

# On the Contribution of International Shocks in Australian Business Cycle Fluctuations<sup>1</sup>

Jamie L. Cross, Aubrey Poon

*Centre of Applied Macroeconomics and Commodity Prices (CAMP), BI Norwegian Business School*

*University of Strathclyde*

*Centre for Applied Macroeconomic Analysis (CAMA)*

---

## Abstract

What proportion of Australian business cycle fluctuations are caused by international shocks? We address this question by estimating a panel VAR model that has time-varying parameters and a common stochastic volatility factor. The time-varying parameters capture the inter-temporal nature of Australia's various bilateral trade relationships, while the common stochastic volatility factor captures various episodes of volatility clustering among macroeconomic shocks, e.g., the 1997/98 Asian Financial Crisis and the 2007/08 Global Financial Crisis. Our main result is that international shocks from Australia's five largest trading partners: China, Japan, the EU, the US and the Republic of Korea, have caused around half of all Australian business cycle fluctuations over the past two decades. We also find important changes in the relative importance of each country's economic impact. For instance, China's positive contribution increased throughout the mining boom of the 2000s, while the overall US influence has almost halved since the 1990s.

---

## 1. Introduction

Following the seminal work of [Dungey and Pagan \(2000\)](#), structural vector autoregression (SVAR) models have been the primary tool for modeling macroeconomic linkages between Australia and the rest of the world (see, e.g., [Dungey and Pagan, 2000, 2009](#); [Dungey and Fry, 2003](#); [Nimark, 2009](#); [Voss and Willard, 2009](#); [Liu, 2010](#); [Leu, 2011](#); [Dungey et al., 2014](#)). Throughout most of this literature, foreign fluctuations are assumed to originate from a single entity. This entity is often represented by the United States of America (US) ([Dungey and Pagan, 2000, 2009](#); [Voss](#)

---

<sup>1</sup>We thank Sharada Davidson and Yiqiao Sun for useful comments on the paper.

and Willard, 2009; Leu, 2011), or a single conglomerate, such as the G7 countries (Claus et al., 2008; Nimark, 2009; Liu, 2010; Leu and Sheen, 2011). The primary reason for this assumption is that it allows for a parsimonious model specification that avoids estimation concerns regarding the curse of dimensionality. However, this simplification does not come without costs. If the researcher focuses on a single economy then they forego any information relating to international transmissions from alternative trading partners. Alternatively, if the researcher considers a single conglomerate, then they forego any inference on country-specific sources of variation. For instance, using a trade-weighted index of G7 countries as a single foreign conglomerate, Liu (2010) and Nimark (2009) independently find that international shocks account for more than half of all Australian business cycle fluctuations. Under their specification, however, there is no way of disentangling the individual significance of each country. Such a decomposition is important. For instance, using a three-country SVAR Dungey et al. (2014) find that the transmission mechanism from the Euro area and the US to Australia differ substantially. Since macroeconomic shocks from distinct economies may have a substantially different impact on the Australian economy, a comprehensive investigation into the sources and effects of foreign shocks has important policy implications.

With this in mind, our objective in this paper is to quantify the proportion of Australian business cycle fluctuations that can be attributed to international shocks. To that end, we consider macroeconomic data on Australia's five largest trading partners: China, Japan, the EU, the US and the Republic of Korea. While the inclusion of multiple economies in an SVAR model is a trivial theoretical extension, in practice, the curse of dimensionality quickly becomes an issue. To overcome this problem, Bańbura et al. (2010) show how Bayesian shrinkage techniques developed by De Mol et al. (2008) can be used to estimate large scale reduced form VAR models. While such VARs are a useful tool in studying single economies such as the US, an important shortcoming of these models in the context of a multi-country analysis such as ours, is that they give no consideration to the existence of cross-sectional heterogeneity in the panel dimension of the data. In this paper, we consequently move away from the large VAR framework and instead opt to estimate a panel VAR (PVAR) model.<sup>2</sup>

---

<sup>2</sup>When choosing the model for our study we also considered estimating international factor augmented VAR

As in traditional VARs, PVAR models provide a multivariate framework through which the propagation of exogenous shocks can be analyzed. In the context of our research question, the advantage of the PVAR specification is that it allows us to capture intra- and inter-country variable interdependencies as well as cross-sectional heterogeneity. In addition to these features, we also allow for time-varying coefficients and stochastic volatility. Both features are important. The time-varying coefficients allow us to capture the evolving inter-temporal nature of Australia's various bilateral trade relationships, while the common stochastic volatility factor captures any volatility clustering among idiosyncratic macroeconomic shocks.<sup>3</sup> While the addition of time-varying parameters exacerbates the curse of dimensionality, we overcome this concern by adopting an economically meaningful factorization structure, details of which are deferred to Section 2.2.

Our main results are as follows. From a modeling perspective, the common stochastic volatility factor is capable of capturing significant exogenous shocks in and around the 1997/98 Asian Financial Crisis and the 2007/08 Global Financial Crisis, which a constant volatility model would fail to detect. Similarly, the time-varying parameters show various interesting changes in the effects of international shocks over our sample period. For instance, over the past two decades, the contribution of international shocks stemming from the US is found to have declined, while contributions from China, Japan, and the EU have increased. Interestingly, contributions from Korea have been constant over the sample period. In relation to our primary research question, we find that international shocks are responsible for around half of all Australian business cycle fluctuations over the sample period. While this aggregate result is consistent with earlier works by Liu (2010) and Nimark (2009) the country-specific contributions are novel. Another new result is that we find an important asymmetry in the effects of international shocks stemming from each country. For instance, China's positive contribution increased by six percent throughout the mining boom of the 2000s, while the overall US influence has almost halved since the 1990s. Taken

---

(FAVAR) (e.g., Mumtaz and Surico, 2009) and global VAR (GVAR) (e.g., Dees et al., 2007) models. The primary reason for choosing the PVAR framework is that it allows us to estimate time-varying coefficients in an economically meaningful way (Canova et al., 2007). We provide further details of this estimation method in Section 2.2. For a more detailed survey of the benefits of the panel VAR approach to FAVARs and GVARs see Canova and Ciccarelli (2013).

<sup>3</sup>Volatility clustering is a phenomenon whereby large changes in observations tend to be followed by large changes and small changes are followed by small changes.

together, the results highlight the dynamic nature of Australia’s trade relationships and suggest that Australian policymakers should be particularly mindful of international economic events when making decisions.

The rest of the paper is structured as follows. Section 2 presents the key steps in the methodology. This includes data transformations, model specifications and model selection. Section 3 presents the main results regarding business cycle drivers, and Section 4 concludes.

## 2. Methodology

### 2.1. Data

The set of countries considered in our analysis are Australia and its five largest trading partners: China, Japan, the EU, (the Republic of) Korea and the US.<sup>4</sup> For each country, we use a standard set of business cycle variables: real GDP, consumption, investment, the trade-balance-to-GDP-ratio and a short-term interest rate (taken to be each country’s bank rate). Due to data limitations, we were unable to include employment and government expenditure data in the analysis (government expenditure data does not exist for the EU, while, to the best of our knowledge, open-source employment data is not available for China). Since we are dealing with a small open economy that engages in inflation targeting, we also consider GDP deflator measured inflation and the real effective exchange rate. The reason for using GDP deflator measured inflation as compared to the consumer price index (CPI), is that the former includes prices of other goods that likely influence the Australian business cycle, e.g. the price of goods and services in the mining sector. Each series is measured at a quarterly frequency, and the sample runs from 1995Q1 to 2017Q4. This sample maximizes the amount of common data between countries with data from China providing the main limitation. Data for China was sourced from the Federal Reserve Bank of Atlanta. Except for the exchange rate, all data for Australia was sourced from the Australian Bureau of Statistics website. All remaining data is collected from the Federal Reserve Bank of St. Louis economic database (FRED).

---

<sup>4</sup>Our reason for limiting the analysis to this set of countries is due to a lack of data availability. We would ideally include the top 10 trading partners but could not obtain quarterly time series data for Singapore, Thailand, and India. We therefore decided to limit the analysis to Australia’s top five trading partners.

The combined bilateral trade of goods and services with these five economies accounts for approximately two-thirds of total Australian international trade over the sample period, making this the largest scale study on the effects of international shocks on Australian business cycle fluctuations. To illustrate the importance of these bilateral trade relationships, we plot the quarterly bilateral trade balance of goods and services between Australia and each of these economies in Figure 1.<sup>5</sup> It is immediately obvious that the trade relationships between Australia and the respective nations exhibit substantial time variation over the sample period. For instance, while Japan has remained a top-two trading partner over most of the sample period, total trade with China has increased exponentially, moving from the fifth largest trading partner in January 1995, to the largest at the end of the sample.<sup>6</sup> The dynamic nature of these bilateral trade relationships highlights the importance of allowing for time-varying parameters in the PVAR model.

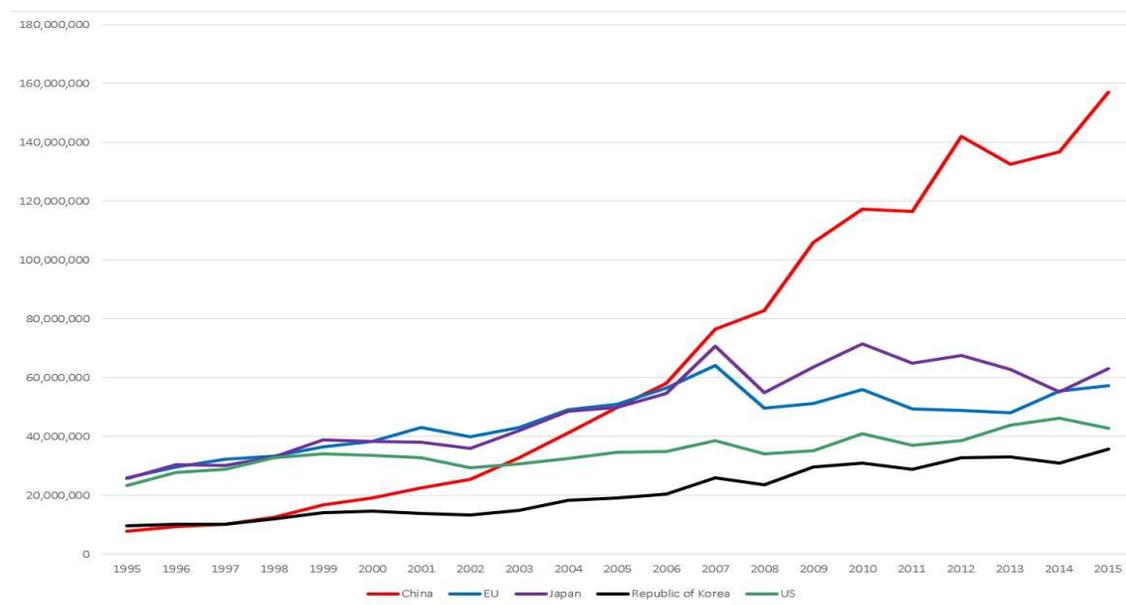


Figure 1: Total bilateral merchandise trade over the sample period in constant 2013 Australian dollar prices (Source: Australian Department of Foreign Affairs and Trade (DFAT) (2017))

<sup>5</sup>All statistics and data used in this paragraph are based on historical trade and economic data available from the Australian Department of Foreign Affairs and Trade (DFAT) website. Link: <https://dfat.gov.au/trade/resources/trade-statistics/Pages/trade-time-series-data.aspx>. Due to data availability, the EU proportion of the statistics is represented by Belgium, Finland, France, Germany, Ireland, Italy, Netherlands, and Sweden. Importantly, the data used in the empirical application covers all 28 EU countries.

<sup>6</sup>For details underlying the nature of the dramatic change in the Australia-China trade relationship we refer the interested reader to [Sheng and Song \(2008\)](#).

Before estimating the model, we annualized and standardized each series. Specifically, we first computed annualized growth rates by taking the quarter on quarter percentage change – i.e., for a given level  $X_t$ , the annualized growth rate is taken to be  $g_t = 400 \times \ln\left(\frac{X_t}{X_{t-1}}\right)$ . Next, each of the series was standardized (i.e., demeaned and divided by the standard deviation). We also tested for cointegration using the Johansen trace statistic on variables within each country and report the results in the online appendix. The results are mixed. For instance, the US variables have cointegration rank 6, while those in Japan and the EU are not cointegrated at all. Since a mixed cointegration panel VAR would be a non-trivial econometric innovation, we use variables in growth rates, however, this points to an important area of future research.

## 2.2. A time-varying parameter panel VAR model

The reduced form time-varying parameter PVAR (TVP-PVAR) model has a state-space representation. The observation equation for dates  $t = 1, \dots, T$ , is:

$$\mathbf{y}_{it} = \mathbf{c}_{it} + \mathbf{A}_{1,it}\mathbf{Y}_{t-1} + \dots + \mathbf{A}_{p,it}\mathbf{Y}_{t-p} + \mathbf{u}_{it}, \quad (1)$$

where  $\mathbf{y}_{it}$ ,  $\mathbf{c}_{it}$  and  $\mathbf{u}_{it}$  respectively denote  $G \times 1$  vectors containing variables of interest, time-varying intercepts and stochastic disturbances for the  $i$ -th country,  $i = 1, \dots, N$ ,  $\mathbf{Y}_t = (\mathbf{Y}'_{1t}, \dots, \mathbf{Y}'_{Nt})'$  is an  $NG \times 1$  vector which stacks the variables of interest for each country, and  $\mathbf{A}_{j,it}$ ,  $j = 1, \dots, p$ , is a  $G \times NG$  matrix of time-varying autoregressive coefficients. Note that when  $i = 1$  the (time-varying parameter) PVAR model becomes a standard (time-varying parameter) VAR model. Thus, the primary difference between the PVAR and VAR models is that the former accounts for country-specific heterogeneity by accounting for the panel dimension of each country's data. For estimation purposes (1) can be written in the form of a seemingly unrelated regression (SUR) model:

$$\mathbf{Y}_t = \mathbf{X}_t\beta_t + \mathbf{u}_t, \quad (2)$$

where  $\mathbf{X}_t = \mathbf{I}_{NG} \otimes \begin{bmatrix} \mathbf{Y}'_{t-1} & \dots & \mathbf{Y}'_{t-p} & \mathbf{1} \end{bmatrix}$ ,  $\beta_t = \text{vec}\left(\begin{bmatrix} \mathbf{A}_t & \mathbf{c}_t \end{bmatrix}'\right)$ ,  $\mathbf{A}_t = \begin{bmatrix} \mathbf{A}'_{1t} & \dots & \mathbf{A}'_{Nt} \end{bmatrix}'$ ,  $\mathbf{A}_{it} = \begin{bmatrix} \mathbf{A}_{1,it} & \dots & \mathbf{A}_{p,it} \end{bmatrix}$ ,  $\mathbf{c}_t = \begin{pmatrix} \mathbf{c}'_{1t} & \dots & \mathbf{c}'_{Nt} \end{pmatrix}'$  and  $\mathbf{u}_t = \begin{pmatrix} \mathbf{u}'_{1,t} & \dots & \mathbf{u}'_{N,t} \end{pmatrix}'$ . Note that  $\otimes$  denotes the Kronecker product and  $\text{vec}(\cdot)$  is a vectorization operation that takes the  $NG \times NG$  matrix of

VAR parameters:  $\mathbf{A}_t$ , and the  $NG \times 1$  intercept matrix:  $\mathbf{c}_t$ , and stacks them block by block into a  $NGk \times 1$  parameter vector:  $\boldsymbol{\beta}_t$ , where  $k = NGp + 1$ .

The cost of the added flexibility in allowing for time-varying parameters is that we must estimate an additional  $(T - 1)NGk$  states. Following [Canova et al. \(2007\)](#), this computational burden can be reduced by implementing a cross-sectional parameter shrinkage procedure, which exploits the panel dimension of the model. More precisely,  $\boldsymbol{\beta}_t$  is factorized as:

$$\boldsymbol{\beta}_t = \boldsymbol{\Xi}_1 \boldsymbol{\theta}_{1,t} + \boldsymbol{\Xi}_2 \boldsymbol{\theta}_{2,t} + \boldsymbol{\Xi}_3 \boldsymbol{\theta}_{3,t}, \quad (3)$$

where  $\boldsymbol{\theta}_{i,t}$ ,  $i = 1, 2, 3$  are mutually orthogonal vectors that induce time variation in the PVAR coefficients, with associated loading matrices  $\boldsymbol{\Xi}_i$  which are respectively of dimension  $NGk \times N_1 < N$ ,  $NGk \times N$  and  $NGk \times G$ . This factorization has the practical advantage of reducing the computational burden from  $NGk$  parameters to just  $N_1 + N + G$  factors. We highlight that the loading matrices in (3) are not estimated, but are instead manually constructed to provide a meaningful economic interpretation. To see this, note that writing (3) in matrix notation and substituting the result into (2) gives:

$$\mathbf{Y}_t = \mathbf{Z}_t \boldsymbol{\theta}_t + \mathbf{u}_t, \quad (4)$$

where  $\mathbf{Z}_t = \mathbf{X}_t \boldsymbol{\Xi}$ ,  $\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Xi}_1 & \boldsymbol{\Xi}_2 & \boldsymbol{\Xi}_3 \end{bmatrix}$ , and  $\boldsymbol{\theta}_t = \begin{pmatrix} \boldsymbol{\theta}_{1,t} & \boldsymbol{\theta}_{2,t} & \boldsymbol{\theta}_{3,t} \end{pmatrix}'$ . The coefficients in (4) provide measures of world, country and variable specific indicators that can be used to assess the relative contributions to Australia's business cycle. In particular, by setting  $N_1 = 1$ ,  $\boldsymbol{\Xi}_1$  to an  $NGk \times 1$  unit vector and  $\theta_{1,t}$  is a scalar, the first term:  $\mathbf{X}_t \boldsymbol{\Xi}_1 \theta_{1,t}$ , can be interpreted as a common or *world leading indicator* (WLI). Similarly, if we were interested in identifying a common component in small and large economies—as in [Section 2.3](#)—then we could set  $N_1 = 2$ ,  $\boldsymbol{\Xi}_1$  to be an appropriately defined  $NGk \times 2$  vector of ones and zeros and  $\theta_{1,t}$  as a  $2 \times 1$  vector. Next, by defining  $\boldsymbol{\Xi}_2$  as a  $NGk \times N$  matrix of zeros with ones on each country's variables in the respective columns,  $\boldsymbol{\theta}_{2,t}$  becomes an  $N \times 1$  vector that captures common movements within each country, implying that the second term:  $\mathbf{X}_t \boldsymbol{\Xi}_2 \boldsymbol{\theta}_{2,t}$ , can be interpreted as *country-specific leading indicators*

(CLI). Since it provides a measure of the business cycle within each country, the CLI is particularly important in addressing our research question. Finally, by defining  $\Xi_3$  as a  $NGk \times G$  matrix of zeros with ones on each variable in the respective columns,  $\theta_{3,t}$  is a  $G \times 1$  vector that captures common movements among variables, implying that the final term:  $\mathbf{X}_t \Xi_3 \theta_{3,t}$  can be interpreted as a vector of *variable specific leading indicators* (VLI).

To complete the state space specification of the time-varying coefficient PVAR model, we need to describe the law of motion for the latent factors (i.e.,  $\theta_i$ ,  $i = 1, 2, 3$ ) along with their prior distributions. To this end, we follow [Canova et al. \(2007\)](#) and specify the state equations for dates  $t = 2, \dots, T$ , as random walks:

$$\theta_t = \theta_{t-1} + \eta_t, \eta_t \sim N(\mathbf{0}, \Omega). \quad (5)$$

Economically, the random walk specification is useful because it allows for the possibility of permanent shifts in the relationships between macroeconomic variables. Computationally, it has the advantage of parsimony. One difference between our specification here and that in [Canova et al. \(2007\)](#) is that  $\Omega$  is a full matrix. This allows for the possibility that the shocks to the time-varying coefficients are contemporaneously correlated. For instance, a shock to the world leading indicator may also generate contemporaneous movements in the country-specific and variable-specific leading indicators too.

Before continuing, we think that it's useful to illustrate the structure of the above indicators. To this end, consider a two country, two variable version of the TVP-PVAR model. The associated factorization in (3) is given by:

$$\beta_t = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \theta_{1t} + \begin{bmatrix} \nu_1 & \mathbf{0} \\ \nu_1 & \mathbf{0} \\ \mathbf{0} & \nu_2 \\ \mathbf{0} & \nu_2 \end{bmatrix} \theta_{2t} + \begin{bmatrix} \nu_3 & \mathbf{0} \\ \mathbf{0} & \nu_4 \\ \nu_3 & \mathbf{0} \\ \mathbf{0} & \nu_4 \end{bmatrix} \theta_{3t}, \quad (6)$$

where  $\nu_1 = [1, 1, 0, 0, 0]'$ ,  $\nu_2 = [0, 0, 1, 1, 0]'$ ,  $\nu_3 = [1, 0, 1, 0, 0]'$ ,  $\nu_4 = [0, 1, 0, 1, 0]'$  and  $\mathbf{0}$  are all  $5 \times 1$  zero vectors. Note that  $\Xi_1$  is a  $20 \times 1$  vector, while  $\Xi_2$  and  $\Xi_3$  are distinct  $20 \times 2$  matrices.

Moreover, if we let  $y_{it}$  and  $x_{it}$ ,  $i = 1, 2$ , denote the variables for each country (i.e., components of  $\mathbf{Y}_{it}$ ), then the associated Eq (4) is given by:

$$\begin{bmatrix} y_{1t} \\ x_{1t} \\ y_{2t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} z_{0,t} \\ z_{0,t} \\ z_{0,t} \\ z_{0,t} \end{bmatrix} \boldsymbol{\theta}_{1t} + \begin{bmatrix} z_{1,t} & 0 \\ z_{1,t} & 0 \\ 0 & z_{2,t} \\ 0 & z_{2,t} \end{bmatrix} \boldsymbol{\theta}_{2t} + \begin{bmatrix} z_{3,t} & 0 \\ 0 & z_{3,t} \\ z_{4,t} & 0 \\ 0 & z_{4,t} \end{bmatrix} \boldsymbol{\theta}_{3t} + \mathbf{u}_t, \quad (7)$$

where and  $z_{0,t} = y_{1t-1} + x_{1t-1} + y_{2t-1} + x_{2t-1} + 1$ ,  $z_{1,t} = y_{1t-1} + x_{1t-1}$ ,  $z_{2,t} = y_{2t-1} + x_{2t-1}$ ,  $z_{3,t} = y_{1t-1} + y_{1t-1}$  and  $z_{4,t} = x_{1t-1} + x_{2t-1}$ . This structure can then be generalized to our six country, seven variable framework.

### 2.2.1. The Common Stochastic Volatility Factor

To allow for the possibility of volatility clustering across macroeconomic shocks, we also include a common stochastic volatility factor. Following [Poon \(2018\)](#), this is done by specifying:

$$\mathbf{u}_t \sim N(\mathbf{0}, e^{h_t} \boldsymbol{\Sigma}_u), \quad (8)$$

where  $\mathbf{u}_t$  is the disturbance term in (4),  $\boldsymbol{\Sigma}_u$  is an  $NG \times NG$  covariance matrix and  $e^{h_t}$  is a latent stochastic volatility factor which is common to all countries. Consistent with both [Poon \(2018\)](#), and other recent studies on large Bayesian VARs (e.g. [Carriero et al. \(2016\)](#) and [Chan \(2018\)](#)), the state equation for the latent log-volatility;  $h_t$ , is set to be a stationary  $AR(1)$  process:

$$h_t = \rho h_{t-1} + \xi_t, \quad \xi_t \sim N(0, \sigma_h^2), \quad (9)$$

where  $|\rho| < 1$ .

### 2.2.2. Priors

To complete the model specification we assume the following independent prior distributions:

$$\boldsymbol{\Sigma}_u \sim IW(\nu_1, \mathbf{S}_1), \quad (10)$$

$$\boldsymbol{\Omega} \sim IW(\nu_2, \mathbf{S}_2), \quad (11)$$

$$\sigma_h^2 \sim IG(\nu_3, S_3), \quad (12)$$

where  $IW(\nu, \mathbf{S})$  denotes the inverse-Wishart distribution with degree of freedom parameter  $\nu > m - 1$ , and positive definite scale matrix  $\mathbf{S}$  and  $IG(\nu, S)$  denotes the inverse-Gamma distribution with shape parameter  $\nu > 0$  and scale parameter  $S > 0$ .

It is well-known that empirical results from Bayesian estimation can be sensitive to the choice of hyperparameters in these prior distributions (see, e.g. [Amir-Ahmadi et al., 2018](#), and references therein). To mitigate this problem, we follow [Amir-Ahmadi et al. \(2018\)](#) and adopt the following hierarchical modeling approach.<sup>7</sup> First, for the measurement covariance matrix, we define  $\mathbf{S}_1 = \nu_1 k_1^2 \mathbf{V}_1$ , in which we set  $\nu_1 = NG$ ,  $\mathbf{V}_1 = Var(\hat{\boldsymbol{\Sigma}}_{OLS})$ , where  $Var(\hat{\boldsymbol{\Sigma}}_{OLS})$  is the OLS estimate of the time-invariant model error covariance matrix, and estimate  $k_1$  within the MCMC algorithm. Second, for the covariance of the time-varying coefficients, we define  $\mathbf{S}_2 = \nu_2 \mathbf{k}_2 \mathbf{V}_2 \mathbf{k}_2$ , in which we set  $\nu_2 = 40$ ,  $\mathbf{V}_2 = Var(\hat{\boldsymbol{\theta}}_{OLS})$ , where  $Var(\hat{\boldsymbol{\theta}}_{OLS})$  is the OLS estimate of the variance-covariance matrix of the parameters in the time-invariant model, and  $\mathbf{k}_2$  is a  $m \times m$  diagonal matrix in which the diagonal elements are estimated. Third, for the variance of the log-volatility, we set  $\nu_3 = k_3^2$  and  $S_3 = k_3^2 V_3$ , in which we set  $V_3 = .01$  and estimate  $k_3$ . Following [Amir-Ahmadi et al. \(2018\)](#), we draw  $k_1$ ,  $\mathbf{k}_2$  and  $k_3$ , with a random walk Metropolis-Hasting step. To that end, let  $\boldsymbol{\kappa} = (k_1, diag(\mathbf{k}_2)', k_3)'$ , where  $diag(\mathbf{k}_2)$  is a column vector consisting of the diagonal elements of  $\mathbf{k}_2$ , making  $\boldsymbol{\kappa}$  a  $(m + 2) \times 1$  vector. The  $j$ -th element of  $\boldsymbol{\kappa}$ ,  $\kappa_j$ , has an independent inverse-Gamma prior of the form:

---

<sup>7</sup>In the online appendix we provide model results with conventional user choices of hyperparameters in time-varying PVAR models as in [Poon \(2018\)](#). The results suggest that such choices may provide different results compared to the modern hierarchical prior framework. Specifically, the time-varying parameter results under the hierarchical prior framework are smoother, and therefore show less time variation compared to conventional user choices. Interestingly, the estimated common stochastic volatility factor is almost identical under both estimation methods.

$$\kappa_j \sim IG(1, 0.1). \quad (13)$$

Finally, the states in the conditional mean are initialized with  $\boldsymbol{\theta}_1 \sim N(\boldsymbol{\theta}_0, \mathbf{V}_\theta)$  where  $\boldsymbol{\theta}_0$  and  $\mathbf{V}_\theta$  are known hyperparameters. We specify a non-informative prior by centering the distribution at  $\boldsymbol{\theta}_0 = \mathbf{0}$  and setting a large variance  $\mathbf{V}_\theta = 10\mathbf{I}_{NG}$ . The states of the conditional variance are initialized according to the stationary distribution of the first-order autoregressive process:  $h_1 \sim N\left(h_0, \frac{\sigma_h^2}{(1-\rho)^2}\right)$ , where  $h_0 = 0$  is a known hyperparameter and  $\sigma_h^2$  and  $\rho$  are drawn from their prior distributions. To ensure stationarity of the states we set  $\rho \sim N(\mu_\rho, V_\rho) \mathbf{1}(|\rho| < 1)$  where  $\mathbf{1}(\cdot)$  denotes the indicator function  $\mu_\rho = 0$  and  $V_\rho = 1$ . Estimation details are deferred to [Appendix A.1](#).

### 2.3. Model Selection

Since our model has many bells and whistles, we first conduct a model selection exercise before taking it to the data. Given our Bayesian estimation framework, this can be accomplished through a direct comparison of each model's marginal likelihood via the *Bayes factor*. In what follows we explain how to compute the Bayes factor for arbitrary models. The results are then presented in Section 3.1.

Let  $M := \{M_1, \dots, M_{\mathcal{I}}\}$  denote the set of models that we wish to compare and  $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 & \dots & \mathbf{Y}_T \end{bmatrix}'$  denote the vector of observations over the full sample. The posterior probability of model  $M_i$ , denoted  $p(M_i|\mathbf{Y})$ , is calculated using Bayes Theorem:

$$p(M_i|\mathbf{Y}) = \frac{p(\mathbf{Y}|M_i)p(M_i)}{p(\mathbf{Y})}, \quad (14)$$

where  $p(\mathbf{Y}|M_i)$  denotes the marginal (model) likelihood,  $p(M_i)$  denote the prior probability of the  $i$ -th model and  $p(\mathbf{Y})$  denotes the marginal data density. The marginal likelihood is calculated by integrating over the parameter space of a given model:

$$p(\mathbf{Y}|M_i) = \int p(\mathbf{Y}|M_i, \Theta_i) p(\Theta_i|M_i) d\Theta_i, \quad (15)$$

where  $\Theta_i$  represents a model-specific parameter vector, while  $p(\mathbf{Y}|\Theta_i, M_i)$  and  $p(\Theta_i|M_i)$  respectively denote the associated likelihood and prior density functions. Taking the ratio of the posterior

probabilities of two models:  $M_i$  and  $M_j$ , gives the *posterior odds ratio*:

$$\frac{p(M_i|\mathbf{Y})}{p(M_j|\mathbf{Y})} = \frac{p(\mathbf{Y}|M_i)p(M_i)}{p(\mathbf{Y}|M_j)p(M_j)}. \quad (16)$$

To avoid any user specified bias in the calculation of the *posterior odds*, we assign equal prior probabilities to all models, i.e.,  $p(M_i) = p(M_j) = \frac{1}{2}$ . Under this assumption, (16) can be written as:

$$\frac{p(M_i|\mathbf{Y})}{p(M_j|\mathbf{Y})} = \frac{p(\mathbf{Y}|M_i)}{p(\mathbf{Y}|M_j)}. \quad (17)$$

The likelihood ratio on the right hand side of (17) is commonly referred to as *Bayes factor* for  $M_i$  against  $M_j$ , and is denoted  $BF_{ij}$ . Since the Bayes factor is a posterior odds ratio, it takes on a probabilistic interpretation. For instance, if  $BF_{ij} = 2$  then conditional on the data,  $M_i$  is twice as likely as compared to  $M_j$ .

To compute the Bayes factor we first need to compute the marginal likelihood of each model. To this end, we follow [Geweke and Amisano \(2011\)](#) and compute the one-step-ahead predictive likelihood for which details are deferred to [Appendix A.2](#). The reason for using the one-step-ahead predictive likelihood instead of Chib's marginal likelihood method as in [Canova et al. \(2007\)](#), or the harmonic mean estimator as in [Canova and Ciccarelli \(2012\)](#), is that recent work has shown that these methods can be extremely inaccurate. Specifically, [Chan and Grant \(2015\)](#) show that the (modified) harmonic mean as in [Gelfand and Dey \(1994\)](#) can have a substantial finite sample bias and can thus lead to inaccurate model selection, while [Frühwirth-Schnatter and Wagner \(2008\)](#) provide similar inference for Chib's marginal likelihood method ([Chib, 1995](#)).

### 3. Results

Our presentation of the main empirical results is partitioned by three themes. First, in [Section 3.1](#) we present the results from our model selection exercise. Second, in [3.2](#) we discuss the economic significance of the model's indicators and common stochastic volatility factor. Finally, in [Section 3.3](#) we address our primary research question surrounding the extent to which international shocks drive Australian business cycle fluctuations.

### 3.1. Model Selection

To commence the analysis we first compare the marginal likelihood of the baseline TVP-PVAR model discussed in Section 2.2 against various alternative specifications. First, since it is not economically relevant when answering our research question, we consider a version of each model with no variable specific indicator—i.e., a TVP-PVAR model with only common and country-specific indicators. Second, instead of using a single common indicator, we consider the case of regional and country size indicators. These specifications enable us to answer the question: “Are business cycles of the selected countries similar on a regional or country size basis?”. Such a question is not only interesting in its own right but is also necessary when addressing our main research question. As discussed in [Canova et al. \(2007\)](#), if regional business cycles deviate from the rest of the world cycles, then failing to account for this feature – through the inclusion of only one world indicator – may spuriously generate significant country-specific indicators because of omitted variable bias. The same logic applies to models that omit an indicator designed to capture distinct cycles for small and large economies. To create the model with regional cycles, we partition the set of countries by continent (i.e., Asia, Europe, The Americas, and Oceania). Since we already have country-specific indicators for the EU, the US, and Australia, this grouping only requires regional indicators for Asian countries (i.e., China, Japan, and Korea). Third, to assess the significance of distinct common behavior between small and large economies, we partition the set of countries by size, with Australia and Korea constituting the set of small open economies. Finally, given its extraordinary macroeconomic performance over the sample period, we exclude China from the rest of the countries when specifying the WLI, and instead treat it as a separate “China-specific indicator” (CSI). The marginal likelihoods of each model against the baseline TVP-PVAR model are presented in Table 1. Given our relatively small sample period, we follow [Canova et al. \(2007\)](#); [Canova and Ciccarelli \(2009\)](#); [Poon \(2018\)](#) and mitigate the risk of over-parameterization by using a lag length of one in each model.

Table 1: Two times log marginal likelihoods and associated Bayes factors subject to a Gaussian TVP-PVAR benchmark

<b>Specification</b>	$2\log ML_i$	<b>(log)</b> $BF_{i,TVP-PVAR}$
Baseline 1	-5170,6	0
Baseline 2	-4811,8	358,8
Regional Indicator 1	-5155,0	15,6
Regional Indicator 2	-4874,4	296,2
Country Size Indicator 1	-5118,6	52
Country Size Indicator 2	-4927,0	243,6
Baseline 1 with CSI	-4869,5	301,1
Baseline 2 with CSI	-5154,1	16,5

Note: The baseline model is the TVP-PVAR model discussed in Section 2. Model variant 1 refers to models with a variable leading indicator and variant 2 refers to models without a variable leading indicator.

Following [Kass and Raftery \(1995\)](#), a (log-) Bayes factor of between 0 and 2 is “not worth more than a bare mention”, while values between 2 and 6 provide “positive” evidence, 6 to 10 provide “strong” evidence and greater than 10 provide “very strong” evidence in favor of  $M_i$  against the baseline model. The results therefore provide very strong evidence that the baseline 2 TVP-PVAR model, i.e., the one without a variable specific leading indicator, is the best. The finding that the country size and regional indicator models do not outperform the baseline model, suggests that neither regional business cycles nor country size business cycles are significant among the nations in our sample. This result is consistent with [Kose et al. \(2003\)](#) who use a sample of 76 countries to show that, except for the North American region, regional business cycles do not play an important role in explaining aggregate world economic volatility. The comparisons of the baseline model with and without China in the common factor are also clear. While the Baseline 1 model with CSI is strongly preferred to the associated Baseline 1 model with the CSI in the common factor, the Baseline 2 model is still strongly preferred by the data. For the remainder of the paper, we will consequently use the Baseline 2 model in the analysis.

### 3.2. Significance of the Indicators

While the Bayes factor is a useful model comparison tool, it provides no inference on whether the model’s parameters at each date in the sample period are statistically different from zero. In this section, we, therefore, examine the world-leading indicator (WLI), the country leading

indicator (CLI) and the common stochastic volatility factor over the sample period. To start, the posterior median and 68percent posterior credible set for the WLI are reported in Figure 2.<sup>8</sup>

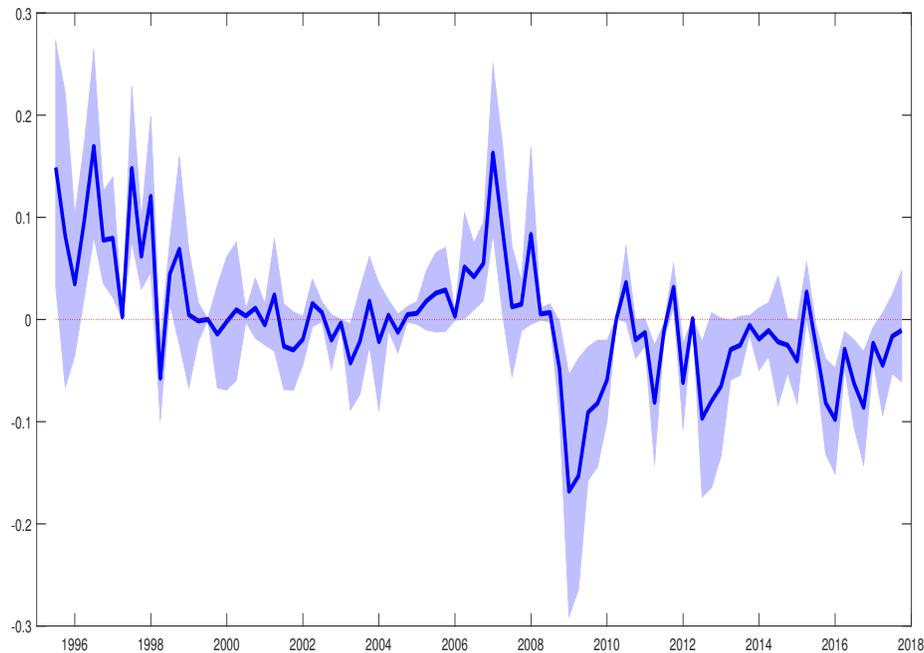


Figure 2: World Leading Indicator (WLI): Posterior median (blue line) and 68percent credible interval (shaded area).

Despite being significant from a model comparison perspective, the posterior distribution of the WLI contains zero for much of the sample. This implies that the indicator is not statistically different from zero for these parts of the sample. We emphasize that this not imply that there were no meaningful global economic events within these periods. Instead, it implies that there was no significant co-movement across all included countries in these periods, at least from a statistical perspective. For example, notable periods of significant co-movement exist around the 1997/98 Asian Financial Crisis and the 2007/08 Global Financial Crisis. In contrast, there is no significant co-movement during the “dotcom bubble” in 2000. To further investigate the behavior of each

---

<sup>8</sup>We provide a 95percent credible interval for each of the figures in the online appendix. The results are similar to those presented here. Interestingly, the common factor becomes less significant when using a two standard deviation credible set, however, the country-specific factors and stochastic volatility factor remain significant.

country during these periods, we now consider each country's real GDP growth data - presented in Figure 3.

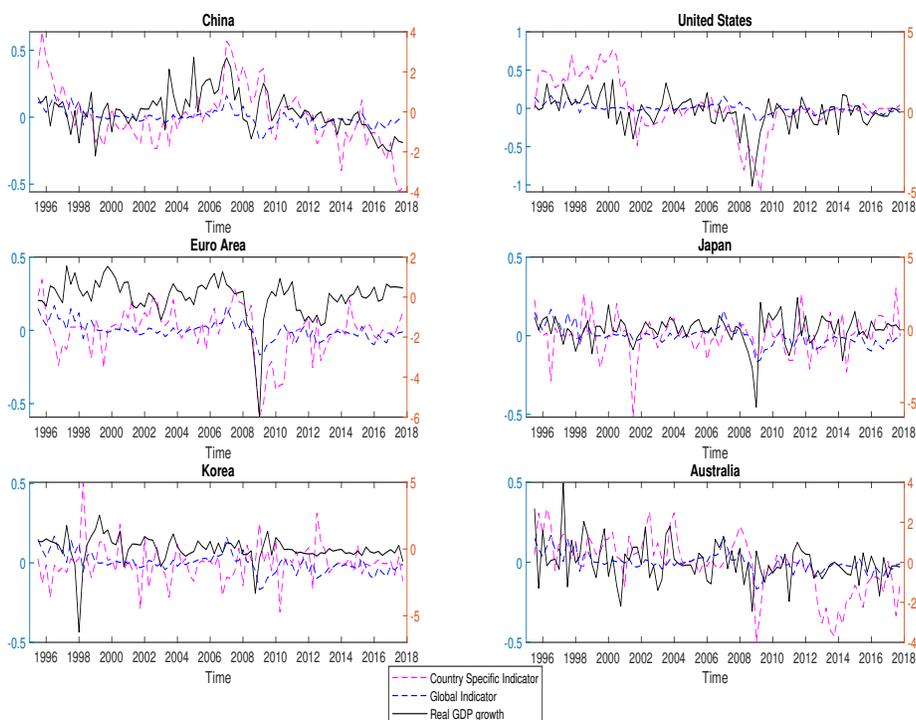


Figure 3: country-specific real GDP growth and posterior median values of the global indicator and country-specific indicator.

In addition to containing real GDP growth time series for each country, Figure 3 also presents the global indicator and country-specific indicator. Consistent with the analysis in [Canova et al. \(2007\)](#), the global indicator tends to track low-frequency movements in the business cycles of the respective nations, while the country-specific indicators replicate more high-frequency-type fluctuations. In particular, while the median of the CLI under predicts those frequencies in the EU, it generally does an excellent job of tracking the real GDP growth of the respective nations. This includes the systematic collapse around the Global Financial Crisis along with the recent slowdown in China's macroeconomic growth since 2015.

Finally, to assess the presence of volatility clustering, Figure 4 presents the posterior median and 68 percent credible interval over the sample period. We highlight two key points. First, since

the credible interval is above zero, from a statistical perspective we conclude that the common stochastic volatility factor is significant across the entire sample. Second, the common stochastic volatility factor can detect two key periods of global economic turbulence over the past two decades. The first is the 1997 Asian Financial Crisis (AFC) and the second is the recent 2007/08 Global Financial Crisis (GFC). While the causes and consequences of these two crises are beyond the scope of this paper (see e.g. [Corsetti et al. \(1999\)](#) for the former and [Mian and Sufi \(2010\)](#) for the latter), a key result of the paper is that the inclusion of the common stochastic volatility factor within the panel VAR framework allows us to detect important structural instabilities that an otherwise constant volatility model would fail to detect.

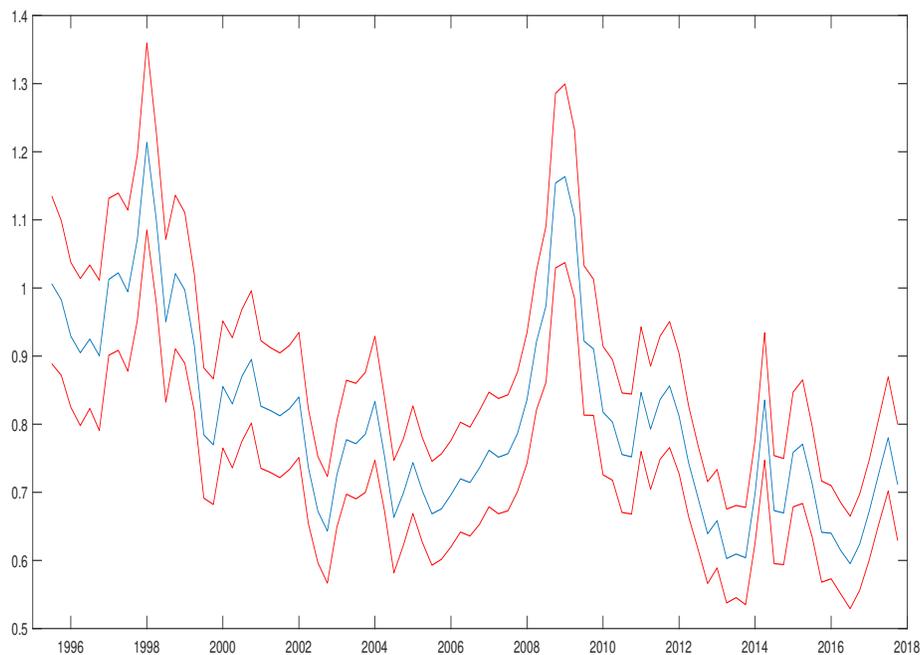


Figure 4: Common stochastic volatility factor: Posterior median (blue line) and 68 percent credible interval (red lines).

### 3.3. Drivers of Australian Business Cycle Fluctuations

In this section, we address our main research objective of quantifying the proportion of business cycle fluctuations attributable to international shocks from Australia's five largest trading

partners: China, Japan, the EU, Korea and the US. To this end, we use a historical decomposition. To construct the decomposition, the model's disturbances are orthogonalized using sign and magnitude restrictions. In the first step, we assume that a positive shock from each of Australia's trading partners results in a non-negative response to Australian real GDP growth. Next, to overcome issues regarding partial identification, we place a magnitude restriction that each country's response to its own shock is larger than its response to any other country's shock. In the online appendix, we provide impulse response functions (IRFs) that exhibit reasonable behavior. For instance, the largest impacts are idiosyncratic, followed by China and the US, with Korea having the smallest impact. For robustness, we also report impulse response functions from a simpler recursive identification strategy in which countries are ordered based on the end of sample trade volumes. While technically simpler, the results suggest that economic expansion in Korea will cause an economic contraction in Australia. This result highlights the importance of using the combination of sign and magnitude restrictions to identify the structural model.<sup>9</sup>

The resulting historical decomposition in Figure 5 shows that international shocks have had significant effects on the Australian business cycle and that these effects have varied over the past two decades. Particularly notable contributions come from China and Japan, while contributions from the remaining countries are relatively modest. There is also substantial time variation in the contribution of each shock. For instance, contributions from the US were mostly positive before the 2007/08 GFC and have since been negative. Also, Korea had a strong influence around the 1997/98 AFC and has had little influence in the most recent decade. In line with the mining boom of the mid-2000s, positive contributions from China are found to have been particularly important around the 2007/08 GFC and appear to have greatly contributed towards offsetting the negative impacts from the US. Our results, therefore, support the idea that one of the reasons Australia avoided recession in this period was at least partially because of strong positive spillovers from China, however positive contributions from Japan, Korea, and the EU also played an important role.

While we are the first paper to consider contributions from this set of countries, [Dungey et al.](#)

---

<sup>9</sup>We are extremely grateful to a referee for this suggestion.

(2014) provide a similar analysis with the EU and the US, while Liu (2010) and Nimark (2009) independently examine the G7 economies as a single conglomerate. While the results here are not directly comparable to Liu (2010) and Nimark (2009), we can draw some comparisons with Dungey et al. (2014). Consistent with their results, we find that the EU and the US respectively have negative and positive contributions during the mid to late 1990s. Also consistent with their result is the fact that US output shocks to variation in Australian output, had a greater impact than the EU in this period. In contrast with their results, however, we find that the domestic contribution is mostly positive during this period. This may highlight the fact that our specification also includes China, which had some significant negative impacts on the Australian economy during the Asian financial crisis.

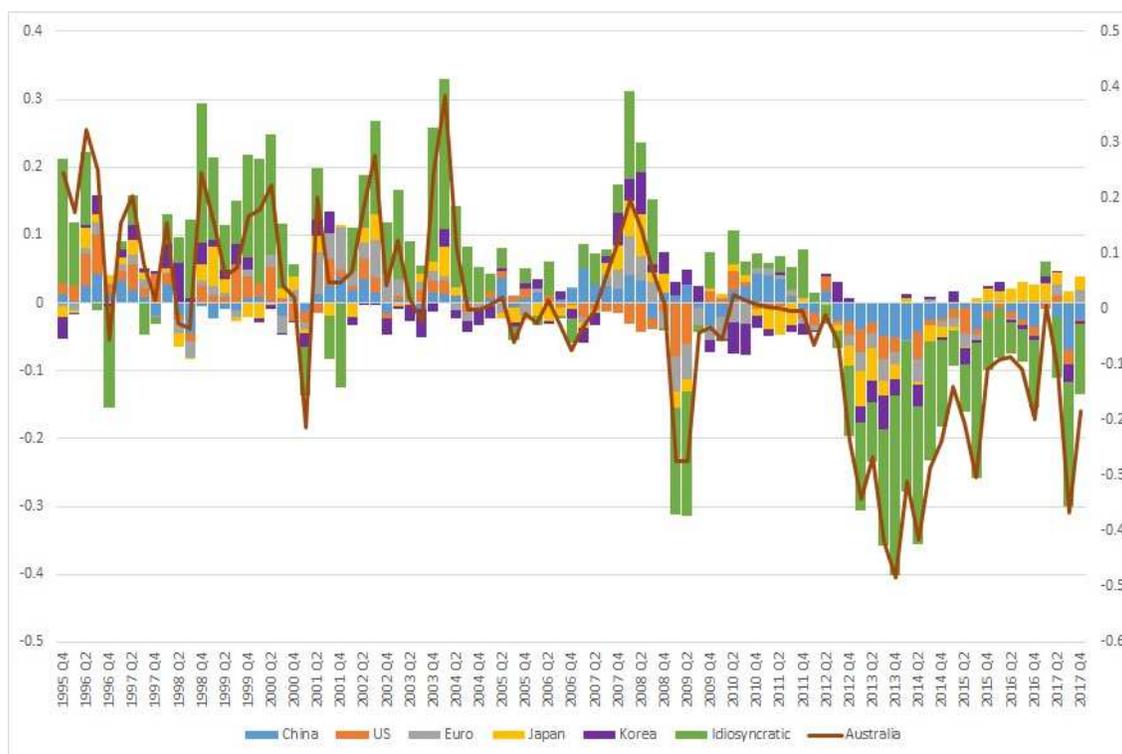


Figure 5: Historical decomposition of Australian business cycle

While this time variation is important, it is also useful to present the average contribution of each shock. To this end, we use an indicator function to partition the historical decomposition in Figure 5 into positive and negative Australian real GDP fluctuations over the sample period. We

then average the contribution of each country over the sample period (as defined by the height of the bars in Figure 5). The results are presented in Table 2. Since the trade statistics in Figure 1 revealed that the quarterly bilateral trade balance of goods and services between Australia and each of these countries has substantially changed over our sample period, we also consider sub-sample results. Specifically, we split the sample into three periods: pre-2000, between 2000 and 2007 and 2008 to 2017.

Table 2: Average country-specific percentage contributions to positive, negative and aggregate Australian business cycle fluctuations over the sample period

	<b>China</b>	<b>US</b>	<b>EU</b>	<b>Japan</b>	<b>Korea</b>	<b>Australia</b>
<i>Aggregate</i>						
Pre-2000	13	16	6	9	11	44
2000-2007	11	10	11	11	12	45
2008-2017	15	10	10	12	11	42
Full Sample	13	11	10	11	11	44
<i>Positive</i>						
Pre-2000	10	21	5	9	13	42
2000-2007	16	10	11	6	7	50
2008-2017	11	5	7	24	20	33
Full Sample	13	10	8	14	14	41
<i>Negative</i>						
Pre-2000	28	1	10	14	4	43
2000-2007	3	17	11	19	34	15
2008-2017	14	18	11	9	9	39
Full Sample	13	14	11	14	17	31

To answer our primary research question, we sum the aggregate contributions of all international shocks to get a measure of the total contribution of international shocks. Based on this measure, international shocks from Australia's top five trading partners are responsible for approximately half (56 percent) of all Australian business cycle fluctuations over the past two decades. While this aggregate result is consistent with earlier works by Liu (2010) and Nimark (2009) an important novelty of our results is that the country-by-country composition of the total is far from stable. For instance, over the past two decades, the contribution of international shocks stemming from the US is found to have declined, while contributions from China, Japan, and the EU have increased. Interestingly, contributions from Korea have been constant over the sample period.

Another novelty of our results is that the positive and negative contributions of international shocks on Australia's business cycle differ across countries. For instance, the large negative contributions from Japan and China in the pre-2000 era are likely due to the Asian Financial Crisis. In comparison, both nations have had a net positive effect on the Australian economy in the most recent decade. For instance, China's positive contribution increased six percent throughout the mining boom of the 2000s, while Japan's positive influence has been especially strong since the 2007/08 GFC. In contrast, the positive US influence has almost halved over the sample period, while the negative impacts have grown significantly. Taken together, the results highlight the dynamic nature of Australia's trade relationships and suggest that Australian policymakers should be particularly mindful of international economic events when making decisions.

#### 4. Conclusion

Our objective in this paper was to quantify the proportion of Australian business cycle fluctuations that can be attributed to international shocks stemming from its five largest trading partners: China, Japan, the EU, the US and the Republic of Korea. To that end, we used a panel VAR model, that allowed for time-varying parameters and common stochastic volatility. In the context of our research question, the advantage of the panel VAR over a traditional VAR framework is that it allowed us to parsimoniously model intra- and inter-country variable interdependencies as well as cross-sectional heterogeneity. The common stochastic volatility factor was useful because it captured various episodes of volatility clustering around the 1997/98 Asian Financial Crisis and the 2007/08 Global Financial Crisis, which a constant volatility model would fail to detect. The time-varying parameter dimension of the model was also able to capture the inter-temporal nature of Australia's various bilateral trade relationships. For instance, the effects of aggregate international shocks stemming from Europe have doubled since the turn of the century, while those from the US have halved. Our main result was that international shocks are responsible for around half of all Australian business cycle fluctuations over the sample period. We also found an important asymmetry in the effects of international shocks stemming from each country. For instance, China's positive contribution increased by six percent throughout the mining boom of the early 2000s, while the US influence has almost halved over the sample period. Taken together,

the results highlight the dynamic nature of Australia's trade relationships and suggest that Australian policymakers should be particularly mindful of international economic events when making decisions.

## References

- Amir-Ahmadi, P., Matthes, C., Wang, M.C., 2018. Choosing prior hyperparameters: with applications to time-varying parameter models. *Journal of Business & Economic Statistics* , 1–13.
- Australian Department of Foreign Affairs and Trade (DFAT), 2017. Australia's trade in goods and services 2014-15.
- Bañbura, M., Giannone, D., Reichlin, L., 2010. Large bayesian vector auto regressions. *Journal of Applied Econometrics* 25, 71–92.
- Canova, F., Ciccarelli, M., 2009. Estimating multicountry var models. *International Economic Review* 50, 929–959.
- Canova, F., Ciccarelli, M., 2012. Clubmed? cyclical fluctuations in the mediterranean basin. *Journal of International Economics* 88, 162–175.
- Canova, F., Ciccarelli, M., 2013. Panel vector autoregressive models: A survey, in: *VAR Models in Macroeconomics—New Developments and Applications: Essays in Honor of Christopher A. Sims*. Emerald Group Publishing Limited, pp. 205–246.
- Canova, F., Ciccarelli, M., Ortega, E., 2007. Similarities and convergence in g-7 cycles. *Journal of Monetary economics* 54, 850–878.
- Canova, F., Ciccarelli, M., Ortega, E., 2012. Do institutional changes affect business cycles? evidence from europe. *Journal of Economic Dynamics and Control* 36, 1520–1533.
- Carriero, A., Clark, T.E., Marcellino, M., 2016. Common drifting volatility in large bayesian vars. *Journal of Business & Economic Statistics* 34, 375–390.

- Chan, J.C., 2018. Large bayesian vars: A flexible kronecker error covariance structure. *Journal of Business & Economic Statistics* , 1–29.
- Chan, J.C., Grant, A.L., 2015. Pitfalls of estimating the marginal likelihood using the modified harmonic mean. *Economics Letters* 131, 29–33.
- Chan, J.C., Hsiao, C.Y., 2014. Estimation of Stochastic Volatility Models with Heavy Tails and Serial Dependence. John Wiley and Sons Inc. pp. 155–176. doi:[10.1002/9781118771051.ch6](https://doi.org/10.1002/9781118771051.ch6).
- Chan, J.C., Jeliaskov, I., 2009. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation* 1, 101–120.
- Chib, S., 1995. Marginal likelihood from the gibbs output. *Journal of the American Statistical Association* 90, 1313–1321.
- Chib, S., Greenberg, E., 1995. Hierarchical analysis of sur models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics* 68, 339–360.
- Claus, E., Dungey, M., Fry, R., 2008. Monetary policy in illiquid markets: Options for a small open economy. *Open Economies Review* 19, 305–336.
- Corsetti, G., Pesenti, P., Roubini, N., 1999. What caused the asian currency and financial crisis? *Japan and the world economy* 11, 305–373.
- De Mol, C., Giannone, D., Reichlin, L., 2008. Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics* 146, 318–328.
- Dees, S., Mauro, F.d., Pesaran, M.H., Smith, L.V., 2007. Exploring the international linkages of the euro area: a global var analysis. *Journal of applied econometrics* 22, 1–38.
- Dungey, M., Fry, R., 2003. International shocks on australia-the japanese effect. *Australian Economic Papers* 42, 158–182.

- Dungey, M., Osborn, D., Raghavan, M., 2014. International transmissions to australia: the roles of the usa and euro area. *Economic Record* 90, 421–446.
- Dungey, M., Pagan, A., 2000. A structural var model of the australian economy. *Economic record* 76, 321–342.
- Dungey, M., Pagan, A., 2009. Extending a svar model of the australian economy. *Economic Record* 85, 1–20.
- Frühwirth-Schnatter, S., Wagner, H., 2008. Marginal likelihoods for non-gaussian models using auxiliary mixture sampling. *Computational Statistics & Data Analysis* 52, 4608–4624.
- Gelfand, A.E., Dey, D.K., 1994. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)* , 501–514.
- Geweke, J., Amisano, G., 2011. Hierarchical markov normal mixture models with applications to financial asset returns. *Journal of Applied Econometrics* 26, 1–29.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *Journal of the american statistical association* 90, 773–795.
- Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies* 65, 361–393.
- Kose, M.A., Prasad, E.S., Terrones, M.E., 2003. How does globalization affect the synchronization of business cycles? *American Economic Review* 93, 57–62.
- Kroese, D.P., Chan, J.C., et al., 2014. *Statistical modeling and computation*. Springer.
- Leu, S.C.Y., 2011. A new keynesian svar model of the australian economy. *Economic Modelling* 28, 157–168.
- Leu, S.C.Y., Sheen, J., 2011. A small new keynesian state space model of the australian economy. *Economic Modelling* 28, 672–684.

- Liu, P., 2010. The effects of international shocks on australia’s business cycle. *Economic Record* 86, 486–503.
- Mian, A., Sufi, A., 2010. The great recession: Lessons from microeconomic data. *The American Economic Review* 100, 51–56.
- Mumtaz, H., Surico, P., 2009. The transmission of international shocks: a factor-augmented var approach. *Journal of Money, Credit and Banking* 41, 71–100.
- Nimark, K.P., 2009. A structural model of australia as a small open economy. *Australian Economic Review* 42, 24–41.
- Poon, A., 2017. The transmission mechanism of malaysian monetary policy: A time-varying vector autoregression approach. *Empirical Economics*. Forthcoming. .
- Poon, A., 2018. Assessing the synchronicity and nature of australian state business cycles. *Economic Record*, Forthcoming .
- Sheng, Y., Song, L., 2008. Comparative advantage and australia-china bilateral trade. *Economic Papers: A journal of applied economics and policy* 27, 41–56.
- Voss, G., Willard, L., 2009. Monetary policy and the exchange rate: Evidence from a two-country model. *Journal of Macroeconomics* 31, 708–720.

## **Appendix A. Appendices**

### *Appendix A.1. Markov Chain Monte Carlo Algorithm*

In this appendix we present the Markov chain Monte Carlo (MCMC) algorithm used to estimate the TVP-PVAR model. All alternative models considered in this paper are nested versions of this model and can therefore be estimated by straightforward modifications of the MCMC procedure. To detail the MCMC procedure, let  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_T)'$ ,  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T)'$ ,  $\mathbf{h} = (\mathbf{h}_1, \dots, \mathbf{h}_T)'$  and  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_T)'$ . The posterior draws are obtained through a six-block Metropolis-within-Gibbs sampler, that sequentially samples each variable from their respective full conditional distribution:

1. Draw from  $p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{h}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$
2. Draw from  $p(\boldsymbol{\Sigma}_u \mid \mathbf{Y}, \boldsymbol{\theta}, \mathbf{h}, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$
3. Draw from  $p(\boldsymbol{\Omega} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \rho, \sigma_h^2, \boldsymbol{\kappa})$
4. Draw from  $p(\mathbf{h} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$
5. Draw from  $p(\rho \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \sigma_h^2, \boldsymbol{\kappa})$
6. Draw from  $p(\sigma_h^2 \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \rho, \boldsymbol{\kappa})$
7. Draw from  $p(\boldsymbol{\kappa} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \rho, \sigma_h^2)$

In our analysis, we use 35,000 posterior draws, discarding the first 15,000 to allow for convergence of the Markov chain to its stationary distribution. Under the previously defined conjugate priors, Steps 2, 3 and 6 can be directly sampled from their resulting posterior distributions. Next, following [Canova et al. \(2007, 2012\)](#), the latent states in Step 1 can be sampled using standard Kalman filter algorithms as in [Chib and Greenberg \(1995\)](#). In this paper, we instead follow [Poon \(2018\)](#) and make use of an efficient *precision sampling* algorithm. The increased efficiency of the algorithm comes from exploiting the fact that the precision matrices of the latent states are block-banded and sparse. This means that computational savings can be made in necessary operations when solving linear systems – such as taking a Cholesky decomposition, matrix multiplication and forward-backward substitution. Next, Step 4 involves a non-linear non-Gaussian measurement equation, and the standard linear Kalman filter can not be applied. To overcome this issue we follow [Poon \(2017\)](#) and make use of the *auxiliary mixture sampler* developed by [Kim et al. \(1998\)](#) along with an efficient sampling algorithm in [Chan and Hsiao \(2014\)](#). The auxiliary mixture sampler uses a seven-Gaussian mixture to convert the non-linear measurement equation in the stochastic volatility model, into a log-linear equation that is conditionally Gaussian. [Chan and Hsiao \(2014\)](#) then show how to adapt to sample the log-volatilities using an efficient *precision sampling* algorithm. Finally, the full conditional distribution in Steps 5 results in non-standard distribution. Sampling is therefore achieved through an independence-chain Metropolis-Hastings Algorithm adapted from the univariate models in [Chan and Hsiao \(2014\)](#). For completeness, we now discuss the derivation of the conditional posterior distribution of each block in the Gibbs sampler. In each case, we use the fact that the posterior distribution for a parameter of interest can be obtained by working with

the kernel of the resulting product from the prior and likelihood functions.

**Step 1: Sample from**  $p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{h}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$

To sample  $\boldsymbol{\theta}$ , first note that (4) can be rewritten as:

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\theta} + \mathbf{u}, \mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \quad (\text{A.1})$$

where  $\mathbf{Z} = \text{diag}(\mathbf{Z}_1, \dots, \mathbf{Z}_T)$ ,  $\mathbf{u} = \begin{bmatrix} \mathbf{u}'_1 & \mathbf{u}'_2 & \dots & \mathbf{u}'_T \end{bmatrix}'$ , and  $\boldsymbol{\Sigma} = \text{diag}(e^{h_1} \boldsymbol{\Sigma}_u, \dots, e^{h_T} \boldsymbol{\Sigma}_u)$ . By a change of variable:

$$\mathbf{Y} \sim N(\mathbf{Z}\boldsymbol{\theta}, \boldsymbol{\Sigma}). \quad (\text{A.2})$$

Next, rewrite (5) as:

$$\mathbf{H}_\theta \boldsymbol{\theta} = \tilde{\boldsymbol{\alpha}}_\theta + \boldsymbol{\eta}, \boldsymbol{\eta} \sim N(\mathbf{0}, \mathbf{S}_\theta), \quad (\text{A.3})$$

where  $\tilde{\boldsymbol{\alpha}}_\theta = \begin{bmatrix} \boldsymbol{\theta}'_0 & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}$ ,  $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_T)'$ ,  $\mathbf{S}_\theta = \text{diag}(\mathbf{V}_\theta, \boldsymbol{\Omega}, \dots, \boldsymbol{\Omega})$  and  $\mathbf{H}_\theta$  is a  $Tm \times Tm$  block diagonal matrix, where  $m = N_1 + N + G$ , with  $\mathbf{I}_m$  on the main diagonal,  $-\mathbf{I}_m$  on the lower diagonal and  $\mathbf{0}_m$  elsewhere. Since  $\mathbf{H}_\theta$  is a lower triangular matrix with ones along the main diagonal,  $|\mathbf{H}_\theta| = 1$ , implying that it is invertible. Using this result, (A.3) can be rewritten as:

$$\boldsymbol{\theta} = \boldsymbol{\alpha}_\theta + \mathbf{H}^{-1}\boldsymbol{\eta}, \quad (\text{A.4})$$

where  $\boldsymbol{\alpha}_\theta = \mathbf{H}_\theta^{-1} \tilde{\boldsymbol{\theta}}_0$ . By a change of variable:

$$\boldsymbol{\theta} \mid \boldsymbol{\Omega} \sim N\left(\boldsymbol{\alpha}_\theta, \left(\mathbf{H}'_\theta \mathbf{S}_\theta^{-1} \mathbf{H}_\theta\right)^{-1}\right). \quad (\text{A.5})$$

Combining (A.2) and (A.5) gives the conditional posterior distribution:

$$\begin{aligned} p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{h}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2) &\propto p(\mathbf{Y} \mid \boldsymbol{\theta}, \mathbf{h}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2) p(\boldsymbol{\theta} \mid \boldsymbol{\Omega}) \\ &\propto \exp\left\{-\frac{1}{2} \left[ (\mathbf{Y} - \mathbf{Z}\boldsymbol{\theta}) \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{Z}\boldsymbol{\theta})' + (\boldsymbol{\theta} - \boldsymbol{\alpha}_\theta) \left( \mathbf{H}'_\theta \mathbf{S}_\theta^{-1} \mathbf{H}_\theta \right) (\boldsymbol{\theta} - \boldsymbol{\alpha}_\theta)' \right]\right\} \\ &\propto \exp\left\{-\frac{1}{2} \left[ \boldsymbol{\theta}' \left( \mathbf{Z}' \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) \boldsymbol{\theta}' - 2\boldsymbol{\theta}' \left( \mathbf{Z}' \boldsymbol{\Sigma}^{-1} \mathbf{Y} + \mathbf{H}'_\theta \mathbf{S}_\theta^{-1} \mathbf{H}_\theta \boldsymbol{\alpha}_\theta \right) \right]\right\} \end{aligned}$$

Thus, by standard linear regression results (Kroese et al., 2014, p.237-240):

$$\left(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{h}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa}\right) \sim N\left(\hat{\boldsymbol{\theta}}, \mathbf{D}_\beta^{-1}\right), \quad (\text{A.6})$$

where  $\hat{\boldsymbol{\theta}} = \mathbf{D}_\beta^{-1} \left(\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} + \mathbf{H}'_\theta \mathbf{S}_\theta^{-1} \mathbf{H}_\theta \boldsymbol{\alpha}_\theta\right)$  and  $\mathbf{D}_\beta = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z} + \mathbf{H}'_\theta \mathbf{S}_\theta^{-1} \mathbf{H}_\theta$ . Following [Poon \(2018\)](#), sampling from this distribution is conducted with the precision sampling algorithm in [Chan and Jeliazkov \(2009\)](#).

**Step 2: Sample from  $p(\boldsymbol{\Sigma}_u \mid \mathbf{Y}, \boldsymbol{\theta}, \mathbf{h}, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$**

To sample,  $\boldsymbol{\Sigma}_u$ , combine the inverse-Wishart prior distribution with the Gaussian likelihood function. Since this is a conjugate distribution it is easy to show that:

$$\left(\boldsymbol{\Sigma}_u \mid \mathbf{Y}, \boldsymbol{\theta}, \mathbf{h}, \sigma_h^2, \boldsymbol{\Omega}, \rho, \boldsymbol{\kappa}\right) \sim IW\left(\nu_2 + T, \sum_{t=1}^T \frac{(\mathbf{Y}_t - \mathbf{Z}_t \boldsymbol{\theta}_t)(\mathbf{Y}_t - \mathbf{Z}_t \boldsymbol{\theta}_t)'}{e^{h_t}} + \nu_2 k_2^2 \mathbf{V}_2\right). \quad (\text{A.7})$$

**Step 3: Sample from  $p(\boldsymbol{\Omega} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \rho, \sigma_h^2, \boldsymbol{\kappa})$**

To sample  $\boldsymbol{\Omega}$ , we combine the inverse-Wishart prior distribution with the Gaussian likelihood function to get

$$\left(\boldsymbol{\Omega} \mid \mathbf{Y}, \boldsymbol{\theta}, \mathbf{h}, \sigma_h^2, \boldsymbol{\Sigma}_u, \rho, \boldsymbol{\kappa}\right) \sim IW\left(\nu_1 + T - 1, \mathbf{H}_\theta \boldsymbol{\theta} \boldsymbol{\theta}' \mathbf{H}'_\theta + \nu_1 \mathbf{k}_1 \mathbf{V}_1 \mathbf{k}_1\right). \quad (\text{A.8})$$

**Step 4: Sample from  $p(\mathbf{h} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \rho, \sigma_h^2, \boldsymbol{\kappa})$**

To sample  $\mathbf{h}$ , we follow [Poon \(2018\)](#) and apply the auxiliary mixture sampler from [Kim et al. \(1998\)](#) along with the precision sampler from [Chan and Hsiao \(2014\)](#). To this end, note that the measurement equation in (4) can be written as:

$$\mathbf{P}^{-1}(\mathbf{Y}_t - \mathbf{Z}_t \boldsymbol{\theta}_t) = e^{\frac{1}{2}h_t} \boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_t \sim N(0, \mathbf{I}_n), \quad (\text{A.9})$$

where  $\mathbf{P}$  is the (lower) Cholesky factor of  $\boldsymbol{\Sigma}_u$ . Squaring both sides of (A.9) and taking the (natural) logarithm gives the log-linear equation:

$$\mathbf{y}_t^* = \boldsymbol{\iota}_{NG} h_t + \boldsymbol{\epsilon}_t^*, \quad (\text{A.10})$$

where  $\mathbf{y}_t^* = \ln \left( \left( \mathbf{P}^{-1} (\mathbf{Y}_t - \mathbf{X}_t \Xi \boldsymbol{\theta}_t) \right)^2 \right)$ ,  $\boldsymbol{\nu}_{NG}$  is a  $NG \times 1$  unit vector and  $\boldsymbol{\epsilon}_t^* = \left[ \ln(\boldsymbol{\epsilon}_{1,t}^2) \ \dots \ \ln(\boldsymbol{\epsilon}_{n,t}^2) \right]'$ . In practice, it is common to set some small constant;  $c$ , to  $\mathbf{y}_t^*$  to avoid numerical problems when  $\mathbf{y}_t^*$  is close to zero – in this paper, we set  $c = 0.0001$ . Note that the resulting disturbance term in (A.10) is no longer Gaussian distributed but instead follows a  $\log \chi_1^2$  distribution. This means that despite being linear in the log-volatility term, standard linear Gaussian state space algorithms can not be directly applied. To overcome this difficulty, Kim et al. (1998) show that the moments of the  $\log \chi_1^2$  distribution can be well approximated through a seven Gaussian mixture, in which an auxiliary random variable, denoted  $s_t$ , serves as the mixture component indicator – hence the name of the algorithm. That is:

$$f(\boldsymbol{\epsilon}_t^*) \approx \sum_{j=1}^7 p_j f_N(\boldsymbol{\epsilon}_t^* | \mu_j - 1.2704, \sigma_j^2), \quad (\text{A.11})$$

where  $p_j = P(s_t = j)$ ,  $j = 1, \dots, 7$ ,  $f_N(\cdot | \mu, \sigma^2)$  is a Gaussian density with mean  $\mu$  and variance  $\sigma^2$  and the values of the probabilities and moments associated with each Gaussian distribution are given in Table A.3.

Table A.3: A seven component Gaussian mixture for approximation the  $\log - \chi_1^2$  distribution

Component	$p_j$	$\mu_j$	$\sigma_j^2$
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Source: Kim et al. (1998, p. 371)

In summary, given the vector of mixture component indicators;  $\mathbf{s} = \left[ s_1 \ \dots \ s_T \right]'$ , and parameter values in Table A.3, the state space model in (A.10) and (9) is linear and conditionally Gaussian, thus standard sampling methods can be used. Instead of adopting traditional Kalman filter based algorithms, we follow Poon (2018) and implement the efficient precision sampling based algorithm in Chan and Jeliazkov (2009); Chan and Hsiao (2014).

To summarize, define  $\mathbf{y}^* = (\mathbf{y}_1^*, \dots, \mathbf{y}_T^*)'$ ,  $\mathbf{h} = (h_1, \dots, h_T)'$  is a  $T \times 1$  and  $\boldsymbol{\epsilon}^* = (\boldsymbol{\epsilon}_1^*, \dots, \boldsymbol{\epsilon}_T^*)'$ .

Stacking the measurement equation in (A.10) over dates  $t = 1, \dots, T$  gives:

$$\mathbf{y}^* = \mathbf{X}_h \mathbf{h} + \boldsymbol{\epsilon}^*, \quad (\text{A.12})$$

where  $\mathbf{X}_h = \mathbf{I}_T \otimes \boldsymbol{\iota}_{NG}$  and  $\boldsymbol{\epsilon}^* \sim N(\mathbf{d}_s, \boldsymbol{\Sigma}_{\mathbf{y}^*})$  where  $\mathbf{d}_s = (\boldsymbol{\mu}_{s_1}, \dots, \boldsymbol{\mu}_{s_T})'$  and  $\boldsymbol{\Sigma}_{\mathbf{y}^*} = \text{diag}(\boldsymbol{\sigma}_{s_1}, \dots, \boldsymbol{\sigma}_{s_T})$  in which  $\boldsymbol{\mu}_{s_t} = (\mu_{s_t^1} - 1.2704, \dots, \mu_{s_t^n} - 1.2704)$  and  $\boldsymbol{\sigma}_{s_t} = (\sigma_{s_t^1}^2, \dots, \sigma_{s_t^n}^2)$ . Thus, by a change of variable:

$$(\mathbf{y}_i^* | \mathbf{s}_i, \mathbf{h}_i) \sim N(\mathbf{h}_i + \mathbf{d}_i, \boldsymbol{\Sigma}_{\mathbf{y}_i^*}). \quad (\text{A.13})$$

To complete the state space representation, stack the state equation for the log stochastic volatility factor in (9) over all dates  $t = 1, \dots, T$  to get:

$$\mathbf{h} = \boldsymbol{\alpha}_h + \mathbf{H}_h^{-1} \boldsymbol{\xi}, \quad \boldsymbol{\xi} \sim N(0, \boldsymbol{\Phi}), \quad (\text{A.14})$$

where  $\boldsymbol{\alpha}_h = \mathbf{H}_h^{-1} \begin{pmatrix} h_0 & 0 & \dots & 0 \end{pmatrix}'$ ,  $\boldsymbol{\Phi} = \text{diag}\left(\frac{\sigma_h^2}{(1-\rho^2)}, \sigma_h^2, \dots, \sigma_h^2\right)$  and:

$$\mathbf{H}_h = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & \vdots \\ 0 & -\rho & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\rho & 1 \end{bmatrix}.$$

By a change of variable:

$$(\mathbf{h} | \boldsymbol{\Phi}, h_0) \sim N\left(\boldsymbol{\alpha}_h, (\mathbf{H}_h' \boldsymbol{\Phi}^{-1} \mathbf{H}_h)^{-1}\right). \quad (\text{A.15})$$

Finally, combining (A.13) and (A.15) gives the conditional posterior distribution:

$$(\mathbf{h} | \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}_u, \sigma_h^2, \rho, \boldsymbol{\kappa}) \sim N(\hat{\mathbf{h}}, \mathbf{K}_h^{-1}), \quad (\text{A.16})$$

where  $\hat{\mathbf{h}} = \mathbf{K}_h^{-1}(\mathbf{H}_h' \boldsymbol{\Phi}^{-1} \mathbf{H}_h' \boldsymbol{\alpha}_h + \mathbf{X}_h' \boldsymbol{\Sigma}_{\mathbf{y}_i^*}^{-1}(\mathbf{y}^* - \mathbf{d}_s))$  and  $\mathbf{K}_h = \mathbf{H}_h' \boldsymbol{\Phi}^{-1} \mathbf{H}_h + \mathbf{X}_h' \boldsymbol{\Sigma}_{\mathbf{y}_i^*}^{-1} \mathbf{X}_h$ . As in step 1

sampling from this distribution is conducted with the precision sampling algorithm in [Chan and Jeliazkov \(2009\)](#).

**Step 5: Sample from**  $p(\rho \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \sigma_h^2, \boldsymbol{\kappa})$

To sample  $\rho$ , first note that combining the prior with (9) gives:

$$\left(\rho \mid \mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{h}, \sigma_h^2, \boldsymbol{\Sigma}_u, \boldsymbol{\kappa}\right) \propto p(\rho)g(\rho)\exp\left\{-\frac{1}{2\sigma_h^2}\sum_{t=2}^T(h_t - \rho h_{t-1})^2\right\}, \quad (\text{A.17})$$

where  $g(\rho) = (1 - \rho^2)^{\frac{1}{2}}\exp(-\frac{1}{2\sigma_h^2}(1 - \rho^2)(h_1 - h_0)^2)$  and  $p(\rho)$  is a truncated normal. Since this conditional distribution is non-standard, we follow [Chan and Hsiao \(2014\)](#) and implement an independence-chain Metropolis-Hastings step with a truncated normal distribution. More precisely,  $\rho \sim N(\hat{\rho}, D_\rho)1(|\rho| < 1)$ , where  $D_\rho = (V_\rho + \mathbf{X}'_\rho \mathbf{X}_\rho / \sigma_h^2)^{-1}$  and  $\hat{\rho} = D_\rho (V_\rho^{-1} \mu_\rho + \mathbf{X}'_\rho \mathbf{z}_\rho / \sigma_h^2)$ , in which  $\mathbf{X}_\rho = (h_1, \dots, h_{T-1})'$  and  $\mathbf{z}_\rho = (h_2, \dots, h_T)'$ . Then, given the current draw  $\rho^d$ , a proposal draw  $\rho^c$  is accepted with probability  $\min\left\{1, \frac{g(\rho^c)}{g(\rho^d)}\right\}$ , otherwise the Markov chain stays at the current draw.

**Step 6: Sample from**  $p(\sigma_h^2 \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \rho, \boldsymbol{\kappa})$

To sample  $\sigma_h^2$ , we combine the inverse-Gamma prior distribution with the Gaussian likelihood function to get

$$(\sigma_h^2 \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{h}, \sigma_h^2, \boldsymbol{\Sigma}_u, \rho, \boldsymbol{\kappa}) \sim IG\left(k_3^2 + \frac{T-1}{2}, +[(1-\rho)^2(h_1)^2 + \sum_{t=2}^T(h_t - \rho h_{t-1})^2 + k_3^2 V_3]/2\right). \quad (\text{A.18})$$

**Step 7: Sample from**  $p(\boldsymbol{\kappa} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_u, \mathbf{h}, \boldsymbol{\Omega}, \rho, \sigma_h^2)$

Finally, to sample the hyperparameters in  $\boldsymbol{\kappa}$  we follow [Amir-Ahmadi et al. \(2018\)](#) and employ a version of the (Gaussian) random walk Metropolis-Hastings algorithm for the proposal variance in a burn-in phase. In what follows, we describe the sampling procedure for a generic scaling factor  $\boldsymbol{\kappa}_X$  where  $X \in \{\boldsymbol{\Sigma}_u, \boldsymbol{\Omega}, \sigma_h^2\}$ .<sup>10</sup> Starting from an initial condition in which we set  $\boldsymbol{\kappa}_X^0$  to be the unit vector, the details are as follows:

1. At step  $i$ , take a candidate draw  $\boldsymbol{\kappa}_X^c$  from  $N(\boldsymbol{\kappa}_X^{i-1}, \sigma_{\boldsymbol{\kappa}_X}^2 \mathbf{I})$  where  $\sigma_{\boldsymbol{\kappa}_X}^2$  is a tuning parameter

---

<sup>10</sup>Note that  $\boldsymbol{\kappa}_X$  will be a scalar when associated with  $\boldsymbol{\Sigma}_u$  or  $\sigma_h^2$  and a vector when associated with  $\boldsymbol{\Omega}$ .

which is changed in the burn-in phase to achieve a target acceptance rate.<sup>11</sup>

2. Calculate the acceptance probability  $p^i = \left\{ 1, \frac{p\langle X|\kappa_X^c \rangle p(\kappa_X^c)}{p\langle X|\kappa_X^{i-1} \rangle p(\kappa_X^{i-1})} \right\}$
3. Accept the candidate draw by setting  $\kappa_X^i = \kappa_X^c$  with probability  $p^i$ . Otherwise set  $\kappa_X^i = \kappa_X^{i-1}$ .

### Appendix A.2. Model Comparison

In this appendix we discuss how to compute the marginal likelihood using the one-step-ahead predictive likelihood. To this end, let  $\mathbf{Y}_t^o$  denote a vector of observed variables up to date  $t$ . Following [Geweke and Amisano \(2011\)](#), the one step ahead predictive likelihood for model  $M_i$ , given data up to date  $t - 1$ , is given by:

$$p(\mathbf{Y}_t|\mathbf{Y}_{t-1}^o, M_i) = \int p(\mathbf{Y}_t, \boldsymbol{\Theta}_i|\mathbf{Y}_{t-1}^o, M_i) d\boldsymbol{\Theta}_i, \quad (\text{A.19})$$

$$= \int p(\mathbf{Y}_t|\mathbf{Y}_{t-1}^o, M_i, \boldsymbol{\Theta}_i) p(\boldsymbol{\Theta}_i|M_i) d\boldsymbol{\Theta}_i \quad (\text{A.20})$$

where  $t = 1$  is evaluated by:

$$p(\mathbf{Y}_1^o|M_i) = \int p(\mathbf{Y}_1^o) p(\boldsymbol{\Theta}_i|M_i) d\boldsymbol{\Theta}_i, \quad (\text{A.21})$$

which is entirely driven by the marginal data density:  $p(\mathbf{Y}_1^o)$ . Given this value, we then approximate  $p(\mathbf{Y}_t|\mathbf{Y}_{t-1}^o, M_i)$  for dates  $t = 2, \dots, T$  by the Monte Carlo average:

$$p(\mathbf{Y}_t|\widehat{\mathbf{Y}}_{t-1}^o, M_i) = R^{-1} \sum_{r=1}^R p(\mathbf{Y}_t|\mathbf{Y}_{t-1}^o, M_i, \boldsymbol{\Theta}_{i,t-1}^{(r)}), \quad (\text{A.22})$$

where  $\{\boldsymbol{\Theta}_{i,t-1}^{(r)}\}_{r=1}^R$  is a sequence of draws from the Metropolis-Hastings within Gibbs sampler described in [Appendix A.1](#).

Finally, to see how the predictive likelihood is related to the Bayes factor, note that the marginal

---

<sup>11</sup>We set the tuning parameter  $\sigma_{\kappa_X} = 0.001$  and we found that the acceptance ratio is between 80-90 percent for each of the models used in the paper.

likelihood of the model  $M_i$  is given by:

$$p(\mathbf{Y}_T^o | M_i) = \prod_{t=1}^T p(\mathbf{Y}_t | \widehat{\mathbf{Y}}_{t-1}^o, M_i). \quad (\text{A.23})$$

Thus, the Bayes factor between models  $M_i$  and  $M_j$  is:

$$BF_{i,j} = \prod_{t=1}^T \frac{p(\mathbf{Y}_t | \widehat{\mathbf{Y}}_{t-1}^o, M_i)}{p(\mathbf{Y}_t | \widehat{\mathbf{Y}}_{t-1}^o, M_j)}. \quad (\text{A.24})$$