard errors are reported in parenthesis.
*, **, and *** are significant at 10%, 5%, and 1%, respectively.

Table 3. Lag selection criteria of the state variable.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-257.299	NA	8.825	5.016	5.041	5.026
1	-5.118	494.568	0.067	0.138	0.189	0.159
2	0.084	10.102	0.062	0.057	0.133	0.088
3	7.206	13.689*	0.055	-0.062	0.040*	-0.021
4	8.953	3.325	0.054*	-0.076*	0.051	-0.024*
5	8.970	0.032	0.055	-0.058	0.096	0.004

Note: *suggests the lag selection.

Table 4. DFM results.

Parameter Name	Coef.	SD	z-Statistic	P-Value
α_{11}	-0.002	0.004	-0.578	0.563
α_{12}	0.109	0.012	8.718	0.000
α_{13}	-0.029	0.013	-2.194	0.028
α_{23}	0.151	0.025	6.116	0.000
α_{33}	0.129	0.133	0.972	0.331
α_{43}	-0.034	0.053	-0.641	0.522
α_{53}	0.327	0.213	1.536	0.125
α_{71}	-0.029	0.024	-1.244	0.214
α_{81}	0.047	0.032	1.438	0.150
Var($\mathbf{v}_t = \mathbf{e}^{c12}$)	-7.873	0.365	-21.584	0.000
Var($\mathbf{u}_{1t} = \mathbf{e}^{c1}$)	-8.388	0.170	-49.402	0.000
Var($\mathbf{u}_{2t} = \mathbf{e}^{c2}$)	-3.650	0.225	-16.218	0.000
Var($\mathbf{u}_{3t} = \mathbf{e}^{c3}$)	-1.889	0.393	-4.807	0.000
Var($\mathbf{u}_{4t} = \mathbf{e}^{c4}$)	-4.020	0.284	-14.146	0.000
Var($\mathbf{u}_{5t} = \mathbf{e}^{c5}$)	-1.360	0.504	-2.698	0.007
Var($\mathbf{u}_{6t} = \mathbf{e}^{c6}$)	-3.965	0.195	-20.284	0.000
Var($\mathbf{u}_{7t} = \mathbf{e}^{c7}$)	-2.851	0.163	-17.513	0.000
Var($\mathbf{u}_{8t} = \mathbf{e}^{c8}$)	-2.710	0.133	-20.382	0.000
Var($\mathbf{u}_{9t} = \mathbf{e}^{c9}$)	-2.293	0.129	-17.770	0.000
Var($\mathbf{u}_{10t} = \mathbf{e}^{c10}$)	-1.207	0.113	-10.662	0.000
Var($\mathbf{u}_{11t} = \mathbf{e}^{c11}$)	-2.318	0.113	-20.467	0.000
θ_1	2.693	0.021	129.587	0.000
θ_2	-3.257	0.028	-115.488	0.000
θ_3	2.420	0.008	295.654	0.000
θ_4	-0.856	0.005	-160.035	0.000

Note: DFM estimation.

State variable name	Final State	Root MSE	z-Statistic	Prob.
SV1	6.088	0.183	33.292	0.000
SV2	6.022	0.151	39.841	0.000
SV3	5.922	0.124	47.703	0.000
SV4	5.787	0.098	59.142	0.000
Log likelihood	135.599	Akaike info criterion		-1.940
Parameters	25	Schwarz criterion		-1.340
Diffuse priors	0	Hannan-Quinn criteria.		-1.697

Note: DFM estimation.

Table 5. The impact of the state of the economy on FDI inflow (Conditional error correction form and bound test).

Dependent variable Δ FDI	State (Kalman smoother)		State (Kalman filter)		State (Principal component)	
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
FDI (-1)	-1.201***	-1.522***	-1.099***	-1.221***	-0.888***	-0.600**
	(0.246)	(0.335)	(0.237)	(0.289)	(0.213)	(0.216)
State (-1)	-11.761***	-17.756***	-10.073***	-13.657**	-6.932***	-4.686*
	(3.696)	(5.353)	(3.466)	(4.687)	(2.238)	(2.540)
Δ State	-734.797**	-1022.636**	-645.575**	-910.541**	-482.953**	-494.055*
	(311.017)	(448.914)	(293.555)	(398.822)	(231.853)	(240.514)
Δ State (-1)	825.674**	1163.456**	712.135**	974.933**	503.326**	445.233*
	(323.556)	(463.364)	(304.871)	(413.627)	(228.825)	(231.869)
Δ State (-2)	1.375	2.046	1.488	2.731*	0.960	1.808**
	(1.369)	(2.095)	(1.222)	(1.453)	(0.786)	(0.715)
Δ State (-3)	0.466	0.503	0.758	1.111	0.528	0.945*
	(1.010)	(1.151)	(0.910)	(0.936)	(0.490)	(0.452)
Δ Exchange rate		-0.027		-0.133		-0.159
		(0.153)		(0.112)		(0.112)
Δ Expenditure		-0.032		-0.043		-0.024
		(0.067)		(0.048)		(0.046)
Inflation		-0.027		-0.033		-0.056**
		(0.037)		(0.027)		(0.022)
Inflation (-1)		-0.064*		-0.061**		-0.005
		(0.031)		(0.021)		(0.020)
Δ Revenue (-1)		0.114		0.141		0.029
		(0.130)		(0.092)		(0.101)
Δ Interest rate		0.021		0.059		0.240*
		(0.177)		(0.121)		(0.119)
Δ Size of trade		-0.070		-0.067		-0.337
		(0.312)		(0.213)		(0.193)
Δ Size of trade (-1)		-0.940**		-0.672*		0.139
		(0.416)		(0.309)		(0.351)
Δ Size of trade (-2)		-0.732		-0.277		0.865**
		(0.443)		(0.383)		(0.362)
Δ Size of trade (-3)		-0.209		-0.008		0.569**
		(0.361)		(0.235)		(0.211)
Output gap	736.209**	1025.025**	647.716**	912.868**	484.548**	496.239*
	(310.759)	(448.680)	(293.343)	(398.615)	(231.612)	(240.306)
Output gap (-1)	-1549.478**	-2169.146**	-1349.155**	-1873.795**	-980.082**	-936.142*
	(630.894)	(906.397)	(594.872)	(807.153)	(458.646)	(469.490)
Output gap (-2)	824.530**	1161.452**	710.029**	971.596**	501.853**	441.851*
	(323.331)	(462.318)	(304.397)	(412.465)	(228.108)	(231.175)
Election (dummy)	0.047	-0.012	0.026	-0.006	0.041	0.128*
	(0.075)	(0.094)	(0.064)	(0.068)	(0.059)	(0.058)
State of the economy (dummy)	-0.078	-0.071	-0.018	0.026	-0.032	-0.094
	(0.102)	(0.123)	(0.092)	(0.089)	(0.075)	(0.073)
Trend	1.272***	1.900***	1.083***	1.419**	0.737***	0.421
	(0.371)	(0.540)	(0.347)	(0.472)	(0.215)	(0.253)
C	-33.059***	-49.644***	-27.217***	-33.666**	-16.266***	-4.152
	(9.446)	(13.726)	(8.697)	(11.964)	(4.717)	(6.377)

R²	0.648	0.858	0.723	0.920	0.751	0.937
Adjusted R²	0.437	0.547	0.557	0.743	0.602	0.797
Prob (F-statistic)	0.013	0.050	0.002	0.005	0.001	0.002
Obs.	37	37	37	37	37	37

Note: Standard errors are in parentheses. Data are normalized with mean 0 and standard deviation 1.
 *, **, and *** are significant at 10%, 5%, and 1%, respectively.

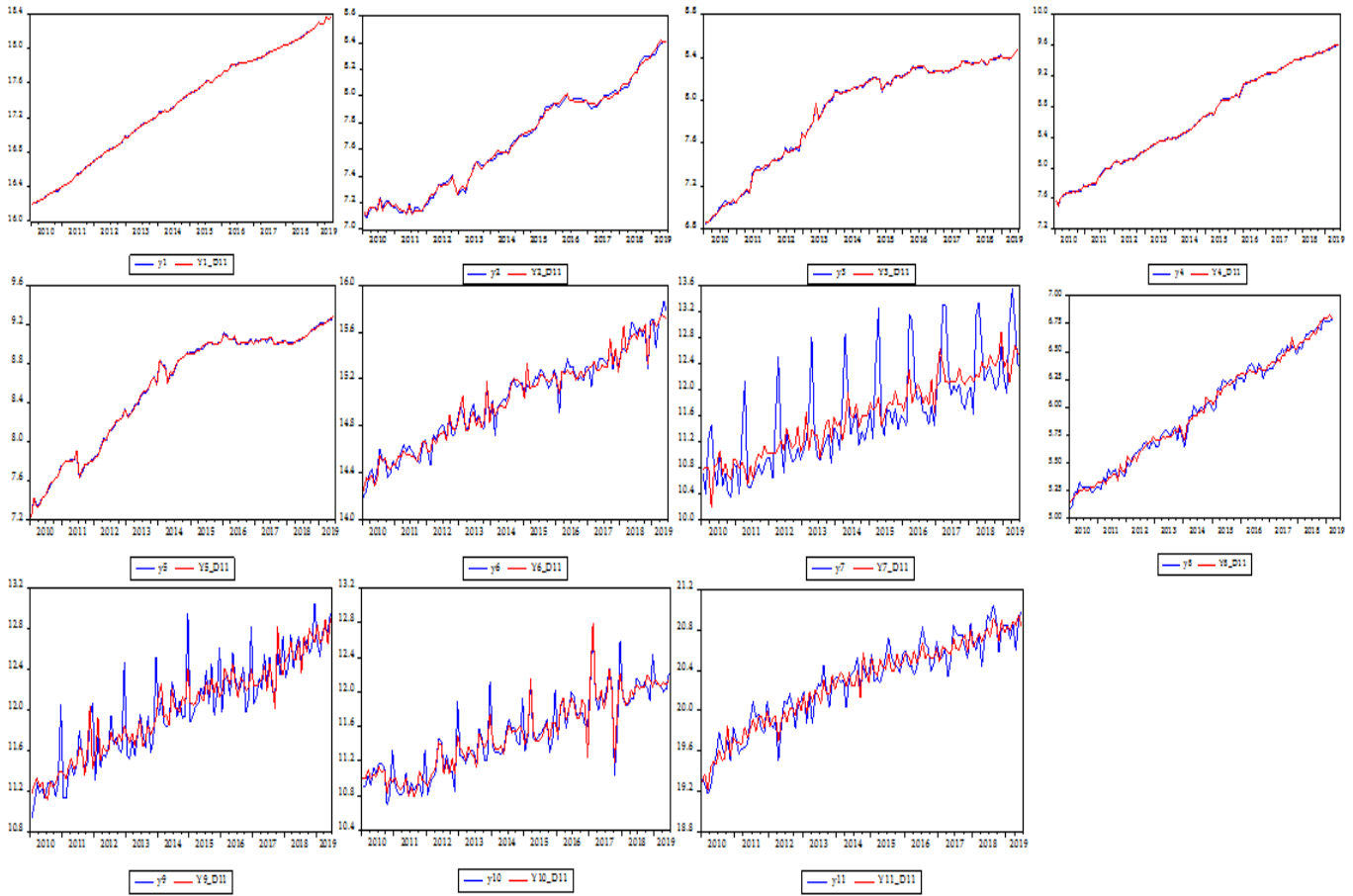


Figure 2. Seasonal and non-seasonal adjusted data.

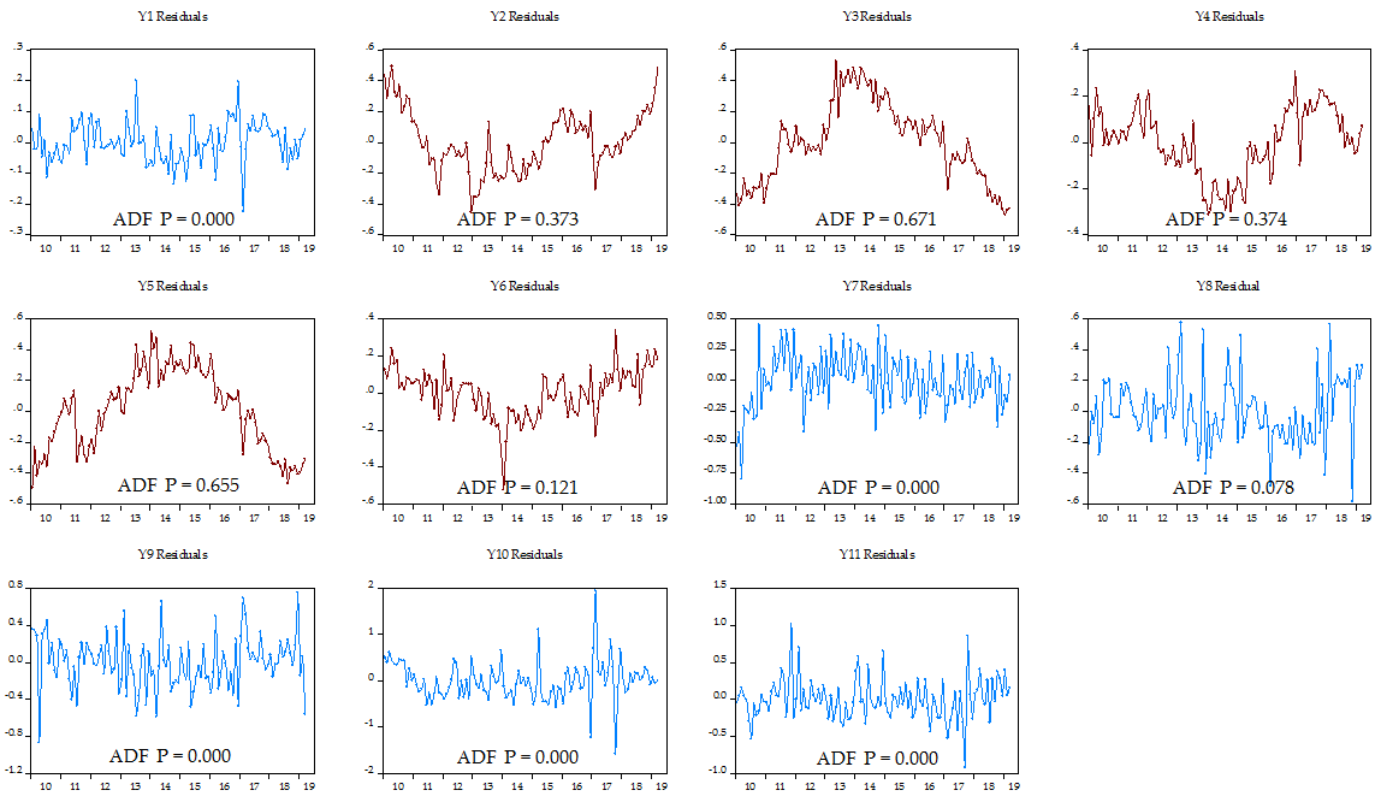


Figure 3. Residuals of OLS regression of each series I(1) on the state variable I(1).

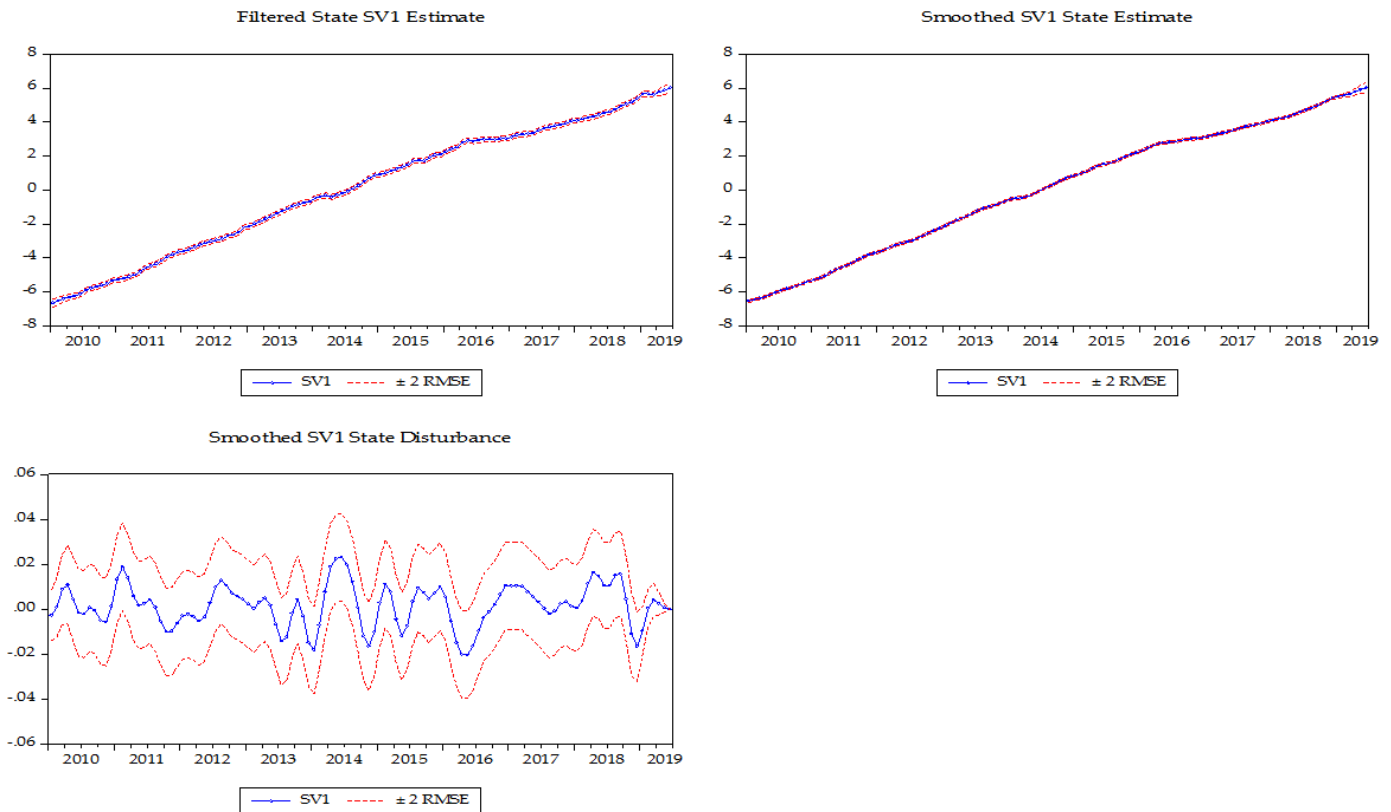


Figure 4. State estimation by Kalman smoother and Kalman filter.

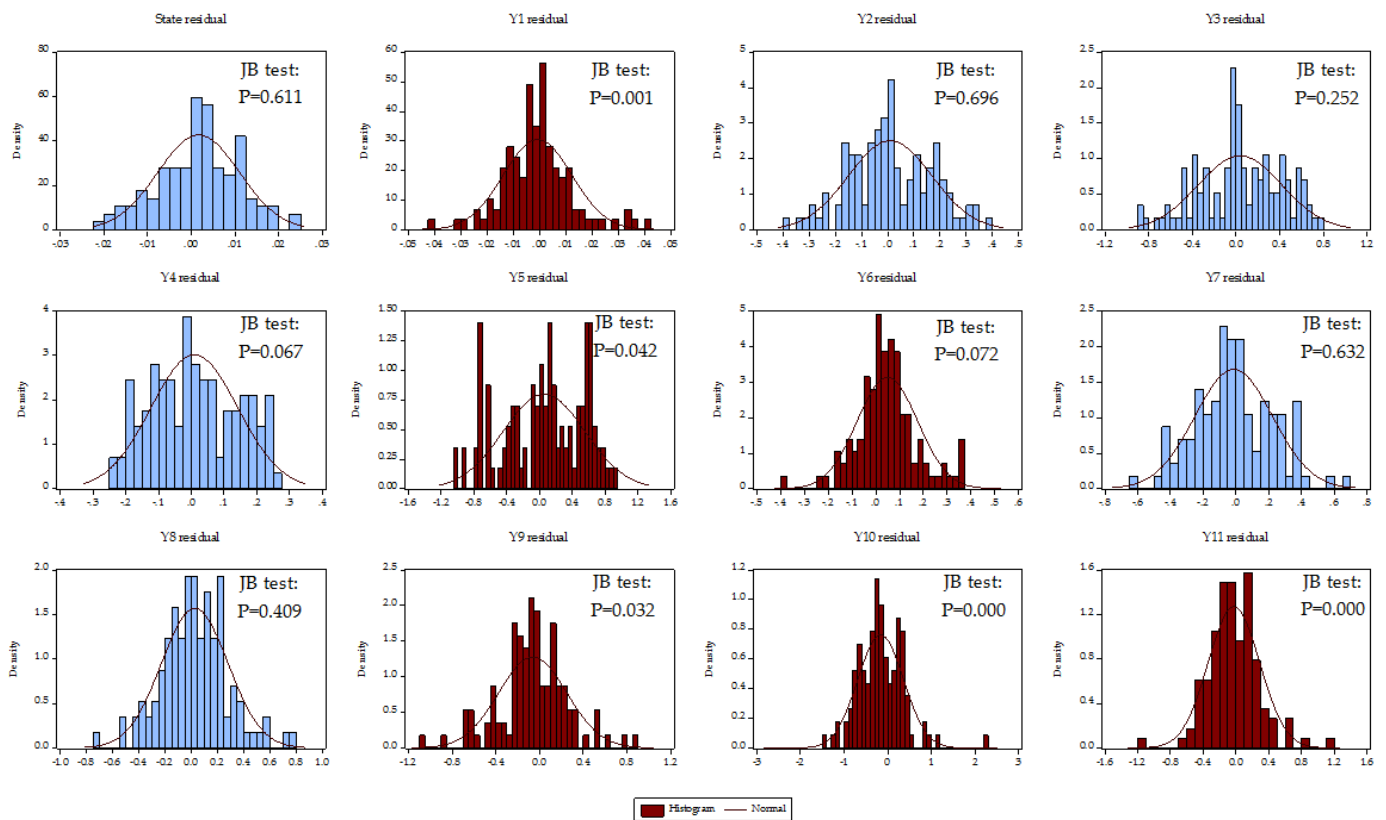


Figure 5. Residual diagnostics.

Asian Online Journal Publishing Group is not responsible or answerable for any loss, damage or liability, etc. caused in relation to/arising out of the use of the content. Any queries should be directed to the corresponding author of the article.