

The University of Tokyo, Graduate School of Public Policy

Tokyo, June 2020

Estimating Cambodia Economic Condition by Dynamic Factor Model

SA Kimleng (ID 51-188215)

Supervisor: Assoc. Prof. Ueda Kenichi

Master Thesis, Master of Public Policy

Stream: Economic Policy, Finance and Development

This thesis was written as a part of the Master of Public Policy, International Program, at GraSPP. Please note that neither the institution nor the examiners are responsible, through the approval of this thesis, for the theories and methods used, or results and conclusions drawn in this work.

Abstract

The study constructs a coincident indicator (CI), the unobserved state of the economy, for Cambodia by combining the principal component (PC) and the dynamic factor model (DFM). In the first step, it estimates the factor loadings, coefficients of the unobserved state variable, by the Ordinary Least Squares (OLS) and Feasible Generalized Least Squares (FGLS) methods using the state variable produced by the PC. In the second step, it estimates the unobserved state variable through the DFM by replacing the coefficients with their consistent estimators in the first step. Doz, Giannone, & Reichlin (2011) introduced this hybrid approach. This method could improve the estimation of the state variable substantially as mentioned in Stock and Watson (2011). The coincident indicator shows that the economy fell below its potential level between 2016 and 2017 and started recovering after mid-2017.

By exploiting the coincident index, the study examines comovement between the government fiscal expenditure and the state of the economy as well as how foreign direct investment (FDI) inflow comoves with the state of the economy using the Autoregressive Distributed Lags (ARDL) model. The study finds out evidence of procyclical fiscal position by looking at the relations between the output gap and government expenditure. Moreover, by allowing for different slope, it reveals a steeper slope when the economy is in a bad position compared to when it is in good shape. For the FDI aspect, the result shows an acceleration of the state of the economy contributes to an increase in FDI inflow in short-run. The long-run coefficient turns negative. A reason could be due to the diminishing marginal product of capital that makes foreign capital investment becomes less attractive.

Contents

I. Introduction.....	6
1. Background of the study	6
2. Research objectives.....	6
3. Research questions.....	7
II. Literature review	7
III. Methodology	10
1. Description of data.....	10
2. Treatment of the data	11
3. Model specification.....	11
IV. Result of the study	13
1. Descriptive statistics	13
2. State estimation.....	13
2.1. Initial coefficients estimation	13
2.2. Dynamic factor model (ARMAX).....	14
3. Fiscal policy and state of the economy	15
4. Foreign direct investment inflow and the state of the economy	16
V. Discussion and limitation of the study.....	17
VI. Conclusion	18
References	19
Appendices.....	23

List of Tables

Table 1 Summary statistics	26
Table 2 Cointegration result (OLS result of individual series on state variable).....	26
Table 3 Lag selection criteria of the state variable	27
Table 4 DFM result (parameter estimation by MLE)	27
Table 5 The impact of the state of the economy on the fiscal stance (Conditional error correction form and bound test).....	28
Table 6 The impact of the state of the economy on the fiscal stance (Different slope estimation).....	29
Table 7 The impact of the state of the economy on FDI inflow (Conditional error correction form and bound test)	30

List of Figures

Figure 1 State estimation by Kalman smoother, Kalman filter, and PC and its comovement to GDP.....	15
Figure 2 Seasonal and non-seasonal adjusted data	32
Figure 3 Residuals of OLS regression of each series I(1) on the state variable I(1)	33
Figure 4 State estimation by Kalman smoother and Kalman filter.....	34
Figure 5 Residual diagnostics	35
Figure 6 Government expenditure and output gap	36
Figure 7 FDI and state of the economy.....	37

I. Introduction

1. Background of the study

The early warning system plays a significant role in monitoring economic activity in Cambodia. Despite high economic growth, the economy faced vulnerability to shocks (World Bank, 2018).¹ IMF (2018) identified the export and fiscal shocks as well as the contingency liability as major risks that could slow down Cambodia's economy. Without an early warning signal, the countercyclical policy may be plausible and sometimes even push the economy into a worsening position. Wyplosz (2005) emphasized the use of fiscal policy as a countercyclical policy could be inefficient and possibly did more harm than good. For these reasons, a good design of an early warning system could lower the magnitude of shocks and improve the monitoring capacity of economic activity and signalling when a recession comes.

Monitoring economic activity, mainly observing the business cycle, requires a high-frequency indicator that represents the cyclical movement of the economy. GDP would be a suitable candidate, as mentioned in OECD (2010). However, monthly and quarterly GDP is not available in Cambodia's context, mainly due to resource and time constraints for compiling such data. For this reason, there is a need to create a similar indicator that closely correlates to GDP to monitor the economic condition as well as to observe the business cycle.

Like GDP, the coincident indicator (CI) could be a useful tool in the study of the business cycle. The CI, unobserved state of the economy, is constructed by exploiting its relationship with high-frequency observable variables. It compiles high-frequency data into a single index that captures the current state of the economy and real-time economic performance. OECD (2010) used macroeconomic variables that hold both economic and statistical relevances to quarterly real GDP to establish the coincident indicator. By proper design, it could help policymakers in surveillant activity and monitoring the real sector through signalling when the economy moves into recession. Additionally, it could be a useful tool for policymakers to deal with unemployment and inflation issues during the peak and trough (Zarnowitz & Moore, 1983). It also became a useful indicator in the analysis of short-term macroeconomic dynamic (Guo, Ozyildirim, & Zarnowitz, 2009).

Although coincident index helps monitor economic performance, there are some challenges to be considered. One concern involves the number of variables to be used for its construction. Caggiano, Kapetanios, & Labhard (2011) suggested that using many variables was not always a better solution. Moreover, how well this index can signal the recession is unknown. This concerns when the real economy will respond or how long it will take effect. OECD (2010) suggested that CI signalled five to six months earlier before the economy went into recession. Besides, the use of CI should be cautious and requires improvement when additional information is available. Despite its limitation, there is no doubt that the coincident index plays a crucial role for policymakers in monitoring the economic condition.

2. Research objectives

There are three main objectives of this study. First, it constructs a coincident indicator (CI) to estimate Cambodia's economic condition by the dynamic factor model (DFM). As mentioned early, this index helps policymakers to improve surveillant activity. Second objective concerns on the fiscal position. While economic theories suggested a country should pursue a countercyclical fiscal policy, empirical studies overwhelmingly showed a procyclical fiscal position in many developing countries. This study explores the comovement between government expenditure and output gap by exploiting the state of the economy. This could help policymakers to reevaluate their fiscal stance. Lastly, it examines the comovement of foreign direct investment (FDI) inflow and the state of the economy. Large capital mobility could put

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

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 one tends to get out of the country when the economy moves into a bad situation, for example, a recession. Large capital mobility could put the balance of payment at risk due to exchange rate depreciation, especially for net debtor country.

3. Research questions

- How is the current state of Cambodia's economic performance?
- How does Cambodia's fiscal position correlate with the state of the economy?
- How does FDI inflow comove with the state of the economy?

This paper has the following structures: Section II reviews the empirical literature of the coincident index and the dynamic factor model (DFM). Part III indicates the methodology, the state-space or dynamic factor model, in establishing the coincident indicator for Cambodia. Part IV shows the results of the study. Part V discusses some limitations of the study. The last section makes the concluding remarks.

II. Literature review

The coincident index (CI), unobserved state of the economy, has often been used in the study of the business cycle, for example, in Stock & Watson (1989), Kim & Nelson (1998), Altissimo, et al. (2001), Aruoba, Diebold, & Scotti (2009), OECD (2010), Rasic, Tkale, & Vizek (2016). The intuitive behind the CI is that many macroeconomic variables comove with a single unobserved variable, 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 & Watson (1989) examined the business cycle as a comovement of aggregated time-series data that coincided with a latent variable, the state of the economy. Stating the business cycle as a latent variable, Aruoba, Diebold, & Scotti (2009) followed a similar method in constructing a CI using high-frequency data via the dynamic factor model (DFM).

The choice of indicators used for creating the coincident indicator remains controversial. The selected variables should have significant relevancy as stated in OECD (2010). The National Bureau of Economic Research (NBER) constructed a CI using four leading indicators: the industrial production, real personal income less transfer, real manufacturing and trade sales, and employment in the non-agriculture (Stock & Watson, 1989). However, the chosen variables vary across studies, for example, in Nilsson & Brunet (2006), Albu & Dinu (2009), Guo, Ozyildirim, & Zarnowitz (2009), and Rasic, Tkale, & Vizek (2016) especially in developing countries where data availability is limited. Some studies used a large number of variables, for example, Altissimo et al. (2010), and Gupta & Kabundi (2011).

The methodologies in establishing coincident index also differ across studies. These include an Adhoc procedure, weighted average of aggregate time-series data, to more complicated methods. An issue in the Adhoc approach is weighting. More important variables should get higher weights (Freudenberg, 2003). Various approaches can be used for weighting such as regression analysis, correlation coefficient, and dimensional reduction approaches such as principal component and factor analysis. On the other hand, standard methods for producing CI receive much attention. For example, Stock & Watson (1989) followed the dynamic factor model (DFM) to integrate aggregated time-series data and used the Kalman filter algorithm to estimate the unobserved state while parameters were estimated by maximum likelihood estimation. Rua & Nunes (2005) implemented the band-pass filter and principal component (PC) to combine multivariate variables into a single index. The PC works on a linear combination of multiple variables into a new set of variables that are linear independence. 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 in estimating recession probability. Bujosa, Garcia-Ferrer, Juan, & Martin-Arroyo (2018) used Linear Dynamic Harmonic Regression (LDHR) based on the spectral approach. They constructed CI using log-transformation without adjusting seasonality for forecasting purpose.

The Dynamic Factor Model (DFM) receives much 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 indirect tracking objects in spacecraft to improve the accuracy of position or navigation purposes. Later, it became popular in economic time-series study as an application to estimate the unobserved variables, for example, the state of the economy, asset pricing, and permanent income hypothesis. By expressing the observable variables as a linear function of the unobserved variables and unobservable error, and the movement of the unobserved variables across time, mainly in the autoregressive structure, Kalman filter algorithm could estimate the unobserved variables with minimum mean squared error (MSE) while parameters were estimated by maximum likelihood estimation (Stock & Watson, 1989; Durbin & Koopman, 2012; Shumway & Stoffer, 2017).

For asymptotic property, the estimated factors by DFM and PC are consistent with an increase in cross-sectional and time dimensions.² Forni, Hallin, Lippi, & Reichlin (2000) showed the asymptotic consistency of the estimated common factors of the dynamic factor model. Stock & Watson (2002), Bai & Ng (2002), and Bai (2003) derived asymptotic consistency and normality of estimated common factors and factor loadings using PC with serial and cross-sectional correlations in the idiosyncratic noises.³ With an increase in cross-sectional dimensions, Doz, Giannone, & Reichlin (2011) showed the consistency of the estimate unobserved common factors and factor loadings by the two-step procedure of combining the principal component and Kalman smoother. As the 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 is not necessarily large for a consistent estimation. For instance, Caggiano, Kapetanios, & Labhard (2011) showed that 12 to 22 variables could achieve the best result in extracting common factors. Poncela & Ruiz (2012) showed that variables did not have to be large to achieve consistency under the Kalman filter. Besides, with the Gaussian assumption, the parameters are estimated by the maximum likelihood estimation. The asymptotic consistency and normality of estimated parameters of the DFM were shown in Caines (1988 p. 426), and Durbin & Koopman (2012).⁴ The Expectation-Maximization (EM) algorithm becomes a common tool in estimating 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 iterative process unless the initial guesses were close to the true values (Wilks, 2019, p. 128).

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 that the uncertainty in extracted procedure will approach to zero (Poncela & Ruiz, 2012).

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

⁴ By consistency, the $\text{plim}_{T \rightarrow \infty}(\hat{\varphi}) = \varphi$.

Although one method is not necessarily superior to another, DFM receives more attention for many reasons in the study of common factors. Rodríguez & Ruiz (2012) pointed out that under the Gaussian assumption of idiosyncratic noises with known parameters, Kalman filter provided best linear unbiased predictions of the common factors in the context of the linear state-space model. Additionally, DFM provided flexible specifications compared to the principal component such as working with non-stationary datasets and strong correlation of idiosyncratic noises, imposing restrictions, and handling irregular elements as well as missing observations (Poncela & Ruiz, 2012). The Kalman filter could produce MSE 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 & Watson (2002) found out that strong serial and cross-sectional correlations (0.9) caused a deterioration of the estimators and forecasting quality using the principal component. Poncela & Ruiz (2012) showed that no matter weak or strong correlations in errors, Kalman filter could produce the efficient minimum MSE when the number of variables was around 30.⁵ Doz, Giannone, & Reichlin (2011) combined PC and Kalman filter. This hybrid approach could substantially improve the estimation of common factors if they were small and persistent (Giannone, Reichlin, & Small, 2008; Stock & Watson, 2011). Another method was to use the bootstrap approach proposed by Rodríguez & Ruiz (2012) to improve the predicted MSE of the unobserved variables, and it also gave a better finite sample property.

Critics that both time and cross-sectional domains lack satisfied property in the finite sample lead to a more rigorous study of DFM in a small sample. It is worth to mention that the MSE under Kalman filter has two sources of uncertainties: one comes from a stochastic process of the filtering, and the other one is from an estimation of the unknown parameters.⁶ The second source came from substituting the consistent parameters when true values were unknown.⁷ With known parameters and non-persistent serial correlation in idiosyncratic noises, the filter uncertainty was a non-increasing function in the cross-sectional dimension no matter 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 10.⁸

The estimation of the coincident indicator comes with some challenges and shortcomings. Munda & Nardo (2005) explained the weighting issue of linear aggregation rule, a weakness that data normalization did not capture in making the coincident index. Another issue involves the stationary assumption in constructing the CI. With this assumption, it throws away some important information if cointegration exists.⁹ Ignoring the long-run relationship generated a detrimental effect on forecasting quality (Smeekes & Wijler, 2019). It is worth to mention that the unobserved state variable can be stationary or non-stationary in DFM context. For this reason, using the hybrid approach (combining PC and Kalman filter or smoother) for non-stationary data may improve the estimation of common factors in the finite sample as shown in Corona, Poncela, & Ruiz (2020). Peña & Poncela (2004), and Moon & Perron (2007)

⁵ They found out that as MSE approaches zero in cross-sectional dimensions, the total uncertainty has a U shape because as more variables include, the number of estimated parameters increase, and this induces the uncertainty.

⁶ As an increase in cross-sectional dimensions will increase the numbers of parameters to be estimated, the DFM performance deteriorates.

⁷ This uncertainty accounted for about 5% ($T=100$) of total uncertainty in univariate non-stationary one factor model (Rodríguez & Ruiz, 2012).

⁸ Poncela & Ruiz (2012) used the sample sizes (T) between 100-200 in their simulation.

⁹ If cointegration exists between the state and observed variables, the error is stationary. In this case, both state (S_t) and observed variables (Y_{it}) are $I(1)$ while the error (u_{it}) is $I(0)$.

worked on a non-stationary series in estimating common factors using DFM. An additional issue concerns on measurement unit. The process involved using data normalization or standardization to combine a group of variables into a single index (Altissimo, et al., 2001; Freudenberg, 2003). Freudenberg (2003) mentioned various normalization methods. However, as CI is unit free by its construction, it causes a problem on the interpretation. Mariano & Murasawa (2003) pointed out a shortcoming on the economic interpretation of the standard coincident index. Lastly, the study implements the Kalman filter using the linearity and normality assumptions. Many studies assumed the linear projection of CI for simplicity; however, if the function is non-linear, this creates the misspecification of the functional form, so weights would be not only inconsistent but also bias. With unknown parameters, the model would become non-linear when expressing in the state-space form (Murphy, 2012, p. 647).

For this reason, other versions of Kalman filters have been initiated to deal with the non-linear context, for example, the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). The intuitive of EKF is linearization the function using the Taylor series and applies the standard Kalman filter to solve the system. The performance of EKF could improve through the iterative process. However, it performed poorly for large prior 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 & Uhlmann (1997). The UKF approximates the Gaussian distribution using the unscented transformation by creating several sample points called sigma points.¹⁰ UKF became more accurate than EKF in capturing mean and covariance at least to the second order of any non-linear function (Murphy, 2012).

The next section describes the treatment of data and model specification in establishing a coincident indicator for Cambodia using the state-space or dynamic factor model.

III. Methodology

1. Description of data

With economic relevancy to the state of the economy, the study uses monthly macroeconomic and banking data, from 2010 to mid-2019, to construct the coincident indicators.¹¹ These data are available on the official websites of the Ministry of Economy and Finance and National Bank of Cambodia. They consist of both stock and flow values. The study creates a coincident indicator by using 11 observed variables. These include:

- Total bank credits (Y_1)
- Bank lending to the service-related sector (Y_2)
- Bank lending to the manufacturing sector (Y_3)
- Bank lending to the retail trade sector (Y_4)
- Bank 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})
- Import value-added tax (Y_{11})

Additionally, the study controls for exogenous variables such as:

- Official exchange rate (Z_1)
- Broad money supply (Z_2)
- Interest rate of bank lending (Z_3)

¹⁰ See: Wan & Van Der Merwe (2000): The Unscented Kalman Filter for nonlinear estimation.

¹¹ Data are available from 2007 to first quarter of 2019; however, many missing observations may affect the result, so the study selects only the period between 2010-2019.

2. Treatment of the data

The study treats the data in the following ways. First, to reduce the issue of irregular elements, it uses seasonally adjusted data.¹² Next, it proceeds with data normalization. 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).

$$Y_{i,new} = \frac{Y_{i,old} - \bar{Y}_i}{\delta_{Y_i}}$$

The study uses non-stationary data in extracting common factor. Series are stationary if their mean, variance, and covariance are constant over time.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, Poncela, & Ruiz (2020) showed that combining PC and Kalman filter to extract the common factors, using the original series could improve the estimation than differencing the series when cointegration exists in a finite sample. In the first step, it uses PC to estimate the initial unobserved state. The first principal component captures the highest proportion of the variation of the series (about 94% of the total proportion).

$$Y_{it} = \beta_i S_{t,pc} + \varepsilon_{it}, \text{ for } i = 1, 2, \dots, 11$$

Y_{it} are the observed variables. $S_{t,pc}$ is the common factor, the state of the economy produced by the principal component. β_i are factor loadings of the first principal component.¹⁴ The subscripts i and t represent cross-sectional and time domains, respectively. The use of PC improved the estimation of common factors substantially in the DFM, especially for small common factors extraction (Giannone, Reichlin, & Small, 2008; Stock & Watson, 2011).

3. Model specification

The study follows the state-space or dynamic factor model. Estimating the unobserved state follows the two-step procedure proposed by Doz, Giannone, & Reichlin (2011).¹⁵ They showed that this hybrid approach yields consistent estimators of common factors in DFM 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:

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

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

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

- Y_t is $m \times 1$ vector of the observed variable
- S_t is $n \times 1$ the unobserved state of the economy
- Z_t is $p \times 1$ vector of the exogenous variable
- u_t and v_t are idiosyncratic noises with serial and contemporaneous independences

¹² For seasonal adjustment, the data used the ARIMA (X-13) method which is available in E-views package.

¹³ Trend and seasonality are the main issues of non-stationary data.

¹⁴ The study estimated the weights separately by using the OLS and FGLS.

¹⁵ In the first step, it used PC to extract the common factor and uses the OLS and FGLS to estimate the parameters or factor loadings. In the second step, it replaced this consistent estimators into the DFM to estimate the unobserved state of the economy.

- 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 coefficients while the other parameters are restricted as zero).

The state-space model consists of two types of equations: the observation or signal equation and transition or state equation. Equation (1) is the observation equation. It explains the relationship of the observed variables as a linear function of the unobserved state. Equation (2) is the transition equation. It expresses the movement of the state variable over time. Although the most common form of the transition equation is in the autoregressive (AR) representative, it could include the autoregressive moving average (ARMA) model as well.¹⁶ Exogenous variables can enter either observation or transition equations without losing any interpretation. Including exogenous variables improve the model goodness of fit. The study refers to the Akaike Information Criteria (AIC) for model selection of the ARMAX linear state-space structure.¹⁷ Shumway and Stoffer (2017) showed the consistency of common factors in the linear state-space model (ARMAX).

This study imposed some assumptions. First, S_t and Y_t cointegrated, so both are I(1) series while the idiosyncratic noise, u_t , is stationary I(0).¹⁸ Second, u_t and v_t are serial and contemporaneous uncorrelated (mutually independence).¹⁹ Third, for simplicity, the study examines the DFM in the context of the linear system. For a non-linear system, other versions of Kalman filters could be implemented, for example, the EKF and UKF. Forth, the model also assumes the initiate 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 a diffuse initial state condition.²⁰

Any system of equations that can be expressed in state-space form can be solved using the Kalman filter. The study uses the Kalman filter to estimate the unobserved state, S_t .²¹ The intuitive of Kalman filter is to update the state from S_t^t to S_{t+1}^{t+1} when new 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 ARMAX linear state-space model in this study is:²²

$$S_{t+1}^t = \theta_t S_t^t \quad (3)$$

$$P_{t+1}^t = \theta_t P_t^t \theta_t' + Q_t \quad (4)$$

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

$$P_{t+1}^{t+1} = (I - K_{t+1} \rho_{t+1}) P_{t+1}^t; \text{ where } I \text{ is the identity matrix} \quad (6)$$

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

¹⁶ The ARMA linear state-space model varies across studies depending on the interest of authors.

¹⁷ MARSS package in R allows for a flexibility of adding the exogenous vector into the state-space model. See: Holmes, Ward, & Scheuerell (2014, chapter 10)

¹⁸ This is for cointegrated series. If they are not cointegrated, the residual is not I(0). In this case, instead of using OLS, the study uses FGLS to estimate the parameters.

¹⁹ This is for simplicity. The Cov (u_t, v_t) $\neq 0$ could be the case, but it does not affect the updating process. See: Shumway & Stoffer (2017 p.319).

²⁰ Durbin & Koopman (2012, chapter 5) provided a comprehensive treatment of the initialization process of the linear state-space model.

²¹ Kalman smoother can be implemented to estimate the state variable as well. In estimating 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, 2017, p. 292).

²² 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 .

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

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

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

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

Deriving these equations is based on Shumway & Stoffer (2017) 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 Kalman filter and smoother in equations (5), (6), (8), and (9).²³

To implement the Kalman filter, it replaces parameters, ρ_t , in the system by their consistent estimators. As mentioned above, using the state variable generated by PC, the study estimates the initial weights by OLS and FGLS. 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 uses the FGLS to estimate the parameters as suggested in Wu, You, & Zou (2016). The remaining parameters are estimated by maximum likelihood estimation in the DFM. Let $\varphi = \{Q_t, \alpha_t, R_t\}$ is 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 serial and contemporaneous independences. The likelihood is calculated from the innovations $\epsilon_1, \epsilon_2, \dots, \epsilon_t$.

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

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

The next section shows the result of DFM in estimating the state of the economy of Cambodia. Besides, it also examines how the government expenditure and FDI inflow comove with this coincident indicator.

IV. Result of the study

1. Descriptive statistics

Table (1) summarizes the statistical properties of the series. Additionally, it uses the Augmented Dickey-Fuller (ADF) test to check the stationary of individual series. All series are I(1). In the case of cointegration between the state and observed variables, using the original series may improve the estimation of common factors as pointed out in Corona, Poncela, & Ruiz (2020).

2. State estimation

2.1. Initial coefficients estimation

The increase in cross-sectional dimension induces the number of parameters to be estimated in the system that causes the DFM less feasible in practice, especially for a finite

²³ See: Shumway & Stoffer (2017) textbook for the case of correlated noises.

sample. Doz, Giannone, & Reichlin (2011) came with an idea of replacing the parameters by their consistent estimators. With stationary assumption, they estimated parameters by the OLS method using the state variable produced by the principal component. This method improves 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, the study uses the non-stationary series. Table (2) shows the result of the Ordinary Least Squares (OLS) of each series on the state variable and their residual tests. The OLS results show that some series cointegrate with the state variable. For the cointegrated series, the OLS parameters estimation holds although its inference is not valid. For series that are not cointegrated, OLS estimation is spurious, so the study refers to the FGLS to estimate the parameters.

2.2. Dynamic factor model (ARMAX)

So far, the study has not indicated a specific form of ARMAX linear state-space model yet. Using the state variable generated by the PC, it constructs the state equation of the DFM in AR form. Table (3) shows various lag selection criteria. The Akaike Information Criteria (AIC) suggests the AR(4) model for the state equation. For non-stationary AR model, the asymptotic distribution of AIC held while the Bayesian Information Criterion (BIC) was weakly consistent (Tsay, 1984). For the observation equation, the study introduces two lags of the state variable. The study controls for exogenous variables to improve the model fitness.²⁴ Together, the linear state-space model in this study is ARMAX (4,2,0). Appendix B shows the additional parameter restrictions. Moreover, it replaces the parameters ρ_t with $\hat{\rho}_t$ estimated by the OLS and FGLS.

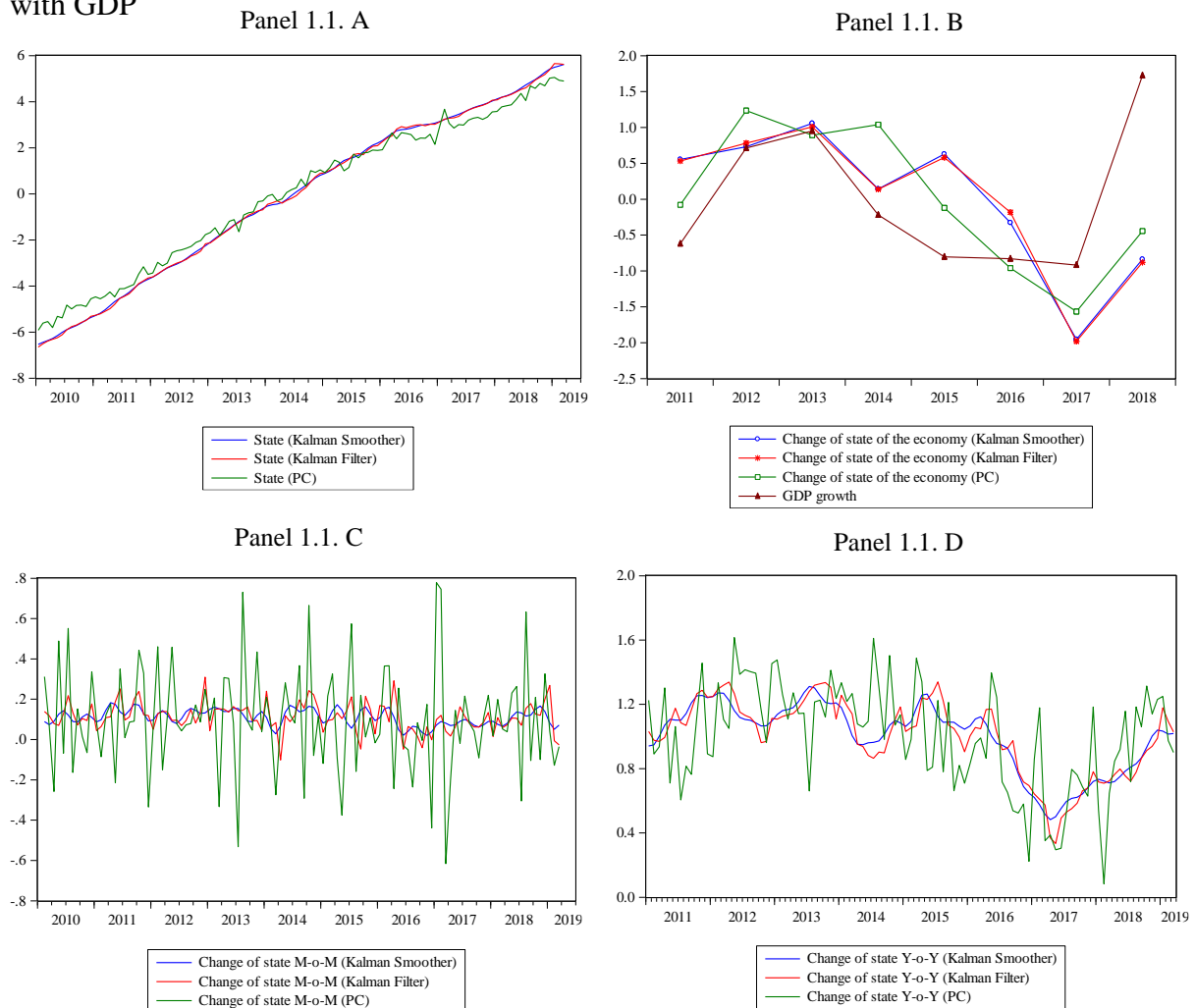
The study uses a diffuse initial state condition. Additionally, it restricts the variances of the idiosyncratic noises of both observed and state equations to be non-negative. Table (4) shows the results of the state-space model. Figure (4) shows the movement of the state of the economy estimated by the Kalman filter and smoother along with their confidence intervals as well as the residual result. Figure (5) shows the distribution of the estimated disturbances of 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 are not Gaussian. Durbin & Koopman (2012) showed that even without normality assumption, from the minimum variance linear unbiased estimates 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 Cambodia economic condition. Panel 1.1A shows the estimation of the state variable by various methods. Panel 1.1B shows its comovement to the GDP growth rate.²⁵ All methods tend to capture well when the economy moves into a bad time. Using the state estimation by various methods, the study compares the change of the state of the economy to its average level. The Kalman smoother gives a smoother state compared to other methods with a lower variation. Panel 1.1C displays the result of the month-on-month change. As expected, the Kalman smoother shows a lower variation compared to the Kalman filter and the principal component. Panel 1.1D shows the ratio of the year-on-year change to the average level. All methods reveal a similar pattern. It shows that the economy performed below the average level during 2016 and 2017. The economy somehow recovers after mid-2017.

²⁴ The AIC value of the model with exogenous variables is -1.940 while it is 3.129 for the model without exogenous variables.

²⁵ Data is aggregated into an annual data for comparison.

Figure 1 State estimation by Kalman smoother, Kalman filter, and PC and its comovement with GDP



3. Fiscal policy and state of the economy

It is worth to emphasize that from a theoretical standpoint, the fiscal policy should be countercyclical to the economic condition. However, empirical studies revealed a procyclical fiscal position in many developing countries, for example, Ilzetzki & Végh (2008), and Dabla-Norris, et al. (2010).

This study examines Cambodia's fiscal position using the state variables estimated by the Kalman smoother, Kalman filter, and principal component using the autoregressive distributed lags (ARDL) model. Pesaran & Shin (1998) mentioned that the cointegration in the ARDL model must be unique. The model was not appropriate when more than one cointegration existed. The study applies the differencing method for regressors that have unit root to restrict the cointegration among the regressors. By doing so, it ensures that only the state variable and dependent variable are I(1) series. For a unique cointegration, the ARDL model can be reparameterized into the error correction model (ECM) (Nkoro & Uko, 2016). For small samples, ARDL is more efficient than Johansen's cointegration test. Pesaran & Shin (1998) indicated that ARDL model could address the serial correlation and endogeneity issues by adding appropriate lags of regressors from ARDL(p, q) to ARDL (p,m), for $m \geq q$. The study refers to the AIC for lag selection.²⁶

²⁶ The ARDL (p, q) model is: $\Delta Y_t = c + \delta t + \phi Y_{t-1} + \lambda X_{t-1} + \sum_{i=1}^{p-1} \omega_i \Delta Y_{t-i} + \sum_{i=1}^{q-1} \tau_i \Delta X_{t-i} + \gamma \Delta X_t + \alpha Z_t + \rho(S_t - S_{t,hp}) + \epsilon_t$.

The study generates an output gap variable defined as the deviation of the state of the economy from its potential level, where it uses the Hodrick-Prescott (HP) filter to estimate the state's potential level. It also generates the state dummy variable.²⁷ Besides, the study generates a dummy variable for the election period.

Table (5) shows the impact of the state of the economy on government expenditure using the three different estimated state variables. Each contains two different models: one controls additional regressors while the other one does not. The study finds out evidence of the procyclical fiscal policy. When the economy accelerates, expenditure tends to loose due to an overall increase in revenue collection. Additionally, the ARDL conditional error correction reveals a cointegration between fiscal expenditure and the state of the economy. The parameters of other exogenous variables are well-behaved although most of them are not significant. Interestingly, expenditure tends to be lower during the election period. The result is significant at a lower level for the state variables estimated by Kalman filter and smoother. Higher inflation tends to slow down government expenditure.

Although the government fiscal condition tends to be procyclical, there may be a good reason to believe that the government tends to react more strongly during bad time compared to when the economy is in a good shape. To examine this, the study allows for different slope.²⁸ Table (6) shows the result of different slope estimation of the relationship between the output gap and government expenditure. The study finds out that the slope is relatively steeper when the economy is in bad condition. The result is robust since all models show the same pattern. The coefficients of other variables are similar to Table (5).

4. Foreign direct investment inflow and the state of the economy

Empirical studies of relations between foreign direct investment (FDI) and economic growth have a long discussion. Some studies founded the impact of FDI on economic growth as in De Mello (1999), and Devajit (2012) while another revealed that economic growth as a factor of FDI inflow, for example, Roy & Mandal (2012). Additionally, Srinivasan, Kalaivani, & Ibrahim (2010), and Hossain & Hossain (2012) found out cointegration between the two variables. Türkcan, Duman, & Yetkiner (2008) pointed out the simultaneous causation between FDI and economic growth. Alfaro (2003) studied the heterogeneity across sectors on the relationship between FDI and growth. This study skips the discussion for the sake of time and page limits.

This study explores the impact of the economic condition on FDI inflow by exploiting the coincident index.²⁹ It uses the ARDL model just like the study on the fiscal part. Bevan & 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.

²⁷ When the state is above its potential level, it records as a good state (dummy = 1). When the ratio is less than 1, it is recorded as a bad state (dummy = 0).

²⁸ The estimated equation can be written as:

$$\Delta Y_t = \begin{cases} a_0 + \delta t + \phi Y_{t-1} + \lambda X_{t-1} + \sum_{i=1}^{p-1} \omega_i \Delta Y_{t-i} + \sum_{i=1}^{q-1} \tau_i \Delta X_{t-i} + \gamma \Delta X_t + \alpha Z_t + \rho_1 (S_t - S_{t, hp}) + \epsilon_t, & S_t > S_{t, hp} \\ b_0 + \delta t + \phi Y_{t-1} + \lambda X_{t-1} + \sum_{i=1}^{p-1} \omega_i \Delta Y_{t-i} + \sum_{i=1}^{q-1} \tau_i \Delta X_{t-i} + \gamma \Delta X_t + \alpha Z_t + \rho_2 (S_t - S_{t, hp}) + \epsilon_t, & S_t \leq S_{t, hp} \end{cases}$$

By assuming noises to be the same, we can combine this two equations into a single equation by using the indicator function: $\Delta Y_t = b_0 + \mu I_t + \delta t + \phi Y_{t-1} + \lambda X_{t-1} + \sum_{i=1}^{p-1} \omega_i \Delta Y_{t-i} + \sum_{i=1}^{q-1} \tau_i \Delta X_{t-i} + \gamma \Delta X_t + \alpha Z_t + \rho_1 I_t (S_t - S_{t, hp}) + \rho_2 (1 - I_t) (S_t - S_{t, hp}) + \epsilon_t$; where I_t equals 1 for $S_t > S_{t, hp}$ and 0 for $S_t \leq S_{t, hp}$. we can estimate this equation by generating two more variables: $S_t^+ = I_t (S_t - S_{t, hp})$ and $S_t^- = (1 - I_t) (S_t - S_{t, hp})$.

²⁹ The study uses quarterly data because only quarterly FDI is available.

Table (7) shows the dynamic relationship between the state of the economy and FDI inflow. It finds out that the state of the economy has a positive impact on FDI inflow in the short-run. All the models using the state variables produced by the Kalman smoother, Kalman filter, and principal component point to a similar tendency. The short-term effect is highly significant. The coefficient of the output gap shows the same pattern. Additionally, the study reveals a cointegration between the state of the economy and FDI inflow. The long-run coefficient turns negative. One explanation of this negative impact could be due to the diminishing marginal product of capital. From the supply side, capital investment becomes less attractive that scares FDI inflow. Additionally, the coefficients of exogenous variables are well-behaved. For example, inflation shows a negative effect on FDI inflow. High inflation may indicate a high cost of investment that often associates with 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, it is still ambiguous to interpret the impact of interest rate on FDI inflow. One reason is that the high rate of return often relates to a higher country risks of investment, especially for a small open economy. The effect of trade is diverse across models. The issue could be due to the short lags inclusion.

V. Discussion and limitation of the study

The construction of the coincident index in this study comes with some limitations. First, it estimates the unobserved state in a linear context. Many macroeconomic variables comove in a non-linear structure. Morphy (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 Kalman filter dealing with the non-linear structure are the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). György, Kelemen, & Dávid (2014) discussed how each algorithm works. Julier & Uhlmann (1997) introduced UKF as a more superior version to EKF. The UKF is a derivative-free filter that does not require to calculate the Jacobian compared to the EKF. Both EKF and UKF approximate the distribution with the Gaussian assumption. Another type of application, particle filter, could also be used for a non-linear system. Unlike EKF and UKF, particle filter does not require Gaussian assumption.

The second limitation of this study involves the implementation of time-invariant parameters in DFM. 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, & Watson (2013) categorized conditions that a standard estimation of factors could tolerate temporal parameters instability. Stock & Watson (2002) showed that estimated factors under PC were consistent with small time-variate parameters. The structural break may be less an issue in this study, for it uses a sample within a short and stable period.

Third, although all methods in constructing coincident index highly correlated with GDP growth, it does not necessarily mean one is better than another. Kalman smoother used all information about the past, present, and future to estimate the state variable while Kalman filter used information of the past or present for estimation (Shumway & Stoffer, 2017). The principal component uses a linear transformation, so it does not take into account the misspecification of functional form.

Another issue relates to the quality of data and how a real economy behaves. The theory depends mainly on the data-generating process for verification. However, real-time data subject to measurement errors. One potential challenge in this study is the availability and quality of data. As a developing country, Cambodia faces a constraint in collecting information of the informal sectors. Facing this problem, how well CI could capture the state of the economy is unknown. Although the quality of data in constructing the CI remains a topic for discussion, its establishment is crucial in monitoring economic activity. CI is a useful tool for early warning

and signalling when the economy moves into a recession. The effect may come with lags, longer or shorter, depending on the characteristics of each economy. The policymakers should use this index with cautious. A similar issue occurs in the study of the fiscal stance. The models based on the lag selection criteria do not suggest a longer lag inclusion. Some economies may decide their expenditure based on a fiscal framework, annually or three to five years beforehand. In this case, models may fail to capture the reality. Besides, inside and outside lags are additional issues that this study may not properly take into account. Incorporating longer lags may resolve the issue. However, with relatively small samples, the tradeoff could be high.

Lastly, there are two other problems regarding the ARDL model. One of them involves a cointegration issue. Pesaran & Shin (1998) mentioned 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 introduces the first differencing method to transform all I(1) regressors, except the state variable. The other issue relates to the serial correlation and endogeneity problems. Empirical studies on fiscal stance and state of the economy concerns on endogeneity issue, for example, Gemmell, Kneller, & Sanz (2016). The same issue arises in the study of comovement between FDI inflow and the state of the economy, for example, in Türkcan, Duman, & Yetkiner (2008). It turns out that the ARDL model can be a potential approach to address these problems. The ARDL could resolve the serial correlation and endogeneity issues by adding appropriate lags regressors, for example, ARDL (p, q) to ARDL (p, m), for $m \geq q$ (Pesaran & Shin, 1998). The study conducted a residual's diagnostic by checking serial correlation using the LM test.

VI. Conclusion

This study constructs a coincident indicator as an unobserved state of the economy using a two-step procedure proposed by Doz, Giannone, & Reichlin (2011). In the first step, it estimates the parameters by the OLS and FGLS methods using the state variable generated by the principal component. In the second step, it estimates the unobserved state via the dynamic factor model (DFM) by substituting parameters with their estimators in the first step. This approach could improve the estimation of common factors substantially as emphasized in Giannone, Reichlin, & Small (2008), and Stock & Watson (2011). The study uses non-stationary data because it will throw away a large amount of information if cointegration exists. Corona, Poncela, & Ruiz (2020) showed that in the case of cointegration, using the original series could improve the estimation of state variable compared to first-differencing. The coincident indicator in this study points to a slower economy performance during 2016 and 2017 within the period of study. The economy somehow rebounded after mid-2017.

It also examines the relations between Cambodia's fiscal position and the state of the economy. It turns out that Cambodia fiscal position tends to be procyclical. By introducing the different slope, the study finds out that the fiscal stance reacts more strongly when the economy is in a bad time. Lastly, the study observes the comovement between FDI inflow and the state of the economy. It finds out a positive impact of the state of the economy on FDI inflow in the short-run. All models using the state variable produced by various methods show a similar tendency while the long-run coefficient becomes negative. One reason could be due to a diminishing marginal rate of return on capital.

References

- Albu, L. L., & Dinu, V. (2009). How deep and how long could be the recession in Romania. *Amfiteatru Economic*, 11, 675-683.
- Alfaro, L. (2003). *Foreign direct investment and growth: Does the sector matter*. Harvard Business School.
- 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 Paper.
- 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.
- Anderson, B. D., & Moore, J. B. (1979). Time invariant filters. In *Optimal filtering* (pp. 62-89). Prentice-Hall, Inc.
- Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real-time measurement of business conditions. *Business & Economic Statistics*, 27, 417-427.
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1), 135-171.
- 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*.
- Bujosa, M., Garcia-Ferrer, A., Juan, A. d., & Martin-Arroyo, A. (2018). Evaluating early warning and coincident indicators of business cycles using trends. Do forecasts combinations help?
- 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.
- Caines, P. E. (1988). Maximum likelihood estimation of Gaussian ARMAX and state-space system. In P. E. Caines, *Linear stochastic systems* (pp. 411-480). SIAM.
- Choi, I. (2012). Efficient estimation of factor models. *Econometric Theory*, 28(2), 274-308.
- Corona, F., Poncela, P., & Ruiz, E. (2020). Estimating non-stationary common factors: implications for risk sharing. *Computational Economics*, 55(1), 37-60.
- Dabla-Norris, E., Allen, R., Zanna, L.-F., Prakash, T., Kvintradze, E., Lledo, V., . . . Gollwitzer, S. (2010). *Budget institutions and fiscal performance in low-income countries*. IMF working papers.
- De Mello, L. R. (1999). Foreign direct investment-led growth: evidence from time series and panel data. *Oxford economic papers*, 51(1), 133-151.
- 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.
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.

- Forni, M., Giannone, D., Lippi, M., & Reichlin, L. (2009). Opening the black box: Structural factor models with large cross sections. *Econometric Theory*, 25(5), 1319-1347.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The Generalized Dynamic-Factor Model: Identification and Estimation. *The Review of Economics and Statistics*, 82(4), 540-554.
- Freudenberg, M. (2003). *Composite indicators of country performance: A critical assessment*. Paris: OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/405566708255>
- Gemmell, N., Kneller, R., & Sanz, I. (2016). Does the Composition of Government Expenditure Matter for Long-Run GDP Levels? *Oxford Bulletin of Economics and Statistics*, 78(4), 522-547.
- Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665-676.
- Goshima, K., Ishijima, H., Shintani, M., & Yamamoto, H. (2019). *Forecasting Japanese inflation with a news-based leading indicator of economic activities (No. CARF-F-458)*. Center for Advanced Research in Finance, Faculty of Economics, The University of Tokyo.
- 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.
- Gupta, R., & Kabundi, A. (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting*, 27(4), 1076-1088.
- György, K., Kelemen, A., & Dávid, L. (2014). Unscented Kalman Filters and particle filter methods for nonlinear state estimation. *Procedia Technology*, 12(1), 65-74.
- Holmes, E. E., Ward, E. J., & Scheuerell, M. D. (2014). *Analysis of multivariate time-series using the MARSS package*. NOAA Fisheries, Northwest Fisheries Science Center, 2725, 98112.
- 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.
- Ilzetzi, E., & Végh, C. A. (2008). *Procyclical fiscal policy in developing countries: Truth or fiction? (No. w14191)*. National Bureau of Economic Research.
- IMF. (2018). *Cambodia: 2018 Article IV consultation-press release; staff report; staff statement; and the statement by the executive director for Cambodia*. International Monetary Fund.
- Julier, S. J., & Uhlmann, J. K. (1997, July). New extension of the Kalman filter to nonlinear systems. *In Signal processing, sensor fusion, and target recognition VI*. 3068, pp. 182-193. International Society for Optics and Photonics.
- 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.
- King, G., & Roberts, M. E. (2015). How robust standard errors expose methodological problems they do not fix, and what to do about it. *Political Analysis*, 23(2), 159-179.
- Levanon, G. (2010). Evaluating and comparing leading and coincident economic indicators. *Business Economics*, 45(1), 16-27.

- 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.
- 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.
- 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.
doi:10.1111/j.1468-0084.2009.00567.x
- 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.
- Munda, G., & Nardo, M. (2005). Constructing consistent composite indicators: The issue of weights.
- Murphy, K. P. (2012). State space models. In K. P. Murphy, *Machine learning: A probabilistic perspective* (pp. 631-660).
- Nilsson, R., & Brunet, O. (2006). Composite leading indicators for major OECD non-member economies: Brazil, China, India, Indonesia, Russian Federation, South Africa. *OECD Statistics Working Paper*.
- Nkoro, E., & Uko, A. K. (2016). Autoregressive Distributed Lag (ARDL) cointegration technique: application and interpretation. *Journal of Statistical and Econometric Methods*, 5(4), 63-91.
- OECD. (2010, September 13). Asian Business Cycles Quarterly. pp. 1-5. Retrieved from <http://www.oecd.org/dev/asia-pacific/45623832.pdf>
- 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. *Econometric Society Monographs*, 31, 371-413.
- Poncela, P., & Ruiz, E. (2012). *More is not always better: Back to the Kalman dynamic factor models*.
- Rasic, I., Tkale, M., & Vizek, M. (2016). Constructing a composite coincident indicator for a post-transition country. *Economic Research-Ekonomska Istrazivanja*, 29(1), 434-445.
- 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.
- 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.
- 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.
- Shumway, R. H., & Stoffer, D. S. (2017). State space models. In *Time series analysis and its applications: with R examples* (pp. 324-411). Springer.

- Smeeke, S., & Wijler, E. (2019). High-dimensional forecasting in the presence of unit roots and cointegration.
- Srinivasan, P., Kalaiyani, 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. (2011). Dynamic factor models. In M. P. Clements, & D. F. Hendry (Eds.), *Oxford Handbook on Economic Forecasting*. Oxford: Oxford University Press.
- Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. In NBER, *NBER Macroeconomics Annual* (Vol. 4, pp. 351 - 409). 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.
- Tsay, R. S. (1984). Order selection in nonstationary autoregressive models. *The Annals of Statistics*, 12(4), 1425-1433.
- Türkcan, B., Duman, A., & Yetkiner, I. H. (2008). *How does FDI and economic growth affect each other? The OECD case*. 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*. 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 *Statistical methods in the atmospheric sciences (Fourth Edition)* (pp. 77-141). Elsevier Inc.
- World Bank. (2018, September 27). *The World Bank in Cambodia*. Retrieved from The World Bank: <https://www.worldbank.org/en/country/cambodia/overview>
- Wu, M., You, P., & Zou, N. (2016). *On spurious regressions with trending variables*.
- 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.

Appendices

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 and smoother. Shumway & Stoffer (2017) showed variance produced by the Kalman smoother is lower than a filtering process.

ARMAX linear state-space model with initial condition $S_0 \sim N(S_0^0, P_0^0)$ can be written as in equations (1) and (2)

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

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

$$\text{where } \begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim (\text{i. i. d}) N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_t & 0 \\ 0 & Q_t \end{bmatrix} \right); S_0 \sim N(S_0^0, P_0^0)$$

Because the sum of Gaussian distributions is 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, a 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

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

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

$$= \theta_t S_t^t; \text{ equation (3)}$$

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

$$= E\{(\theta_t(S_t - S_t^t) + v_t)(\theta_t(S_t - S_t^t) + v_t)' | Y_t\}$$

$$= \theta_t P_t^t \theta_t' + Q_t; \text{ equation (4)}$$

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

$$\text{Where } E(\epsilon_{t+1}) = 0$$

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

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

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

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

$$= \text{Cov}[S_{t+1} - S_{t+1}^t, \rho_{t+1}(S_{t+1} - S_{t+1}^t) + u_{t+1}]$$

$$= P_{t+1}^t \rho_{t+1}'$$

So, the joint distribution between S_{t+1} and ϵ_{t+1} condition on Y_t is:

$$\begin{pmatrix} S_{t+1} \\ \epsilon_{t+1} \end{pmatrix} | Y_t \sim N \left(\begin{bmatrix} S_{t+1}^t \\ 0 \end{bmatrix}, \begin{bmatrix} P_{t+1}^t & P_{t+1}^t \rho_{t+1}' \\ \rho_{t+1} P_{t+1}^t & \rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1} \end{bmatrix} \right)$$

$$S_{t+1}^{t+1} = E(S_{t+1} | Y_1, \dots, Y_t, Y_{t+1}) = E(S_{t+1} | Y_t, \epsilon_{t+1})$$

$$S_{t+1}^{t+1} = S_{t+1}^t + P_{t+1}^t \rho_{t+1}' (\rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1})^{-1} \epsilon_{t+1}$$

$$\text{Let } K_{t+1} (\text{Kalman gain}) = P_{t+1}^t \rho_{t+1}' (\rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1})^{-1}; \text{ equation (7)}$$

$$S_{t+1}^{t+1} = S_{t+1}^t + K_{t+1} \epsilon_{t+1}; \text{ equation (5)}$$

$$P_{t+1}^{t+1} = \text{Cov}(S_{t+1} | Y_t, \epsilon_{t+1}) = P_{t+1}^t - P_{t+1}^t \rho_{t+1}' (\rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1})^{-1} \rho_{t+1} P_{t+1}^t$$

$$= [I - P_{t+1}^t \rho_{t+1}' (\rho_{t+1} P_{t+1}^t \rho_{t+1}' + R_{t+1})^{-1} \rho_{t+1}] P_{t+1}^t$$

$$P_{t+1}^{t+1} = (I - K_{t+1} \rho_{t+1}) P_{t+1}^t; \text{ equation (6)}$$

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} conditional on Y_t is

$$\begin{pmatrix} S_t \\ S_{t+1} \end{pmatrix} | Y_t \sim N \left(\begin{bmatrix} S_t^t \\ S_{t+1}^t \end{bmatrix}, \begin{bmatrix} P_t^t & P_t^t \theta_t' \\ \theta_t P_t^t & P_{t+1}^t \end{bmatrix} \right)$$

$$E(S_t | S_{t+1}, Y_t) = S_t^t + P_t^t \theta_t' (P_{t+1}^t)^{-1} (S_{t+1} - S_{t+1}^t)$$

$$\text{Var}(S_t | S_{t+1}, Y_t) = P_t^t - P_t^t \theta_t' (P_{t+1}^t)^{-1} \theta_t P_t^t$$

$$\text{Let } J_t = P_t^t \theta_t' (P_{t+1}^t)^{-1} \quad (10)$$

For (n) total samples, by the law of iterated expectation: $S_t^n = E(S_t | Y_n) =$

$$E(E(S_t | S_{t+1}, Y_n) | Y_n) = E(E(S_t | S_{t+1}, Y_t) | Y_n); \text{ For } n > t$$

$$= E(S_t^t + J_t (S_{t+1} - S_{t+1}^t) | Y_n)$$

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

$$P_t^n = E(S_t - S_t^n)(S_t - S_t^n)'$$

$$\text{From (8): } S_t - S_t^n = S_t - S_t^t - J_t (S_{t+1}^n - S_{t+1}^t)$$

$$S_t - S_t^n + J_t S_{t+1}^n = S_t - S_t^t + J_t S_{t+1}^t$$

Multiply both sides by its transpose and take expectation, we get

$$E[(S_t - S_t^n + J_t S_{t+1}^n)(S_t - S_t^n + J_t S_{t+1}^n)'] = E[(S_t - S_t^t + J_t S_{t+1}^t)(S_t - S_t^t + J_t S_{t+1}^t)']$$

Because cross-product terms are zero, so

$$P_t^n + J_t E(S_{t+1}^n S_{t+1}^{n'}) J_t' = P_t^t + J_t \theta_t E(S_t^t S_t') \theta_t' J_t'$$

$$E(S_{t+1}^n S_{t+1}^{n'}) = E(S_{t+1} S_{t+1}') - P_{t+1}^n = \theta_t E(S_t S_t') \theta_t' + Q_t - P_{t+1}^n$$

$$E(S_t^t S_t') = E(S_t S_t') - P_t^t$$

$$P_t^n + J_t [\theta_t E(S_t S_t') \theta_t' + Q_t - P_{t+1}^n] J_t' = P_t^t + J_t \theta_t [E(S_t S_t') - P_t^t] \theta_t' J_t'$$

$$P_t^n = P_t^t + J_t P_{t+1}^n J_t' - [J_t (\theta_t P_t^t \theta_t' + Q_t) J_t']$$

$$P_t^n = P_t^t + J_t P_{t+1}^n J_t' - J_t P_{t+1}^t J_t'$$

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

Appendix B. Setting initial parameters for DFM

Using the state variable produced by PC, the study estimates the parameters of the observation equations using the OLS and FGLS. Poncela & Ruiz (2012) showed that total uncertainty exhibited a U shape while the MSE approached zero in a small cross-sectional dimension. This is because of the increase in the number of parameters to be estimated by the model when the cross-sectional dimension increases. So, including too many parameters in the system causes DFM to perform poorly in a finite sample. This is the benefit of substituting the factor loadings by their consistent estimators. Moreover, the study restricts the parameters in the observation equations as below:

$$\begin{bmatrix} \hat{Y}_{1t} \\ \hat{Y}_{2t} \\ \hat{Y}_{3t} \\ \hat{Y}_{4t} \\ \hat{Y}_{5t} \\ \hat{Y}_{6t} \\ \hat{Y}_{7t} \\ \hat{Y}_{8t} \\ \hat{Y}_{9t} \\ \hat{Y}_{10t} \\ \hat{Y}_{11t} \end{bmatrix} = \begin{bmatrix} 0.087 & 0.060 & 0.085 \\ 0.085 & 0.048 & 0.158 \\ 0.085 & 0.135 & 0.128 \\ 0.043 & 0.092 & 0.107 \\ 0.084 & 0.118 & 0.207 \\ 0.085 & 0.052 & 0.118 \\ 0.256 & 0.013 & 0.015 \\ 0.279 & -0.002 & 0.025 \\ 0.619 & -0.372 & 0 \\ 0.930 & -0.203 & -0.555 \\ 0.287 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ 0 & 0 & \alpha_{23} \\ 0 & 0 & \alpha_{33} \\ 0 & 0 & \alpha_{43} \\ 0 & 0 & \alpha_{53} \\ 0 & 0 & 0 \\ \alpha_{71} & 0 & 0 \\ \alpha_{81} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{bmatrix}$$

List of Tables

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 (Y ₁)	17.246	0.717	-0.410	0.903	-14.049	0.000
Bank lending to the service-related sectors (Y ₂)	7.652	0.396	0.763	0.993	-12.542	0.000
Bank lending to the manufacturing sector (Y ₃)	7.885	0.503	-3.017	0.057	-13.758	0.000
Bank lending to the retail trade sector (Y ₄)	8.609	0.640	-0.633	0.858	-13.001	0.000
Bank lending to the wholesale sector (Y ₅)	8.539	0.608	-2.312	0.170	-14.933	0.000
Total electricity supply (Y ₆)	5.796	0.588	0.350	0.980	-12.247	0.000
Export Value (Y ₇)	19.611	0.798	-0.781	0.822	-21.410	0.000
Import Value (Y ₈)	14.252	0.961	-1.130	0.704	-9.796	0.000
Corporate income or profit tax (Y ₉)	11.746	0.588	-0.194	0.935	-8.347	0.000
Domestic VAT (Y ₁₀)	11.521	0.432	-0.412	0.902	-8.870	0.000
Import VAT (Y ₁₁)	12.029	0.452	-0.299	0.920	-9.615	0.000

Table 2 Cointegration result (OLS result of individual series on state variable)

Variables	Coefficient	Residual's unit root test	
		t-statistics	P-value
Total Bank credits (Y ₁)	0.302*** (0.002)	-7.808	0.000
Bank lending to the service-related sectors (Y ₂)	0.292*** (0.006)	-1.812	0.373
Bank lending to the manufacturing sector (Y ₃)	0.298*** (0.008)	-1.204	0.671
Bank lending to the retail trade sector (Y ₄)	0.300*** (0.004)	-1.810	0.374
Bank 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)	0.303*** (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

Standard errors are in the parenthesis

Note. *, **, *** are significance 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

*suggests the lag selection

Table 4 DFM result (parameter estimation by MLE)

	Coef.	S. D	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
$\text{Var}(v_t = e^{c_{12}})$	-7.873	0.365	-21.584	0.000
$\text{Var}(u_{1t} = e^{c_1})$	-8.388	0.170	-49.402	0.000
$\text{Var}(u_{2t} = e^{c_2})$	-3.650	0.225	-16.218	0.000
$\text{Var}(u_{3t} = e^{c_3})$	-1.889	0.393	-4.807	0.000
$\text{Var}(u_{4t} = e^{c_4})$	-4.020	0.284	-14.146	0.000
$\text{Var}(u_{5t} = e^{c_5})$	-1.360	0.504	-2.698	0.007
$\text{Var}(u_{6t} = e^{c_6})$	-3.965	0.195	-20.284	0.000
$\text{Var}(u_{7t} = e^{c_7})$	-2.851	0.163	-17.513	0.000
$\text{Var}(u_{8t} = e^{c_8})$	-2.710	0.133	-20.382	0.000
$\text{Var}(u_{9t} = e^{c_9})$	-2.293	0.129	-17.770	0.000
$\text{Var}(u_{10t} = e^{c_{10}})$	-1.207	0.113	-10.662	0.000
$\text{Var}(u_{11t} = e^{c_{11}})$	-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

	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 criter.		-1.697

Table 5 The impact of the state of the economy on the fiscal stance (Conditional error correction form and bound test)

Dependent Var. Δ Expenditure	State (Kalman Smoother)		State (Kalman Filter)		State (principal component)	
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Expenditure (-1)	-2.033*** (0.369)	-2.490*** (0.450)	-2.049*** (0.373)	-2.553*** (0.448)	-2.342*** (0.380)	-2.671*** (0.432)
State (-1)	-6.740*** (1.364)	-11.031*** (3.252)	-6.675*** (1.405)	-11.898*** (2.865)	-7.047*** (1.281)	-8.272*** (2.077)
Revenue (-1)		0.137 (0.251)		0.106 (0.240)		0.151 (0.250)
Money supply (-1)		-1.054 (0.927)		-1.312 (0.788)		-0.111 (0.587)
Δ Expenditure (-1)	0.951*** (0.326)	1.265*** (0.392)	1.002*** (0.332)	1.352*** (0.394)	1.256*** (0.335)	1.440*** (0.376)
Δ Expenditure (-2)	0.753** (0.287)	1.013*** (0.338)	0.767** (0.296)	1.063*** (0.342)	0.999*** (0.296)	1.165*** (0.327)
Δ Expenditure (-3)	0.631** (0.239)	0.813*** (0.279)	0.607** (0.249)	0.830*** (0.286)	0.759*** (0.248)	0.888*** (0.273)
Δ Expenditure (-4)	0.463** (0.198)	0.595** (0.231)	0.428** (0.205)	0.591** (0.235)	0.553*** (0.203)	0.639*** (0.225)
Δ Expenditure (-5)	0.343** (0.156)	0.440** (0.175)	0.347** (0.157)	0.444** (0.175)	0.423*** (0.155)	0.468*** (0.169)
Δ Expenditure (-6)	0.123 (0.110)	0.180 (0.118)	0.154 (0.109)	0.197* (0.115)	0.181* (0.108)	0.207* (0.115)
Δ State	-25.510*** (9.451)	-20.534** (10.006)	-3.728 (2.903)	-8.617** (3.654)	-6.037*** (1.683)	-7.175** (2.740)
Δ State (-1)	25.562** (13.227)	11.562 (15.018)	-0.388 (2.534)	-0.107 (2.688)	0.818 (0.847)	0.908 (0.970)
Δ State (-2)	-9.563 (9.579)	1.000 (11.382)	2.468 (2.568)	3.579 (2.719)	0.324 (0.712)	0.470 (0.790)
Δ Revenue		0.108 (0.092)		0.110 (0.091)		0.108 (0.092)
Δ Revenue (-1)		-0.005 (0.157)		0.003 (0.151)		0.005 (0.155)
Δ Revenue (-2)		0.000 (0.096)		0.018 (0.094)		0.022 (0.097)
Δ Money Supply		-2.810* (1.633)		-2.944* (1.603)		-2.225 (1.606)
Δ Money Supply (-1)		-0.690 (1.657)		-0.655 (1.654)		-1.134 (1.603)
Δ Money Supply (-2)		0.555 (1.708)		0.121 (1.725)		0.214 (1.632)
Δ Interest rate		0.210 (0.223)		0.176 (0.224)		0.213 (0.226)
Inflation		-0.089* (0.050)		-0.097** (0.048)		-0.109** (0.049)
Exchange rate		0.086		0.072		0.071

		(0.066)		(0.066)		(0.065)
Output gap	4.335	8.665*	5.784**	10.658***	7.114***	8.209**
	(3.318)	(4.665)	(2.537)	(3.432)	(2.040)	(3.170)
Election (dummy)	-0.178	-0.238*	-0.217*	-0.274**	-0.164	-0.156
	(0.122)	(0.141)	(0.117)	(0.123)	(0.112)	(0.115)
State (dummy)	0.079	0.054	-0.014	0.018	-0.085	-0.126
	(0.169)	(0.181)	(0.156)	(0.166)	(0.159)	(0.169)
Trend	0.263***	0.438***	0.263***	0.476***	0.281***	0.328***
	(0.051)	(0.133)	(0.052)	(0.116)	(0.049)	(0.085)
C	-14.583***	-24.396***	-14.627***	-26.372***	-15.536***	-18.020***
	(2.832)	(7.561)	(2.895)	(6.490)	(2.704)	(4.697)
R ²	0.573	0.622	0.559	0.624	0.583	0.638
Adjusted R ²	0.499	0.493	0.483	0.495	0.511	0.514
Prob(F-statistic)	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors are in parentheses. Data are normalized with mean 0 and standard deviation 1.

Note. *, **, *** are significance at 10%, 5%, and 1%, respectively.

Table 6 The impact of the state of the economy on the fiscal stance (Different slope estimation)

Dependent Var. Δ Expenditure	State (Kalman Smoother)		State (Kalman Filter)		State (principal component)	
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Expenditure (-1)	-2.019***	-2.534***	-2.031***	-2.574***	-2.316***	-2.646***
	(0.370)	(0.452)	(0.374)	(0.450)	(0.384)	(0.437)
State (-1)	-6.693***	-11.357***	-6.609***	-11.997***	-6.952***	-8.304***
	(1.369)	(3.268)	(1.406)	(2.875)	(1.297)	(2.088)
Revenue (-1)		0.231		0.165		0.155
		(0.267)		(0.252)		(0.252)
Money supply (-1)		-1.155		-1.349*		-0.164
		(0.932)		(0.791)		(0.598)
Δ Expenditure (-1)	0.936***	1.302***	0.981***	1.370***	1.246***	1.434***
	(0.328)	(0.393)	(0.333)	(0.395)	(0.336)	(0.377)
Δ Expenditure (-2)	0.741**	1.047***	0.743**	1.073***	0.979***	1.149***
	(0.288)	(0.339)	(0.297)	(0.343)	(0.300)	(0.331)
Δ Expenditure (-3)	0.622**	0.848***	0.575**	0.834***	0.755***	0.887***
	(0.240)	(0.281)	(0.251)	(0.287)	(0.249)	(0.274)
Δ Expenditure (-4)	0.461**	0.630***	0.406*	0.594**	0.540***	0.627***
	(0.198)	(0.233)	(0.206)	(0.236)	(0.206)	(0.227)
Δ Expenditure (-5)	0.348**	0.472***	0.333**	0.447**	0.415***	0.459***
	(0.156)	(0.178)	(0.158)	(0.175)	(0.157)	(0.170)
Δ Expenditure (-6)	0.128	0.198	0.153	0.203*	0.171	0.195
	(0.110)	(0.119)	(0.109)	(0.116)	(0.110)	(0.118)
Δ State	-	-21.689**	-3.659	-8.743**	-5.910***	-7.228**
	26.262***					
	(9.530)	(10.070)	(2.903)	(3.667)	(1.703)	(2.755)
Δ State (-1)	26.421**	12.201	-0.241	0.105	0.829	0.884
	(13.313)	(15.029)	(2.538)	(2.708)	(0.851)	(0.975)
Δ State (-2)	-10.751	-0.103	2.241	3.461	0.345	0.464
	(9.738)	(11.433)	(2.577)	(2.730)	(0.716)	(0.794)
Δ Revenue		0.136		0.117		0.114

		(0.096)		(0.092)		(0.093)
ΔRevenue (-1)		-0.039		-0.023		0.000
		(0.160)		(0.155)		(0.156)
ΔRevenue (-2)		-0.016		0.009		0.018
		(0.098)		(0.095)		(0.098)
Δ Money Supply		-2.953*		-2.964*		-2.338
		(1.639)		(1.608)		(1.628)
Δ Money Supply (-1)		-0.804		-0.782		-1.173
		(1.661)		(1.667)		(1.612)
Δ Money Supply (-2)		0.582		0.089		0.310
		(1.708)		(1.729)		(1.650)
Inflation		-0.085*		-0.092*		-0.105**
		(0.050)		(0.049)		(0.050)
ΔInterest rate		0.229		0.195		0.202
		(0.224)		(0.226)		(0.228)
Exchange rate		0.077		0.064		0.069
		(0.067)		(0.067)		(0.065)
Output gap (S ⁻)	7.507	13.654**	8.744**	13.170***	7.649***	8.933**
	(5.435)	(6.791)	(3.855)	(4.702)	(2.250)	(3.472)
Output gap (S ⁺)	2.758	6.371	3.963	9.166**	6.303**	7.603**
	(3.954)	(5.187)	(3.102)	(3.932)	(2.488)	(3.388)
Election (dummy)	-0.180	0.007	-0.241**	-0.011	-0.168	-0.124
	(0.123)	(0.187)	(0.119)	(0.170)	(0.113)	(0.170)
State (dummy)	0.040	-0.245*	-0.053	-0.294**	-0.084	-0.159
	(0.177)	(0.141)	(0.161)	(0.126)	(0.160)	(0.115)
Trend	0.262***	0.451***	0.261***	0.479***	0.277***	0.330***
	(0.051)	(0.133)	(0.052)	(0.117)	(0.049)	(0.086)
C	-	-24.936***	-14.429***	-26.500***	-15.314***	-18.097***
	14.394***					
	(2.851)	(7.579)	(2.901)	(6.509)	(2.742)	(4.722)
R ²	0.575	0.627	0.564	0.627	0.585	0.639
Adjusted R ²	0.496	0.493	0.483	0.492	0.508	0.509
Prob(F-statistic)	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors are in parentheses. Data are normalized with mean 0 and standard deviation 1.

Note. *, **, *** are significance at 10%, 5%, and 1%, respectively.

Table 7 The impact of the state of the economy on FDI inflow (Conditional error correction form and bound test)

Dependent var.	State (Kalman Smoother)		State (Kalman Filter)		State (Principal component)	
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
ΔFDI						
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)
ΔState	-734.797**	-1022.636**	-645.575**	-910.541**	-482.953**	-494.055*
	(311.017)	(448.914)	(293.555)	(398.822)	(231.853)	(240.514)
ΔState (-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 (1.369)	2.046 (2.095)	1.488 (1.222)	2.731* (1.453)	0.960 (0.786)	1.808** (0.715)
Δ State (-3)	0.466 (1.010)	0.503 (1.151)	0.758 (0.910)	1.111 (0.936)	0.528 (0.490)	0.945* (0.452)
Exchange rate		-0.027 (0.153)		-0.133 (0.112)		-0.159 (0.112)
Δ Expenditure		-0.032 (0.067)		-0.043 (0.048)		-0.024 (0.046)
Inflation		-0.027 (0.037)		-0.033 (0.027)		-0.056** (0.022)
Inflation (-1)		-0.064* (0.031)		-0.061** (0.021)		-0.005 (0.020)
Δ Revenue (-1)		0.114 (0.130)		0.141 (0.092)		0.029 (0.101)
Δ Interest rate		0.021 (0.177)		0.059 (0.121)		0.240* (0.119)
Δ Size of trade		-0.070 (0.312)		-0.067 (0.213)		-0.337 (0.193)
Δ Size of trade (-1)		-0.940** (0.416)		-0.672* (0.309)		0.139 (0.351)
Δ Size of trade (-2)		-0.732 (0.443)		-0.277 (0.383)		0.865** (0.362)
Δ Size of trade (-3)		-0.209 (0.361)		-0.008 (0.235)		0.569** (0.211)
Output gap	736.209** (310.759)	1025.025** (448.680)	647.716** (293.343)	912.868** (398.615)	484.548** (231.612)	496.239* (240.306)
Output gap (-1)	-1549.478** (630.894)	-2169.146** (906.397)	-1349.155** (594.872)	-1873.795** (807.153)	-980.082** (458.646)	-936.142* (469.490)
Output gap (-2)	824.530** (323.331)	1161.452** (462.318)	710.029** (304.397)	971.596** (412.465)	501.853** (228.108)	441.851* (231.175)
Election (dummy)	0.047 (0.075)	-0.012 (0.094)	0.026 (0.064)	-0.006 (0.068)	0.041 (0.059)	0.128* (0.058)
State of the economy (dummy)	-0.078 (0.102)	-0.071 (0.123)	-0.018 (0.092)	0.026 (0.089)	-0.032 (0.075)	-0.094 (0.073)
Trend	1.272*** (0.371)	1.900*** (0.540)	1.083*** (0.347)	1.419** (0.472)	0.737*** (0.215)	0.421 (0.253)
C	-33.059*** (9.446)	-49.644*** (13.726)	-27.217*** (8.697)	-33.666** (11.964)	-16.266*** (4.717)	-4.152 (6.377)
R ²	0.648	0.858	0.723	0.920	0.751	0.937
Adjusted R ²	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

Standard errors are in parentheses. Data are normalized with mean 0 and standard deviation 1.

Note. *, **, *** are significance at 10%, 5%, and 1%, respectively.

Useful Graphs

Figure 2 Seasonal and non-seasonal adjusted data

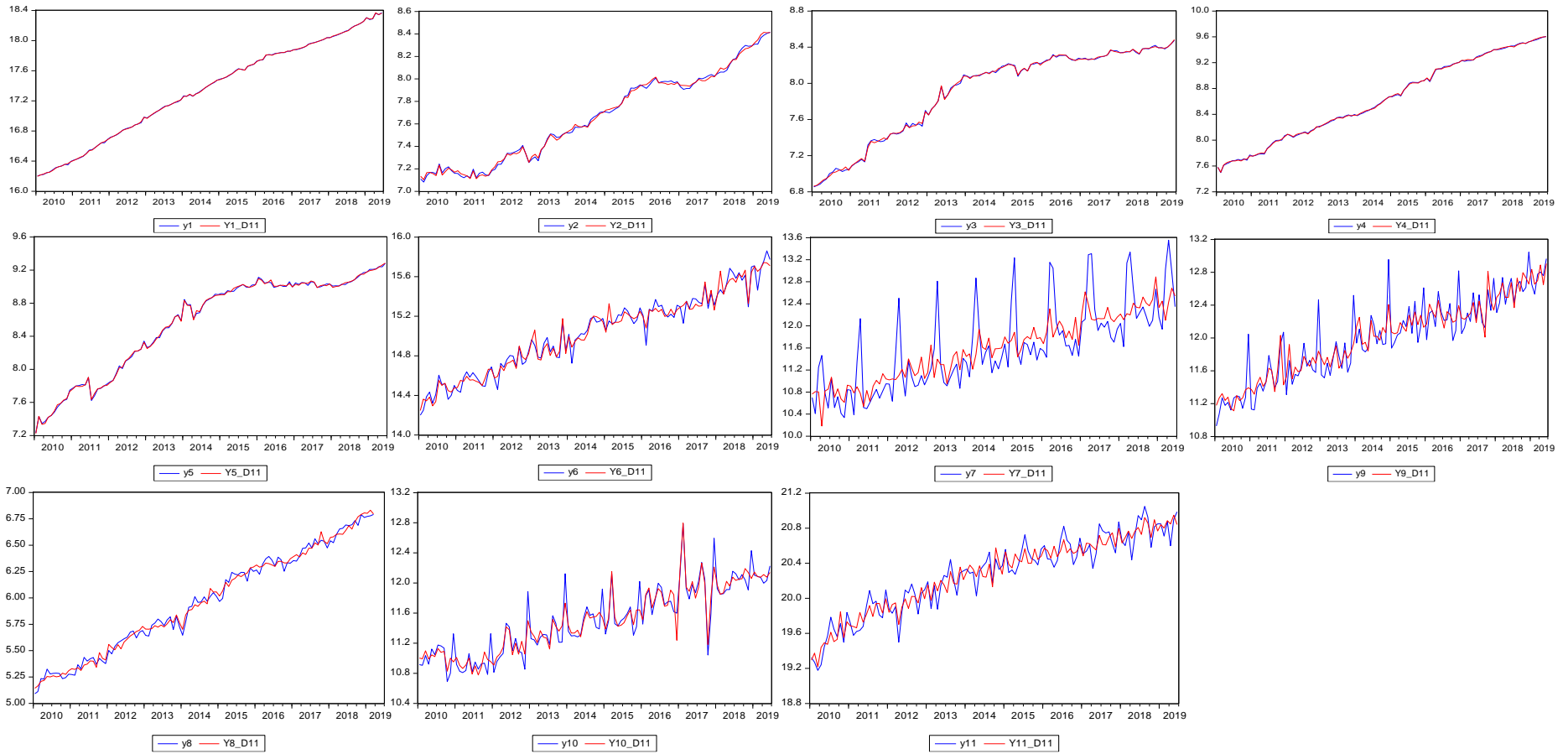


Figure 3 Residuals of OLS regression of each series I(1) on the state variable I(1)

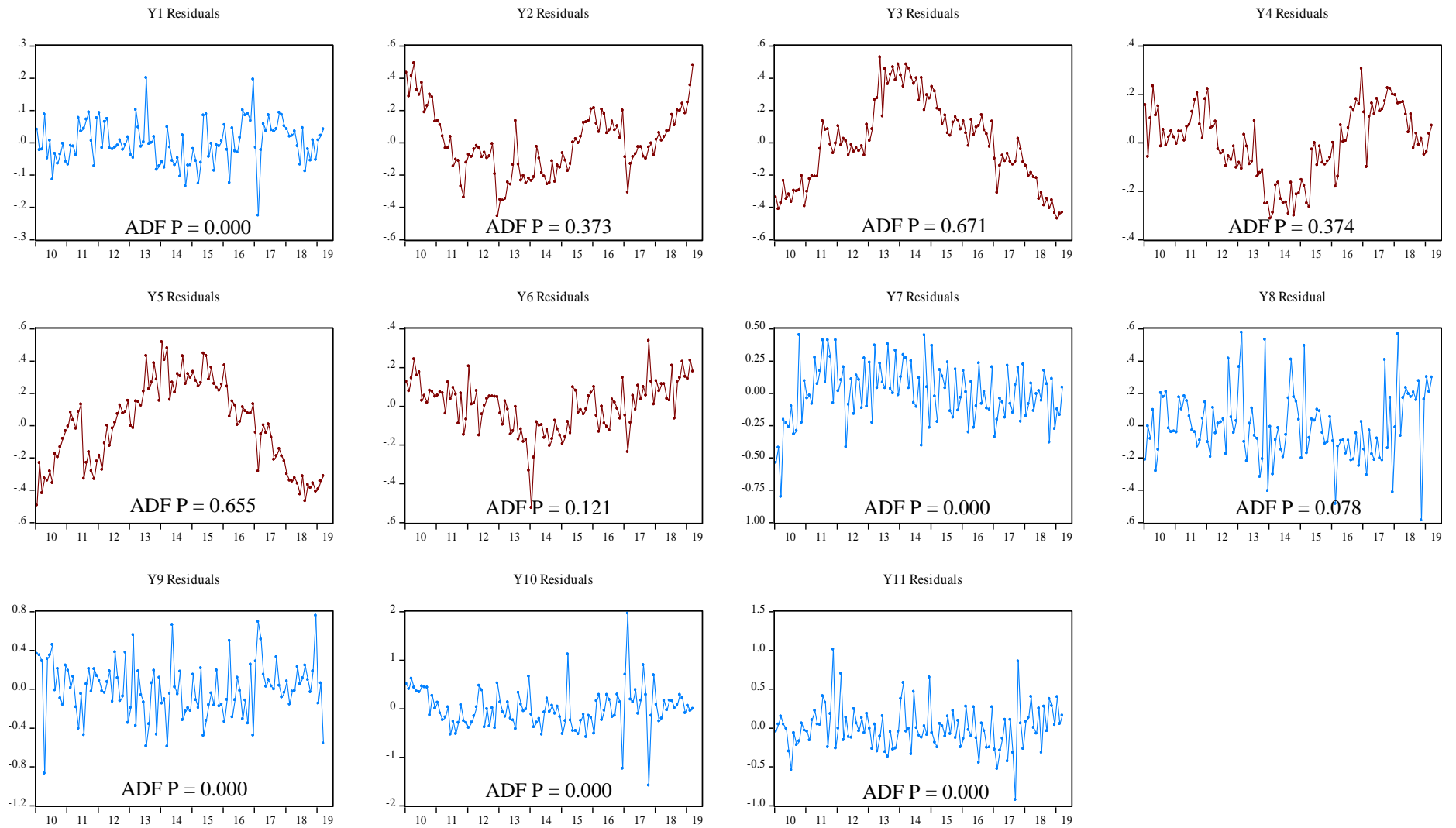
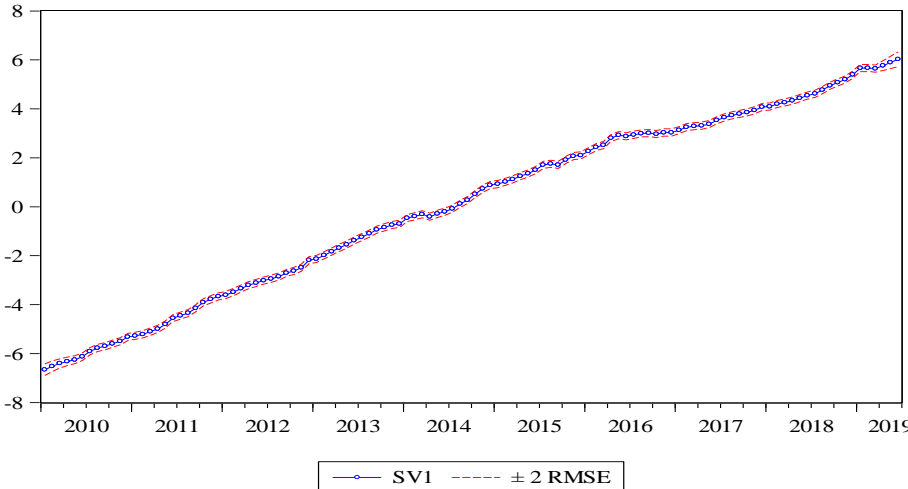
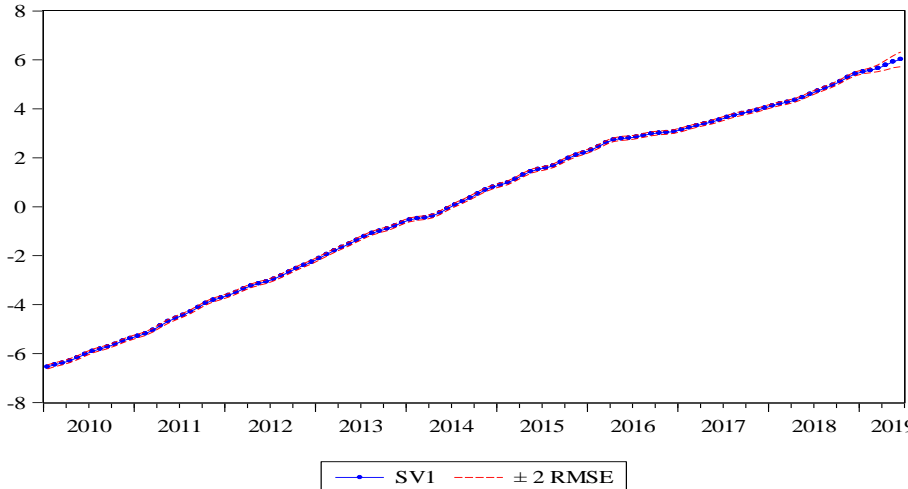


Figure 4 State estimation by Kalman smoother and Kalman filter

Filtered State SV1 Estimate



Smoothed SV1 State Estimate



Smoothed SV1 State Disturbance

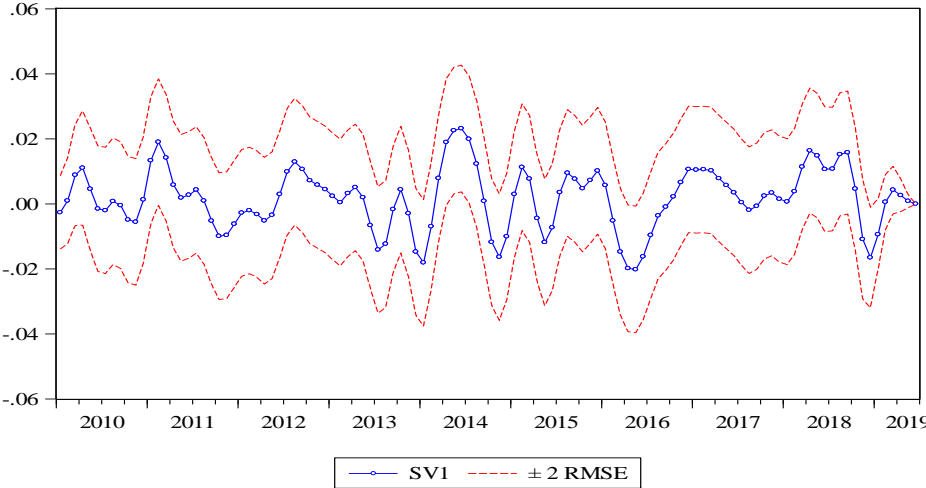


Figure 5 Residual diagnostics

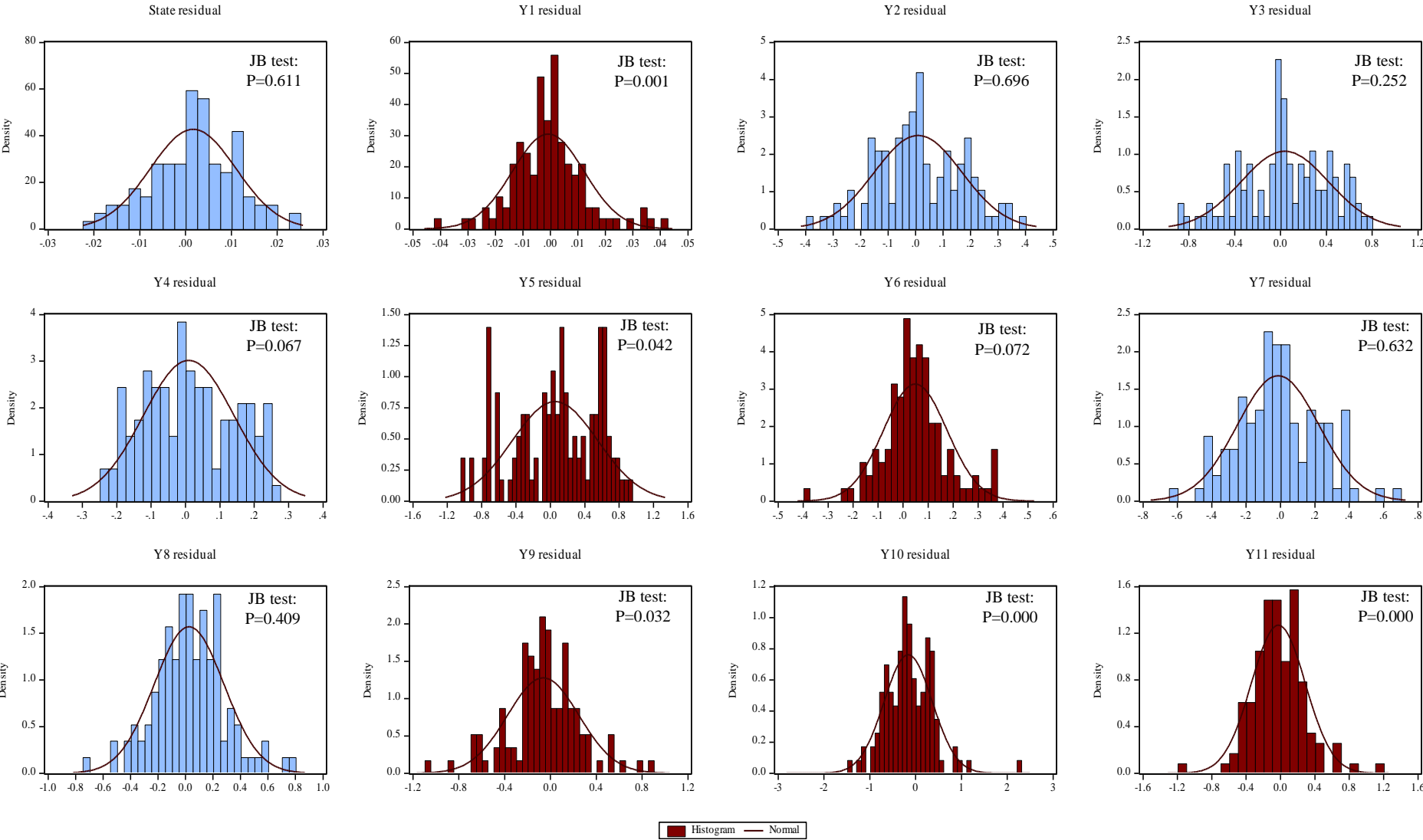


Figure 6 Government expenditure and output gap

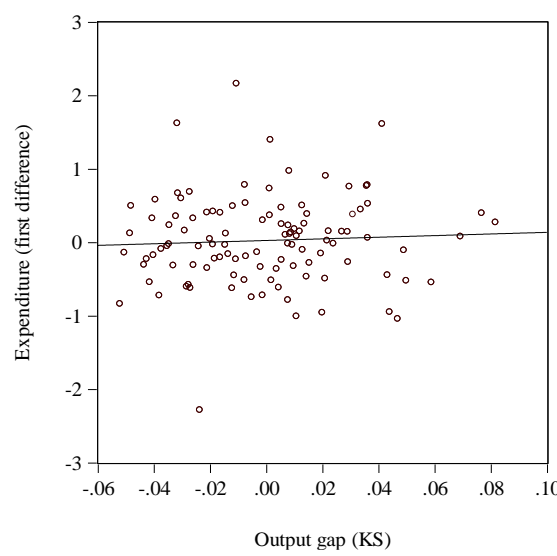
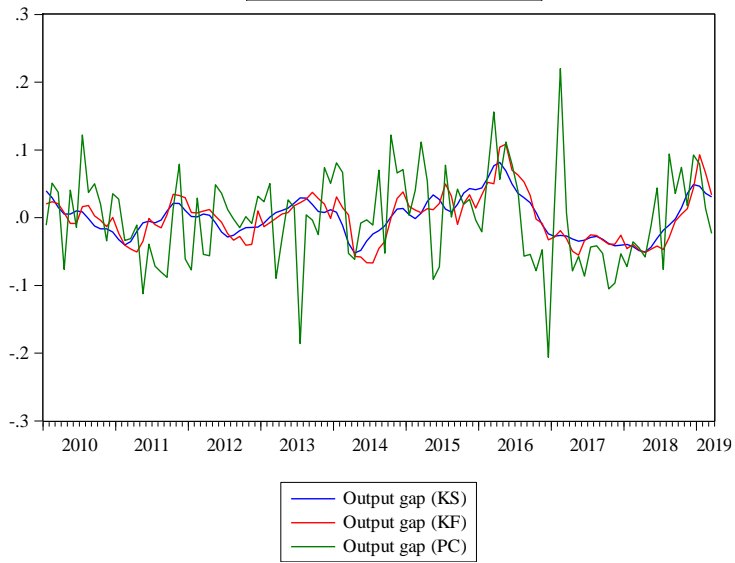
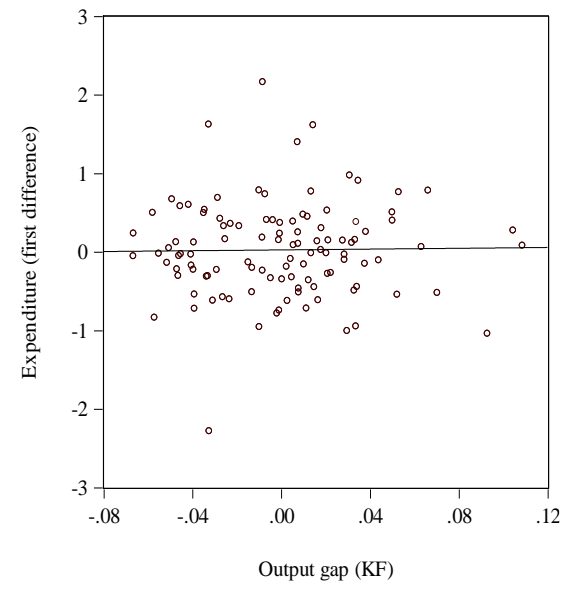
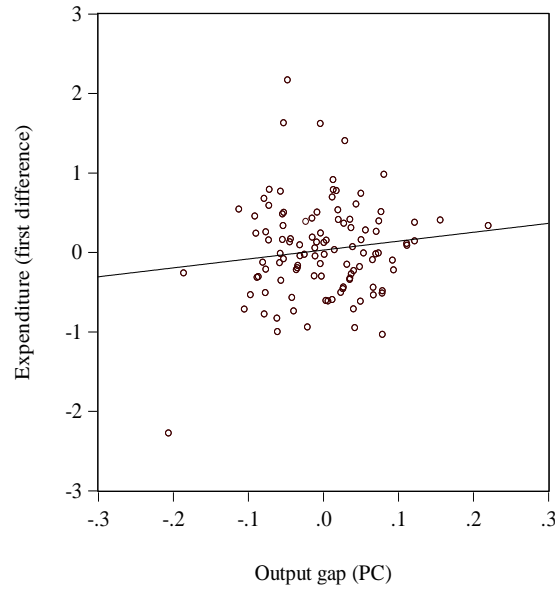
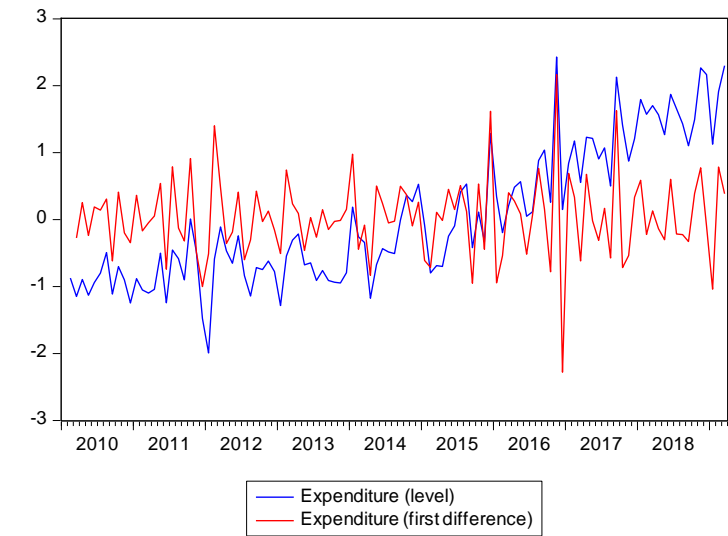


Figure 7 FDI and state of the economy

