Three Applications of Time-Varying Parameter and Stochastic Volatility Models to the Malaysian and Australian Economy

Aubrey Poon

A thesis submitted for the degree of Ph.D Economics
The Australian National University
March 2017
© Aubrey Poon 2017
Except where otherwise indicated, this thesis is my own original work.

Aubrey Poon 9th March 2017
I dedicate this thesis to my Mother Margaret Lee, Father Kee Shiang Poon, Sister Brenda Poon and my Grandfather Sylvester Lee.
Acknowledgment

First, I would like to thank my Chair supervisor Joshua Chan for all the help he has given me during my PhD candidacy. Through his direction and guidance I was able to learn and understand what it takes to become a professional researcher. Also, I owe a debt of gratitude to Joshua in regards to teaching me the fundamentals in regards to Bayesian econometrics and Matlab programming. Joshua’s Bayesian econometrics course early in my PhD studies was very important as it allowed me to build the fundamental skills required for research. I would also like to thank my Panel supervisor Timothy Kam who also gave me great direction and advice in regards to my PhD thesis and also the importance of becoming a high level researcher. I would also like to thank my Panel supervisor Chung Tran for all the feedback he has given me during my PhD seminar presentations.

Second, I would like to thank all my fellow PhD classmates, Jamie Cross, Jim Hancock, Chenghan Hou, Qingyin Ma, Luis Uzeda Garcia, Sanghyeok Lee, Minh Nguyen and Azadeh Abbasi-Shavazi for their support during my PhD seminar presentations. I would like to make a special mention to my coauthor Jamie Cross, who gave me good constructive feedback in regards to my Chapter 2 and 4 papers. I would like to acknowledge financial support by The Economic Society of Australia via the 2015 Peter Kenyon Memorial Postgraduate Scholarship. This scholarship significantly helped me during my PhD candidacy. I would also like to thank Mary Lu Nelson for help in editing my PhD papers.

I would also like to thank my good friends Daniel Pastorelli and Allan Poon for all the support they have given me through my PhD studies and life. Lastly, I would like to thank my Mother Margaret Lee, Father Kee Shiang Poon and Sister Brenda Poon for their incredible support they have given me throughout my PhD studies.
Abstract

After the introductory chapter, this thesis comprises of three chapters that examines the application of time-varying parameter and stochastic volatility models to the Malaysian and Australian economy.

Chapter 2 aims to determine whether the propagation and transmission mechanism of Malaysian monetary policy differed during the Asian Financial Crisis of 1997/98 and the Global Financial Crisis of 2007/08. The methodology employs a time-varying vector-autoregression framework. The primary result is that despite having no evidence of time-variation in the propagation mechanism of Malaysian monetary policy the average contribution of a monetary policy shock to the variability of each macroeconomic variable- Real GDP, Inflation and the Nominal Effective Exchange Rate-differs between the two crises. This finding suggests that despite the propagation mechanism being relatively constant, Malaysia’s monetary policy transmission mechanism evolves over time. We believe that the main mechanism driving this evolution is the time-variation in the variance-covariance matrix of the shocks of the model, not the coefficients. We also find some evidence that the implementation of capital controls reduced the influenceability of monetary policy on the Malaysian economy.

Chapter 3 investigates whether incorporating time variation and fat-tails into a suite of popular univariate and multivariate Gaussian distributed models can improve the forecast performance of key Australian macroeconomic variables: real GDP growth, CPI inflation and a short-term interest rate. The forecast period is from 1992Q1 to 2014Q4, thus replicating the central banks forecasting responsibilities since adopting inflation targeting. We show that time varying parameters and stochastic volatility with Student’s-t error distribution are important modeling features of the data. More specifically, a vector autoregression with the proposed features provides the best interest and inflation forecasts over the entire sample. Remarkably, the full sample results show that a simple rolling window autoregressive model with Student’s-t errors provides the most accurate GDP forecasts.

Chapter 4 estimates a time-varying parameter Panel Bayesian vector autoregression
with a new feature: a common stochastic volatility factor in the error structure, to assess the synchronicity and the nature of Australian State business cycles. The common stochastic volatility factor reveals that macroeconomic volatility or uncertainty was more pronounced during the Asian Financial Crisis as compared to the more recent Global Financial Crisis. Next, the Panel VAR’s common, regional and variable specific indicators capture several interesting economic facts. In the first instance, the fluctuations of the common indicator closely follow the trend line of the Organisation for Economic Co-operation and Development composite leading indicators for Australia making it a good proxy for nationwide business cycle fluctuations. Next, despite significant co-movements of Australian States and Territory business cycles during times of economic contractions, the regional indicators suggest that the average degree of synchronisation across the Australian States and Territories cycles in the 2000s is only half of that presented in the 1990s. Given that aggregate macroeconomic activity is determined by cumulative activity of each of the nation states, the results suggests that the Federal Government should award state governments greater autonomy in handling state specific cyclical fluctuations.
# Table of Contents

## Contents

1 Chapter 1  
  1.1 Introduction ........................................ 15

2 Chapter 2  
  2.1 Introduction ........................................ 20  
  2.2 Empirical Methodology ................................. 23  
    2.2.1 Data ........................................ 24  
    2.2.2 Capital Control Index ......................... 25  
    2.2.3 Priors ........................................ 26  
    2.2.4 Identification ................................. 27  
    2.2.5 Estimations .................................... 29  
  2.3 Empirical Results .................................... 30  
    2.3.1 Model Comparison ............................... 31  
    2.3.2 Time-varying Volatility ......................... 33  
    2.3.3 Time-varying Impulse Responses ............... 34  
    2.3.4 Real GDP ..................................... 35  
    2.3.5 Inflation ..................................... 35  
    2.3.6 Nominal Effective Exchange Rate ............... 36  
    2.3.7 Forecast Error Variance Decomposition ....... 37  
  2.4 Robustness ........................................... 38  
    2.4.1 Prior Sensitivity ................................ 38  
    2.4.2 Extension of the Sample Period ............... 39  
  2.5 Conclusion ........................................... 40  
  2.6 Figures .............................................. 42  
  2.7 Appendix ............................................. 54  
  2.8 Appendix A ............................................ 54  
    2.8.1 Step 1 Drawing $\beta$ .......................... 55  
    2.8.2 Step 2 Drawing $h$ ............................. 56  
    2.8.3 Step 3 Drawing $\alpha$ .......................... 60  
    2.8.4 Step 4, 5, 6 Drawing $\Omega_\beta$, $\Omega_\alpha$, $\Omega_h$  
      .................................................. 62  
  2.9 Appendix B ............................................ 63  
    2.9.1 Generalised Impulse Response Functions and 
      Sign Restrictions ................................. 63  
  2.10 Appendix C ........................................... 65  
    2.10.1 Markov Chain Monte Carlo Convergence ....... 65  
    2.10.2 DIC Estimation ................................ 66

3 Chapter 3 .............................................. 68  
  3.1 Introduction ........................................ 69  
  3.2 Models .............................................. 73  
    3.2.1 Regime Switching Models ....................... 74  
    3.2.2 Time-varying Models ......................... 76  
    3.2.3 Rolling-window Models ......................... 79  
    3.2.4 Stochastic Volatility under Student’s-t Distributed Errors  
      .................................................. 80  
  3.3 Data and Forecast Metrics ............................ 80
List of Figures

2.1 Capital Controls Index for Outflows. 1 refers to restriction and 0 refers to liberalization ................................. 42
2.2 Trace plots of selected parameters: (a) $\beta_{20}$, (b) $\beta_{555}$, (c) $\beta_{1600}$, (d) $a_{21}$, (e) $a_{120}$, (f) $h_{55}$ and (g) $h_{200}$ ........................................ 42
2.3 Posterior mean (blue line), 16th (red line) and 84th (brown line) percentiles of the estimated standard deviations of the stochastic volatility for each variable. .................................................. 43
2.4 The median generalised impulse responses of Real GDP growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals. .......................... 44
2.5 The median generalised impulse responses of Real GDP growth to a contractionary monetary policy shock. .................. 44
2.6 Differences between impulse responses for Real GDP growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval. .................. 45
2.7 The median generalised impulse responses of inflation growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals. .................. 46
2.8 The median generalised impulse responses of inflation growth to a contractionary monetary policy shock. .................. 46
2.9 Differences between impulse responses for inflation growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval. .................. 47
2.10 The median generalised impulse responses of NEER growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals.

2.11 The median generalised impulse responses of NEER growth to a contractionary monetary policy shock.

2.12 Differences between impulse responses for NEER growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval.

2.13 The median generalised impulse responses to a contractionary monetary policy shock for Prior 1 specification.

2.14 The median generalised impulse responses to a contractionary monetary policy shock for Prior 2 specification.

2.15 The median generalised impulse responses, for period of 1996Q1, 1997Q3, 1998Q4, 2006Q1, 2008Q4, and 2010Q1, for all variables to a contractionary monetary policy shock for a TVP-VAR-SV model with a sample period of 1990Q1-2015Q4 and no capital controls index.

2.16 The median generalised impulse responses, for period of 2012Q1, 2013Q4 and 2015Q4, for all variables to a contractionary monetary policy shock for a TVP-VAR-SV model with a sample period of 1990Q1-2015Q4 and no capital controls index.

2.17 20th order sample autocorrelation: (a) $\beta$, (b) $a$, (c) $h$ and (d) $\omega^2$.

3.1 Australian Macroeconomic Time Series.

3.2 Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for CPI Inflation.

3.3 Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for real GDP growth.

3.4 Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for the interest rate.
3.5 Aggregate posterior mean of the degree of freedom parameter for the interest rate, real GDP growth and the inflation rate.

3.6 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; CPI inflation.

3.7 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; CPI inflation.

3.8 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; RGDP Growth.

3.9 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; RGDP Growth.

3.10 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; Interest.

3.11 Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; Interest.

4.1 Trace plots of selected parameters: (a) $\theta_{10}$, (b) $\theta_{550}$, (c) $\theta_{1000}$, (d) $\omega_3^2$, (e) $\omega_7^2$, (f) $(\Sigma_u)_{20}$ and (g) $h_{60}$

4.2 Posterior median (blue line), 16th (red line) and 84th (red line) percentiles of the common factor stochastic volatility (expressed as standard deviation).

4.3 Posterior median (blue line), 16th (red line) and 84th (red line) percentiles of the stochastic volatility (expressed as standard deviation) from AR-SV(1) model.

4.4 Plot of common indicator over time.

4.5 Plot of the posterior median common indicator and the OECD composite leading indicator (CLI) over time.

4.6 Plot of variable-specific indices over time.

4.7 Plot of State Indices over time.

4.8 Plot of pairwise rolling correlations between the State factors.

4.9 Plot of pairwise rolling correlations between the State factors.
4.10 Plot of pairwise rolling correlations between the State factors . . . . . . 142
4.11 Plot of average correlations between a State indicator with all others . 143
4.12 Plot of Historical Decomposition . . . . . . . . . . . . . . . . . . . . . . . 144
4.13 Plot of Historical Decomposition . . . . . . . . . . . . . . . . . . . . . . . 145
List of Tables

2.1 Geweke Convergence Diagnostics Statistic ........................................... 30
2.2 DIC estimates for competing VARs (numerical standard errors in the parentheses) .......................................................... 32
2.3 Forecast Error Variance Decomposition .............................................. 38
2.4 A Seven Component Gaussian Mixture for Approximating the $\log - \chi^2_1$ distribution ................................................................. 60
2.5 Summary Distribution of Inefficiency Factors for different set of parameters 65
3.1 Full sample univariate point forecast for interest, $i$, GDP growth, $y$, and inflation, : relative mean square forecast errors (RMSFE) subject to an AR benchmark. .................................................... 91
3.2 Full sample multivariate point forecast for interest, $i$, GDP growth, $y$, and inflation, : relative mean square forecast errors (RMSFE) subject to a VAR benchmark. .................................................... 92
3.3 Best point forecast model for interest, $i$, GDP growth, $y$, and inflation, over the full sample: relative mean square forecast errors (RMSFE) with multivariate model as benchmark. ........................................... 93
3.4 Full sample univariate density forecast for interest, $i$, GDP growth, $y$, and inflation, : relative log predictive likelihood (RLPL) subject to an AR benchmark. .................................................... 94
3.5 Full sample multivariate density forecast for interest, $i$, GDP growth, $y$, and inflation, : relative log predictive likelihood (RLPL) subject to a VAR benchmark. .................................................... 95
3.6 Best density forecast model for interest, $i$, GDP growth, $y$, and inflation, over the full sample: relative log predictive likelihood (RLPL) with multivariate model as benchmark. ........................................... 96
4.1 Geweke Convergence Diagnostics Statistic ........................................... 122
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>Log marginal likelihood estimates for selected models</td>
</tr>
<tr>
<td>4.3</td>
<td>Percentage of variance explained by the common indicator</td>
</tr>
<tr>
<td>4.4</td>
<td>A Seven Component Gaussian Mixture for Approximating the ( \log -\chi^2_1 ) distribution</td>
</tr>
</tbody>
</table>


1 Chapter 1

1.1 Introduction

Since the seminal works of Primiceri (2005), and Cogley and Sargent (2005), the time-varying parameter vector autoregression (TVP-VAR) with stochastic volatility has become an increasingly popular tool in the macroeconomics literature. Koop and Korobilis (2010) note that maintaining the VAR parameters and error covariances constant over time is too restrictive since there have been many studies in the macroeconomics literature that have documented the occurrence of structural breaks and parameter change in many time series variables. Therefore, when analysing macroeconomic policy issues, they argue that the analysis should be based on multivariate models where both the VAR coefficients and the error covariance matrices evolve over time. For instance, Koop, Leon-Gonzalez and Strachan (2009) estimated a model similar to Primiceri (2005), but their model differ in that it allows them to determine how the nature of the parameters evolved over time. They found overwhelming evidence of gradual changes in all their parameters over time and reinforces the findings in Primiceri (2005). However, this time-varying structure in the parameters and error covariances is also applicable to other models, not only in the VAR model. For instance, Korobilis (2013) implements a similar time-varying structure within a dynamic factor model framework to assess the transmission mechanism of US monetary policy.

As the literature on TVP-VAR progressed, new features and different specifications, Chiu et al. (2015), Clark and Ravazzolo (2015) and Ciccarelli, Ortega and Valderrama (2016) all introduced fat-tails or student-t errors in the error structure of the model. The underlying motivation in incorporating this new feature was to enhance the model’s ability to capture large unanticipated macroeconomic shocks, such as the recent Global Financial Crisis and the oil price shocks of the 1970’s. Recently, studies undertaken by Curdia, Del Negro, and Greenwald (2014) and Chib and Ramamurthy (2014) have shown that models with a multivariate t-distributed shock structure have a better in-sample fit than models with standard Gaussian errors.

Another new development in the literature is the flexible extensions of the Panel
BVAR. A time-varying parameter version was developed by Canova, Ciccarelli and Ortega (2007) and Canova and Ciccarelli (2009) and is mainly used to assess cyclical or business cycle fluctuations across countries. Canova, Ciccarelli and Ortega (2007) highlight two reasons why this econometric model is advantageous when examining business cycles across countries. Firstly, the econometric methodology is designed for large scale dynamic models that display unit specific dynamics and cross country lagged inter-dependencies. Secondly, the parsimonious parameterisation proposed in Canova and Ciccarelli (2009) allows the researcher to endogenously produce an index structure where indicators of common and country specific cycles are recursively constructed and dynamically span cross country interdependencies. Recently, this Panel BVAR literature has been extended by Koop and Korobilis (2015a) and Koop and Korobilis (2015b) whereby they introduce Bayesian Model Averaging or Model Uncertainty to the model and apply it to inflation forecasting across countries.

This thesis examines three applications of different model specifications within the TVP-VAR framework. Specifically, in Chapter 2 we estimate a standard TVP-VAR with stochastic volatility model from Primiceri (2005) to examine the propagation and the transmission mechanism of Malaysian monetary policy. Chapter 3 considers whether incorporating time variation and fat-tails into a class of popular univariate and multivariate Gaussian distributed models can improve the forecast performance of key Australian macroeconomic variables. Chapter 4 estimates a time-varying parameter Panel BVAR with a new feature; a common stochastic volatility factor in the error structure, to assess the synchronicity and the nature of Australian State business cycles. Each of these chapters is self contained paper that includes an introduction and a conclusion. The contents of the individual chapters are outlined below.

The main aim of Chapter 2 is to determine whether the propagation and transmission mechanism of Malaysian monetary policy differed during the Asian Financial Crisis of 1997/98 and the Global Financial Crisis of 2007/08. We estimate a standard TVP-VAR with stochastic volatility model from Primiceri (2005). The primary result we find is that despite having no evidence of time-variation in the propagation mechanism
of Malaysian monetary policy the average contribution of a monetary policy shock to the variability of each macroeconomic variable: Real GDP, Inflation and the Nominal Effective Exchange Rate, differs between the two crises. This finding suggests that despite the propagation mechanism being relatively constant, Malaysia’s monetary policy transmission mechanism evolves over time. We believe that the main mechanism driving this evolution is the time-variation in the variance-covariance matrix of the shocks of the model, not the coefficients. We also find some evidence that the implementation of capital controls reduced the influenceability of monetary policy on the Malaysian economy.

Chapter 3 entails a study that investigates whether incorporating time variation and fat-tails into a class of popular univariate and multivariate Gaussian distributed models can improve the forecast performance of key Australian macroeconomic variables: Real GDP growth, CPI Inflation and a short-term interest rate. Our forecasting period is from 1992Q1 to 2014Q4, which is aligned to the central bank’s forecasting responsibilities since adopting inflation targeting. We found that time varying parameters and stochastic volatility with Student’s-t error distribution are important modeling features of the data. More specifically, a VAR with these proposed features provides the best interest rate and inflation forecasts over the entire sample. Remarkably, the full sample results show that a simple rolling window autoregressive model with Student’s-t errors provides the most accurate GDP forecasts.

In Chapter 4 we estimate a time-varying parameter Panel Bayesian vector autoregression with a new feature; a common stochastic volatility factor in the error structure, to assess the synchronicity and the nature of Australian State business cycles. The common stochastic volatility factor reveals that macroeconomic volatility or uncertainty was more pronounced during the Asian Financial Crisis as compared to the more recent Global Financial Crisis. Next, the Panel VAR’s common, regional and variable specific indicators capture several interesting economic facts. In the first instance, the fluctuations of the common indicator closely follow the trend line of the Organisation for Economic Co-operation and Development composite leading indicators for
Australia making it a good proxy for nationwide business cycle fluctuations. Next, despite significant co-movements of Australian States and Territory business cycles during times of economic contractions, the regional indicators suggest that the average degree of synchronisation across the Australian States and Territories cycles in the 2000s is only half of that presented in the 1990s. Given that aggregate macroeconomic activity is determined by cumulative activity of each of the nation states, the results suggests that federal governments should award state governments greater autonomy in handling state specific cyclical fluctuations.

Finally, Chapter 5 concludes and discusses future research topics.
Chapter 2

The Transmission Mechanism of Malaysian Monetary Policy: A Time-Varying Vector Autoregression Approach

This paper will be published in the Journal *Empirical Economics*
2.1 Introduction

Monetary policy has always played an important role in influencing the Malaysian economy throughout the years. Prior to the 1990s, the main focus of the Bank of Negara Malaysia (BNM), the central bank of Malaysia, was on monetary targeting. However, as rapid globalization of financial markets occurred in the early 1990s, the BNM objective shifted towards interest rate targeting. At the onset of the Asian Financial Crisis (AFC), a large capital flight and speculative pressures on the Malaysian ringgit prevented the BNM from influencing the interest rate for domestic purposes. As a result, this caused a severe contraction in the economy. In response, the Malaysian government implemented selective capital controls on outflows during September 1998. Athukorala and Jongwanich (2012) argued that the controls helped insulate the Malaysian domestic capital markets from the world capital markets, which allowed the BNM to regain policy autonomy and enabled them to pursue an expansionary monetary policy to reflate the economy. In addition, the Malaysian government implemented a series of banking and corporate sector reforms. Through time, these reforms created a sounder and more stable banking and financial system in Malaysia. Therefore, by the time the Global Financial Crisis (GFC) hit Malaysia, the BNM was well placed or in a better position, compared to during the AFC, to respond to the crisis. Athukorala (2010) noted Malaysia was the first country in the region to pursue an expansionary monetary policy in response to the crisis and by the end of 2009 the economy had recovered.

The main objective of this paper is to focus on the periods during the AFC and GFC, and determine whether the propagation and transmission mechanism of Malaysian monetary policy between the two crises are different. To account for the effects of capital controls implemented during September 1998, we used the capital controls index for outflow derived from Athukorala and Jongwanich (2012) and estimate it as an exogenous variable within the model. We estimate a four-variable TVP-VAR-SV model, which consists of Real GDP, Inflation, Nominal Effective Exchange Rate and the Interest Rate as the endogenous variables. Lastly, this paper contributes to the empirical literature on Malaysian monetary policy and extends previous work by allowing for a
time-varying structure on the VAR.

Numerous empirical studies have been undertaken regarding Malaysian monetary policy. The studies by Athanasopoulos, Raghavan, and Silvapulle (2012), Fung (2002) and Ito and Sato (2008) all employed the structural vector autoregression (VAR) methodology to assess Malaysia’s monetary policy transmission mechanism. The studies by Ibrahim (2005) and Domac (1999) also employed the same econometric methodology, but their studies differ in that they focused on the sectoral effects of Malaysian monetary policy. In contrast, the study undertaken by Tang (2006) focused on the relative strengths of different monetary policy transmission channels in Malaysia. In regards to our study, we also assess Malaysia’s monetary policy transmission mechanism. However, our study differs from the previous empirical studies in that we employ a different econometric methodology and identification scheme. In all the studies stated above, they all estimated a constant parameter standard VAR and employed a standard recursive identification scheme. However, in our study we estimate a time-varying VAR with stochastic volatility (TVP-VAR-SV) from Primiceri (2005) and we employ the sign restriction approach for our identification scheme. By employing this time-varying structure, we are able to determine the evolution of Malaysia’s monetary policy transmission mechanism over time. This is very important, since Koop, Leon-Gonzalez and Strachan (2009) found evidence to support that the transmission mechanism, which is a major goal in many macroeconomic papers, changes over time. The TVP-VAR-SV model has become an increasingly popular tool within the macroeconomic literature. For instance, D’Agostino et al. (2013) and Cross and Poon (2016) found that the TVP-VAR-SV specification delivers a more accurate forecast than other VAR models. In terms of structural analysis, Benati (2008) used a TVP-VAR-SV model to investigate the causes of the Great Moderation in the United Kingdom and Baumeister and Poon (2013) used it to explain the relationship between oil supply shocks and the US economy. The TVP-VAR-SV model allows us to capture the time-varying behaviour of the underlying structure in the multivariate data, which enable us to capture any structural breaks or regime shift within the time series variables.
The main results can be summarised as follows. First, the generalised impulse responses for all variables between the two periods of crises are not statistically different. This implies that there is no evidence of time-variation within the propagation mechanism of Malaysian monetary policy on all the variables. This being said, the second key result is that the average contribution of a monetary policy shock to the variability of each variable changes over time. This suggest that the Malaysia’s monetary policy transmission mechanism evolves over time. In the spirit of the TVP-VAR-SV model, the main mechanism driving this evolution is time-variation in the variance-covariance matrix of the shocks of the model, not the VAR coefficients. This result is consistent with the findings of Chan and Eisenstat (2016) and Primiceri (2005) who conduct similar analysis on the US economy. To further investigate this result we then undertake a model comparison exercise in which we compare the TVP-VAR-SV model against three alternative models: a standard fixed coefficients VAR, a time-varying parameter VAR with constant variance (TVP-VAR) and a fixed coefficient VAR with stochastic volatility (VAR-SV). Model comparison is based on the Bayesian deviance information criterion (DIC) measure for each of the four completing models. The results show that the constant or time-invariant parameters VAR with stochastic volatility (VAR-SV) provides the best in sample fit out of the four models. This result further supports the aforementioned argument that the main source of time-variation in the model is through the variance-covariance matrix of the shocks. Lastly, in addition to these time series results we find evidence that the implementation of capital controls, to some extent, reduces the influence of Malaysian monetary policy on the economy. This result contradicts the argument put forward by Athukorala and Jongwanich (2012) who suggest that the imposition of capital controls allowed the BNM to regain monetary policy autonomy, thus enabling them to pursue expansionary policies to reflate the Malaysian economy.

This paper is organized as follows. Section 2 presents the empirical methodology: the estimation procedure for the TVP-VAR-SV model, the identification strategy and the priors for the model. Section 3 describes and discusses the empirical results from
the TVP-VAR-SV model. Section 4 details the robustness checks of the model. Finally, section 5 concludes.

2.2 Empirical Methodology

Following Primiceri (2005), we estimate a time-varying parameter vector autoregression (TVP-VAR-SV) model which allows for time variation from three sources: (1) in the VAR coefficients, (2) in the variance of the errors and (3) in the covariance of the errors. The TVP-VAR-SV model with $n$ variables and $p$ lags is defined by:

$$y_t = v_t + B_{1,t}y_{t-1} + \ldots + B_{p,t}y_{t-p} + \Gamma_{1,t}z_t + \ldots + \Gamma_{q,t}z_{t-q} + \epsilon_t,$$

(2.1)

where $t = 1, \ldots, T$ is denoted as the time periods, $p$ and $q$ are the number of lags for the endogenous and exogenous variables respectively, $z_t$ is the vector of $r \times 1$ exogenous variables and $y_t$ is the vector of $n \times 1$ observed endogenous variables. Both $v_t$ and $B_{i,t}$ are $n \times 1$ and $n \times n$ time varying vector and matrices of the intercepts and coefficients respectively. $\Gamma_{i,t}$ is $n \times r$ time varying matrix of the coefficients for the exogenous variables. The $\epsilon_t$ is a $n \times 1$ vector of heteroscedastic unobservable shocks with a $n \times n$ variance-covariance matrix of $\Sigma_t$, that is $\epsilon_t \sim N(0, \Sigma_t)$. In the empirical estimation we follow Primiceri (2005) and Nakajima, Kasuya and Watanabe (2011) and impose a lag length of two.

We can rewrite equation (1) into a standard linear regression matrix form:

$$y_t = X_t\beta_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma_t),$$

(2.2)

where $X_t = I_n \otimes (1, y_{t-1}', \ldots, y_{t-p}', z_t', \ldots, z_{t-q}')$, $\beta_t = vec([v_t, B_{1,t}, \ldots, B_{p,t}, \Gamma_{1,t}, \ldots, \Gamma_{q,t}])'$, $X_t$ is $n \times b$ matrix and $\beta_t$ is $b \times 1$, where $b$ is the number of $\beta$ parameters. Following Primiceri (2005) the variance-covariance matrix can be decomposed as $\Sigma_t^{-1} = L_t'D_t^{-1}L_t$. For example, in our study $n = 4$, which implies
\[
D_t = \begin{bmatrix}
    e^{h_{1,t}} & 0 & 0 & 0 \\
    0 & e^{h_{2,t}} & 0 & 0 \\
    0 & 0 & e^{h_{3,t}} & 0 \\
    0 & 0 & 0 & e^{h_{4,t}}
\end{bmatrix}, \quad L_t = \begin{bmatrix}
    1 & 0 & 0 & 0 \\
    \alpha_{21,t} & 1 & 0 & 0 \\
    \alpha_{31,t} & \alpha_{32,t} & 1 & 0 \\
    \alpha_{41,t} & \alpha_{42,t} & \alpha_{43,t} & 1
\end{bmatrix}.
\]

Let \( h_t = (h_{1,t}, h_{2,t}, h_{3,t}, h_{4,t})' \) and \( a_t = (\alpha_{21,t}, \alpha_{31,t}, \alpha_{32,t}, \alpha_{41,t}, \alpha_{42,t}, \alpha_{43,t})' \) then the model's time-varying parameters evolve according to:

\[
\beta_t = \beta_{t-1} + \nu_t, \quad \nu_t \sim N(0, \Omega_{\beta}),
\]

(2.3)

\[
a_t = a_{t-1} + \zeta_t, \quad \zeta_t \sim N(0, \Omega_a),
\]

(2.4)

\[
h_t = h_{t-1} + \eta_t, \quad \eta_t \sim N(0, \Omega_h),
\]

(2.5)

where \( \Omega_a = \text{diag}(\omega_{a1}^2, \omega_{a2}^2, \omega_{a3}^2, ..., \omega_{a6}^2)' \), \( \Omega_h = \text{diag}(\omega_{h1}^2, \omega_{h2}^2, \omega_{h3}^2, \omega_{h4}^2)' \), \( \Omega_{\beta} = \text{diag}(\omega_{\beta1}^2, \omega_{\beta2}^2, ..., \omega_{\beta6}^2)' \) are all diagonal matrices. Both \( \beta_t \) and \( a_t \) are modeled as driftless random walks, while \( e^{h_{i,t}} \) is modeled as a geometric random walk. Primiceri (2005) stated that the drifting coefficients are meant to capture the possible non-linearities or the time variation in the lag structure of the model and the multivariate stochastic volatility is supposed to capture the possible heteroscedasticity of the shocks and non-linearities in the simultaneous relation among the variables of model. Further, Primiceri (2005) added that by allowing for time-variation in both the coefficients and the variance-covariance matrix will enable the data to determine whether the time-variation of the linear structure is derived from changes in the size of the shocks or from the changes in the propagation mechanism.

2.2.1 Data

The data frequency is quarterly and the sample period covers 1990Q1 to 2010Q4. The reason the sample period ends at the end of 2010 is due to unavailability of data on the capital controls index past this point of time. The variables of interest are Real GDP growth, CPI inflation growth, Nominal Effective Exchange Rate (NEER) growth
and the Interest rate. This choice of variables is in line with the findings of Alonso-Carrera and Kam (2015) who show that these variables are required to successfully capture the dynamics in a small open economy model. This result is not dissimilar to Franta, Horvath and Rusnak (2014) who also note that these four variables are typically regarded as the minimum set of factors to be considered for an analysis of a small open economy. To take into consideration the implementation of capital controls by the Malaysian authorities, we include a capital control index for outflows as an exogenous variable $z_t$ in the model. This capital control index, discussed below, is a continuous variable and between 0 and 1.

All the data were gathered from the International Financial Statistics database from the International Monetary Fund (IMF). In regards to the monetary policy indicator variable, we used the short-term money market rate. Malaysia did not officially have a policy rate until the mid-2000s, which means data on the policy or interbank rate were unavailable during the 1990s. However, Raghavan and Silvapulle (2008) noted that from the mid-1990s the BNM began to shift towards interest rate targeting. Also, Domac (1999) notes that the BNM directly influences the interbank rate through its intervention in the money market. Therefore, it is a reasonable assumption that the short-term money market rate is an indicator of the BNM’s stance on monetary policy. Lastly, besides the interest rate, real GDP, CPI and the NEER were all seasonally adjusted, logarithmically first differenced and multiplied by 100.

2.2.2 Capital Control Index

The capital controls index for outflow is taken from the study undertaken by Athukorala and Jongwanich (2012). Previously, capital controls indexes had been constructed from the IMF’s Annual Report on Exchange Arrangement and Exchange Restrictions (see Ito and Chinn (2016), Johnston and Tamirisa (1998)). The authors argue that annual information from this report does not capture the variations of capital restrictions that well and as a result they constructed quarterly capital controls indexes on information gathered from notifications, press releases, and speeches related to foreign exchange
and the capital account issued by the BNM. They constructed indexes for both capital inflows and outflows, and within each capital flows indexes they disaggregated the flows into four categories: foreign direct investment, equity securities, debt securities and other investment flows. However, for simplicity, in our study we used the total aggregate capital controls indexes, which is the average of all four categories of indexes.

Athukorala and Jongwanich (2012) state that the indexes are constructed by assigning $+1$ or $-1$ to each announced measure. For instance, policy changes that facilitate inflows and outflows are assigned $+1$ and those that restrict inflows and outflows are assigned $-1$ regardless of whether they relate to transactions by residents or non-residents. This number is then scaled by different weights based on direct and indirect impact criteria. The weights are set between 0 and 2, the higher the weight, the more severe the measure is. Athukorala and Jongwanich (2012) states that a weight of 2 could be assigned when the BNM imposed a tax or lifted a certain policy measure. For a weight of 1, an example could be the BNM requests and/or requires investors or financial institutions to undertake certain measures. Lastly, an example for a weight to be assigned between 0.25 to 0.5 could be when the BNM changes a regulation slightly, seeks the cooperation of investors or provides them a particular option.

Once the number and weight have been assigned to every measure, the weighted numbers are sequentially accumulated over time in order for the computation of the indexes. Athukorala and Jongwanich (2012) rescaled the indexes to lie between 0 and 1. Figure 2.1 shows the graph of the capital controls index for outflows that we use in our study. Economically, a value of 0 represents when capital outflows are liberalized (or not restricted), whilst a value of 1 occurs when capital outflows are restricted.

### 2.2.3 Priors

It is well known that TVP-VARs are not parsimonious models. They have a large number of coefficients and without prior information, Koop and Korobilis (2010) show that it can be very difficult to obtain precise estimates of the VAR coefficients. In his original paper, Primiceri (2005) aimed to mitigate this problem by using a training
sample, consisting of the first 10 years of the sample period, to calibrate the coefficients prior distributions. However, due to our short sample period, it is not possible for us to use a training sample to specify the prior distributions. As a result, we follow Koop and Korobilis (2010) and calibrate the priors’ distributions for the initial conditions of the time-varying parameters as follows:

\[ \beta_1 \sim N(0, 4I_{b \times b}), \]
\[ h_1 \sim N(0, 4I_{n \times n}), \]
\[ a_1 \sim N(0, 4I_{m \times m}), \]

(2.6)

where \( m \) is denoted the number of dimensions of the vector \( a_t \). For the priors of the time-varying parameters error covariances, we implemented conjugate priors. Koop and Korobilis (2010) argued that conjugate priors lead to analytical results for the posterior and predictive densities. Primiceri (2005) stated that a slightly tight prior is needed for the error covariance of \( \Omega_\beta \) in order to avoid the implausible behaviours of the time-varying coefficients. Therefore the priors for \( i-th \) diagonals of the error covariances are:

\[ \omega_{a_i}^2 \sim IG(2, 0.01) \quad \text{for } i = 1, \ldots, m, \]
\[ \omega_{h_i}^2 \sim IG(2, 0.01) \quad \text{for } i = 1, \ldots, n, \]
\[ \omega_{\beta_i}^2 \sim IG(10, 0.01) \quad \text{for } i = 1, \ldots, b. \]

(2.7)

The hyperparameters for the priors for the error covariances are taken from Nakajima, Kasuya and Watanabe (2011).

### 2.2.4 Identification

Under the recursive assumption, the identified monetary policy shock is assumed to affect real GDP, inflation and NEER with at least one period of lag. This assumption is very common within the empirical literature (for instance see Bernanke and Mihov
(1998) and Christiano, Eichenbaum and Evans, (1999)). However, as noted within the literature, an evidence of a price puzzle is commonly associated with this identification scheme (see for instance Hanson (2004)). To overcome this price puzzle problem associated with recursively identified models, we implemented the sign restrictions approach by Faust (1998), Canova and DeNicolo (2002), and Uhlig (2005), whereby the structural shocks are identified by restricting the signs of the impulse responses of selected model variables to structural shocks. When implementing this approach, normally each identified shock is associated with a unique sign pattern. For our study, we simply restrict the signs of the impact/contemporaneous matrix as in Uhlig (2005). For our identifying restrictions, we follow the restrictions commonly set within the empirical literature (for instance see Ellis, Mumtaz and Zabczyk (2014), Benati and Mumtaz (2007), Canova and Gambetti (2009) and Franta, Horvath and Rusnak (2014)). Commonly, a monetary policy shock is identified based on the assumption that a contractionary monetary policy shock will have a non-positive effect on both real GDP and inflation, and a non-negative effect on both the NEER and interest rate. Therefore, assuming the vector $u_t$ and the matrix $A_t$ are the structural shocks and the impact/contemporaneous matrix respectively, the restrictions we impose are:

$$\epsilon_t = A_t u_t,$$  \hspace{1cm} (2.8)

$$\begin{bmatrix}
\epsilon_{GDP,t} \\
\epsilon_{\pi,t} \\
\epsilon_{NEER,t} \\
\epsilon_{int,t}
\end{bmatrix} = \begin{bmatrix}
\times & \times & \times & - \\
\times & \times & \times & - \\
\times & \times & \times & + \\
\times & \times & \times & +
\end{bmatrix} \begin{bmatrix}
u_{GDP,t} \\
u_{\pi,t} \\
u_{NEER,t} \\
u_{int,t}
\end{bmatrix},$$

where $GDP, \pi, NEER, int$ denote as real GDP growth, inflation growth, NEER growth and the interest rate respectively. Both $+$ (positive) and $-$ (negative) denotes the postulated sign of impact response and $\times$ denote no restriction. Since our main objective of this study is to assess the evolution of the transmission mechanism of Malaysian monetary policy, we only identify a monetary policy shock. Further research on modelling
the Malaysian economy needs to be undertaken first to fully identify GDP, inflation and exchange rate shocks by sign restrictions. To implement the sign restriction approach within the time-varying framework, we follow the methodology proposed in Baumeister and Peersman (2013) in drawing candidate solutions of $A_t$ that satisfy the restrictions above and more details about the procedure is discussed in the online appendix. Our acceptance ratio for drawing candidate impact matrix is about 28 per cent for a monetary policy shock. In other words, on average about four draws are needed to draw one solution of the candidate impact matrix that satisfy the sign restrictions above.

2.2.5 Estimations

The TVP-VAR-SV model is estimated through a standard Markov Chain Monte Carlo (MCMC) method and the sampling algorithm we follow is from Chan and Jeliazkov (2009), and Chan and Hsiao (2014). We follow the procedure of Baumeister and Peersman (2013) to store 50,000 draws after the initial 50,000 draws are discarded. More details about the Gibbs Sampler can be found in the online appendix. An important issue when using a Gibbs Sampler is the convergence of the limiting distribution of the sample to the posterior distribution. In theory, the sampler converges as the number of draws reaches infinity. In applied work, however, an infinite number of draws is infeasible. To assess whether our sample has converged, we thus follow Geweke (1992) and compute a finite draw convergence diagnostic. The convergence diagnostic is calculated by taking the difference between the means $\bar{g}_a = \frac{1}{n_a} \sum_{i=1}^{n_a} \theta^{(i)}$, based on the first $n_a$ draws and $\bar{g}_b = \frac{1}{n_b} \sum_{i=1}^{n_b} \theta^{(i)}$, based on the last $n_b$ draws and dividing by the asymptotic standard errors of the difference $\sqrt{\frac{\sigma^2_{n_a}}{n_a} + \frac{\sigma^2_{n_b}}{n_b}}$.

Following Geweke (1992) $n_a$ and $n_b$ are set to be the first 10 percent and last 50 percent of the total draws respectively. Thus, in terms of our estimation, $n_a$ is the first 5,000 draws and $n_b$ is the last 25,000 draws after the burn-in period. If the sequence of the MCMC sampling is stationary, then by the central limit theorem, the distribution of this diagnostic converges to a standard normal. Table 2.1 shows the posterior means, standard deviations, the convergence diagnostics and the inefficiency factors for selected parameter estimates. Notice for all the parameter estimates, the convergence
diagnostics (denoted CD in Table 2.1) are all less than the 5 per cent significance level, which implies that the null hypothesis of the convergence to the posterior distribution is not rejected. Also, all the inefficiency factors (denoted IF in Table 2.1) are less than 20. Primiceri (2005) notes that inefficiency factors below or around twenty are regarded as satisfactory. We also report the trace plots of these selected parameters in Figure 2.2 and for each parameter the chain appears to be stable. Therefore, the results from the Geweke convergence diagnostics, the inefficiency factors and Figure 2.2 show that the parameters and state variables are efficiently drawn from the posterior distributions. We also calculate the 20th order sample autocorrelation and the inefficiency factors for all the parameters in the model in the online appendix.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean ($n_a$)</th>
<th>Stdev. ($n_a$)</th>
<th>Mean ($n_b$)</th>
<th>Stdev. ($n_b$)</th>
<th>CD</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{20}$</td>
<td>-0.32</td>
<td>0.02</td>
<td>-0.32</td>
<td>0.01</td>
<td>-1.08</td>
<td>1.58</td>
</tr>
<tr>
<td>$\beta_{155}$</td>
<td>0.73</td>
<td>0.00</td>
<td>0.74</td>
<td>0.00</td>
<td>1.00</td>
<td>16.62</td>
</tr>
<tr>
<td>$\beta_{1600}$</td>
<td>0.20</td>
<td>0.02</td>
<td>0.18</td>
<td>0.01</td>
<td>1.02</td>
<td>2.41</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.14</td>
<td>0.01</td>
<td>0.14</td>
<td>0.00</td>
<td>0.94</td>
<td>3.22</td>
</tr>
<tr>
<td>$a_{120}$</td>
<td>-0.05</td>
<td>0.00</td>
<td>-0.04</td>
<td>0.00</td>
<td>-1.34</td>
<td>14.70</td>
</tr>
<tr>
<td>$h_{55}$</td>
<td>2.51</td>
<td>0.01</td>
<td>2.53</td>
<td>0.00</td>
<td>0.99</td>
<td>11.93</td>
</tr>
<tr>
<td>$h_{200}$</td>
<td>0.10</td>
<td>0.00</td>
<td>0.09</td>
<td>0.00</td>
<td>1.11</td>
<td>10.83</td>
</tr>
</tbody>
</table>

### 2.3 Empirical Results

In this section we present the empirical results from the TVP-VAR-SV model. In the first sub-section, we undertake a model comparison exercise via the Bayesian deviance information criterion. In the second sub-section, we examine the time-varying volatility of each of the endogenous variables in our model. For the third sub-section, we assess the evolution of Malaysia’s monetary policy transmission mechanism by deriving and examining the generalised impulse response functions from the TVP-VAR-SV model for a shock to monetary policy. We compute the generalised impulse response functions based on the procedure detailed in Koop, Pesaran and Potter (1996), for detail explanation see the online appendix. Our main focus is on the periods during the AFC and GFC. Lastly, we examine the forecast error variance decomposition for the contribution of a monetary policy shock to each variable.
2.3.1 Model Comparison

In this section, we compare the TVP-VAR-SV model to three other models, which are the standard fixed coefficients VAR, time-varying parameters VAR with constant variance (TVP-VAR) and standard fixed coefficient VAR with stochastic volatility (VAR-SV). Our model comparison is carried out via the Bayesian deviance information criterion (DIC) introduced by Spiegelhalter et al. (2002). The DIC can be viewed as a tradeoff between model fit and model complexity. Let $\psi$ denote the model-specific parameter vector. Then the DIC is defined as:

$$DIC = \overline{D(\psi)} + p_D,$$

(2.9)

where:

$$\overline{D(\psi)} = -2\mathbb{E}_\psi[\log f(y|\psi)|y] + 2\log h(y),$$

(2.10)

is the posterior mean deviance and $h(y)$ is some fully specified standardizing term that is function of the data alone. The model complexity is measured by the effective number of parameters $p_D$ of the model, which is defined as:

$$p_D = \overline{D(\psi)} - D(\tilde{\psi}),$$

(2.11)

and:

$$D(\psi) = -2\log f(y|\psi) + 2\log h(y),$$

(2.12)

where $\tilde{\psi}$ is an estimate of $\psi$, which is typically taken as the posterior mean or the mode. Therefore, the DIC can be interpreted as the sum of the posterior mean deviance, which measures the goodness of fit, and the effective number of parameters $p_D$. For model comparison, normally $h(y)$ is set to be unity for all models. The model with the lowest DIC is the preferred model. In most cases, the DIC can be computed by evaluating the likelihood function for each iteration of the MCMC. However, models
with stochastic volatility are difficult to compute since they do not have a closed-form expression. Commonly, one could use the auxiliary particle filter of Pitt and Shephard (1999) to evaluate the likelihood, for instance see Mumtaz and Sunder-Plassmann (2013). However, the major disadvantage of the auxiliary particle filter is that it is very computationally intensive. Recently, Chan and Eisenstat (2016) have developed a more efficient approach to calculating the DIC of the TVP-VAR-SV model, whereby they use an efficient important sampling estimator for evaluating the integrated likelihood. We follow their approach and more details about their methodology can be found in the online appendix and their paper.

Table 2.2: DIC estimates for competing VARs (numerical standard errors in the parentheses)

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>$p_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVP-VAR-SV</td>
<td>997.61(1.01)</td>
<td>49.89(0.47)</td>
</tr>
<tr>
<td>TVP-VAR</td>
<td>1042.74(0.19)</td>
<td>43.31(0.09)</td>
</tr>
<tr>
<td>VAR</td>
<td>1001.63(0.04)</td>
<td>46.09(0.01)</td>
</tr>
<tr>
<td>VAR-SV</td>
<td>860.42(0.77)</td>
<td>58.65(0.46)</td>
</tr>
</tbody>
</table>

Table 2.2 shows the estimated DIC for four models. Each DIC estimate (and corresponding numerical standard error) is computed using 10 parallel chains, each consists of 50,000 posterior draws after a 50,000 burn-in period. For the stochastic volatility case, the integrated likelihood is calculated at every 500-th post burn-in draw, that is, a total of 1,000 evaluations are made. The results show that the TVP-VAR-SV is only slightly preferred in comparison to the fixed coefficients VAR. However, it is clearly evident that the VAR-SV is the preferred model out of the four models\(^1\). Similarly, Chan and Eisenstat (2016) also found that the VAR-SV is the preferred model under both US and Australian data. Chan and Eisenstat (2016) conclude that most of the gains in the model fit appear to have come from allowing for stochastic volatility rather than time-variation in the VAR coefficients or contemporaneous relationship. This result is also consistent with the findings by Primiceri (2005). Table 2.2 also reports the effective number of parameters $p_D$ and as expected both TVP-VAR-SV and VAR-SV are the most complex model. However, the difference between the models are still quite small. In summary, these results imply that model’s with stochastic volatility

\(^1\)We also computed the DIC measure for a VAR-SV with a Minnesota (or non-informative) prior for the fixed coefficients. We found that the DIC measure (DIC = 864.42 & $p_D = 61.07$) to be very similar to the VAR-SV reported in Table 2.2.
are overwhelmingly favored by the data in comparison to models without stochastic volatility.

2.3.2 Time-varying Volatility

Figure 2.3 plots the standard deviation of the estimated stochastic volatility for each of the four variables. We plot the posterior mean with the 16th and 84th percentiles. Both the time-varying volatility for real GDP and the interest rate appear to be relatively constant throughout the sample period. This implies that the fluctuation in shocks in both real GDP and the interest rate are time-invariant. The constant volatility result for real GDP is largely surprising given that during the AFC the Malaysian economy experienced a severe deterioration and we would expect the standard deviation of real GDP volatility to jump during this period. Similarly, we also would expect the standard deviation of the interest rate volatility to be very high during the pre-AFC period due to the large inflow of short-term capital, as mentioned above. However, the standard deviation of the interest rate volatility has remained relatively constant throughout the sample period. Although, the volatility does exhibit a slight declining trend after the period of 1998 which would reflect the BNM policy response to the AFC. In response to the crisis, Malaysia imposed selective capital controls and Athukorala and Jongwanich (2012) argued that the controls helped insulate the Malaysian domestic capital markets from the world capital markets, which allowed the BNM to regain policy autonomy and enabled them to pursue an expansionary monetary policy. Lastly, the declining trend of the interest rate volatility throughout the 2000’s period could be due to the BNM adoption of inflation targeting during the period.

The time-varying volatility for inflation exhibited two humped shapes during the periods of 1993-94 and the GFC. The Malaysian economy was experiencing high level of growth during the pre-AFC period and as expected, inflation would be high during this period too. The standard deviation of the inflation volatility peaked at around 1994. However, since 1995, inflation volatility started to fall and it remained at low levels during the AFC. The reason Malaysia experienced low levels of inflation during
the AFC is due to the BNM managing the exchange rate. At the onset of the AFC, the Malaysian ringgit came under enormous speculative pressures and initially the BNM tried to defend it. By defending the exchange rate, the BNM would have contracted their money supply significantly and this would have lead to deflation in the economy. Since 2005, inflation volatility has started to rise significantly and it peaked during the 2007-08 period. The reason for this high inflation period could be due to the low level of interest rate at this time too, which is mentioned above. It appears that the GFC caused a large sharp deflation on the economy and inflation has remained low ever since.

As expected, the standard deviation of the NEER volatility was very high during the pre-AFC period and it peaked during late 1997. This is reflective of the large inflow of short-term capital during this period. Once the AFC hit the economy, a large capital flight occurred and as a result this caused a severe depreciation in the Malaysian ringgit. The large jump in NEER volatility in Figure 2.3 is reflective of this episode. In addition to the introduction of capital controls, the Malaysian authorities also pegged the Malaysian ringgit to the US dollar in response to the AFC which explains the declining trend in the volatility of the NEER after 1998.

2.3.3 Time-varying Impulse Responses

Since the coefficients are time-varying, there will be a different set of generalised impulse response functions at each date in the sample period. However, for our study, we only focus on the generalised impulse response functions of the periods that are associated with the AFC and GFC. For the AFC, we compute three generalised impulse response functions for the periods of 1996Q1, 1997Q3 and 1998Q4. 1996Q1 represents the pre-AFC period, 1997Q3 represents the period of the AFC and 1998Q4 represents the period after Malaysia imposed capital controls (September 1998). In regards to the GFC, we compute the generalised impulse response functions for the periods of 2006Q1, 2008Q4 and 2010Q1. These periods represent the periods before, during and after the GFC respectively. We focus on a contractionary monetary policy shock and normalized the generalised impulse response functions on the relative interest rate response for the
initial period at each point of time. This normalization allows us to isolate the changes in the transmission mechanism from the changes in the magnitude of the shock over time. Therefore, the magnitude of a monetary policy shock is such that it raises relative interest rate by 1 per cent in the initial period at each point in time.

2.3.4 Real GDP

Figures 2.4 and 2.5 report the generalised impulse response functions of real GDP growth to a 1 per cent increase in the interest rate. For all the periods, a contractionary monetary policy shock has a negative effect on real GDP growth on impact, which is consistent under conventional monetary theory. However, the negative effects of the shock appears to be short-term and not statistically significant. After the 2nd quarter of the initial shock, the impulse responses of real GDP growth oscillates between positive and negative territories and then converges back to zero for all time periods. The magnitude of the oscillation between positive and negative territories appears to be similar in size, which implies that Malaysian monetary policy has on average only a short-term effect on real GDP growth. Figure 2.5 shows there is evidence that the impulse responses are different between the periods. However, Figure 2.6, which reports the differences of the impulse responses between the periods of 1996Q1-1997Q3, 1997Q3-1998Q4, 2006Q1-2008Q4, 2006Q1-2010Q1, 2008Q4-1997Q3 and 2008Q4-1998Q4, shows that for each of the panels, the 68 per cent credible interval includes zero, which means there is no statistically evidence of time-variation within these periods. This result concludes there is no statistical evidence that the propagation mechanism of Malaysian monetary policy on real GDP growth was different during and between the AFC and GFC periods.

2.3.5 Inflation

Figures 2.7 and 2.8 report the generalised impulse response functions of inflation growth to a 1 per cent increase in the interest rate. For all the periods, a contractionary
monetary policy shock has also a negative effect on inflation growth on impact and except for 1996Q1, all the periods initial impact is statistically significant. However, the effects of the monetary policy shock also appear to be of short-term nature. Figure 2.8 shows that monetary policy during the 2008Q4 had the largest initial impact on inflation growth, compared to the other time periods. Figure 2.9 reports the differences of the impulse responses for the corresponding periods similar to Figure 2.6 and for all the panels, the 68 per cent credible interval includes zeros, which means there is no statistically evidence that time-variation is present within the propagation mechanism of Malaysian monetary policy on inflation growth during each of the crises. However, in Figure 2.8, the generalised impulse response functions are clearly different between the two crises and it appears that the shock had a overall larger negative impact on inflation growth during the GFC than the AFC. Economically, this result is consistent with the events of the AFC. As mentioned before, during the AFC, the BNM tried to defend their exchange rate from depreciating which resulted in a large contraction of the money supply. This caused a large deflation in the economy and it would have been very difficult for the BNM to reverse this deflation at that time. Also, Figure 2.3 shows that inflation volatility was very high in the lead up to the GFC compared to the AFC, which could mean that inflationary pressures or expectations were higher during the GFC than the AFC.

2.3.6 Nominal Effective Exchange Rate

Figures 2.9 and 2.10 report the generalised impulse response functions of NEER growth to a 1 per cent increase in the interest rate. For all the periods, a contractionary monetary policy shock has a positive effect on NEER growth on impact and except for the period of 2008Q4, they are all statistically significant. Similar to both real GDP and inflation growth, the effects of the shock also appears to be of short-term nature and it converges back to zero after the 2nd quarter of the initial shock. Figure 2.11 reports the differences of the impulse responses between the two period of crises and they also show that there is no statistical evidence of time-variation present within the
propagation mechanism of Malaysian monetary policy on NEER growth. This result is consistent with Figure 2.10, as all the period's impulse responses exhibit a similar shape. Although majority of the results are statistically insignificant, Figure 2.10 does provide some interesting economic inference or insight. It shows that a contractionary monetary policy shock has a larger initial impact on the period of 1997Q3 than 1998Q4. This result is intuitive since during the period of 1998Q4, the Malaysian authorities, in addition to the implementation of capital controls, pegged the Malaysian ringgit against the US dollar. One would expect that country's monetary policy would have less of an influence on the exchange rate when it is pegged to another country's currency.

### 2.3.7 Forecast Error Variance Decomposition

Table 2.3 reports the forecast error variance decomposition for the contribution of a monetary policy shock to each variable. Notice that on average the contribution of a monetary policy shock to the variability of each variable is different at each period of time. This result implies that there is evidence that the transmission mechanism of Malaysian monetary policy evolves through time. Focusing on the AFC periods, Table 2.3 shows that for the period after the imposition of capital controls (1998Q4), on average the contribution of a monetary policy shock to the variability or fluctuations in all three endogenous variables have decreased compared to the period of 1997Q3. This quantitative result implies that the implementation of capital controls to some extent reduced the effectiveness of Malaysian monetary policy in influencing the economy, which contradicts the argument put forward by Athukorala and Jongwanich (2012) that the imposition of capital controls allowed the BNM to regain monetary policy autonomy and enable them to pursue expansionary policies to reflate the Malaysian economy. Also, Table 2.3 shows that monetary policy on average was slightly more effective on real GDP growth during the GFC than the AFC, which implies that monetary policy played a significant role in the recovery of the Malaysian economy after the GFC. This is consistent within the literature. Athukorala (2010) noted that Malaysia was the first country in the region to pursue an expansionary monetary policy in response to the crisis and by the end of 2009 the economy had recovered.
Table 2.3: Forecast Error Variance Decomposition

<table>
<thead>
<tr>
<th>Dates</th>
<th>1996Q1</th>
<th>1997Q3</th>
<th>1998Q4</th>
<th>2006Q1</th>
<th>2008Q4</th>
<th>2010Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of quarters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.91%</td>
<td>10.01%</td>
<td>9.78%</td>
<td>11.97%</td>
<td>17.44%</td>
<td>15.64%</td>
</tr>
<tr>
<td>4</td>
<td>17.48%</td>
<td>9.45%</td>
<td>4.39%</td>
<td>14.58%</td>
<td>10.29%</td>
<td>16.15%</td>
</tr>
<tr>
<td>8</td>
<td>16.38%</td>
<td>8.39%</td>
<td>3.64%</td>
<td>15.13%</td>
<td>12.31%</td>
<td>17.63%</td>
</tr>
<tr>
<td>12</td>
<td>16.15%</td>
<td>8.10%</td>
<td>3.54%</td>
<td>15.03%</td>
<td>12.34%</td>
<td>17.61%</td>
</tr>
<tr>
<td>20</td>
<td>15.83%</td>
<td>7.79%</td>
<td>3.46%</td>
<td>14.88%</td>
<td>12.32%</td>
<td>17.47%</td>
</tr>
<tr>
<td>Average</td>
<td>15.35%</td>
<td>8.75%</td>
<td>4.96%</td>
<td>14.32%</td>
<td>12.94%</td>
<td>16.90%</td>
</tr>
<tr>
<td>Inflation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12.33%</td>
<td>11.26%</td>
<td>10.62%</td>
<td>10.01%</td>
<td>10.61%</td>
<td>10.01%</td>
</tr>
<tr>
<td>4</td>
<td>17.48%</td>
<td>9.45%</td>
<td>4.39%</td>
<td>14.58%</td>
<td>10.29%</td>
<td>16.15%</td>
</tr>
<tr>
<td>8</td>
<td>15.82%</td>
<td>9.25%</td>
<td>4.96%</td>
<td>15.20%</td>
<td>12.02%</td>
<td>17.63%</td>
</tr>
<tr>
<td>12</td>
<td>15.67%</td>
<td>9.01%</td>
<td>4.82%</td>
<td>15.15%</td>
<td>12.11%</td>
<td>17.13%</td>
</tr>
<tr>
<td>20</td>
<td>15.44%</td>
<td>8.64%</td>
<td>4.65%</td>
<td>15.05%</td>
<td>12.20%</td>
<td>17.14%</td>
</tr>
<tr>
<td>Average</td>
<td>15.04%</td>
<td>9.56%</td>
<td>6.09%</td>
<td>14.14%</td>
<td>11.59%</td>
<td>15.62%</td>
</tr>
<tr>
<td>NEER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.54%</td>
<td>6.25%</td>
<td>2.12%</td>
<td>10.88%</td>
<td>14.10%</td>
<td>18.73%</td>
</tr>
<tr>
<td>4</td>
<td>10.25%</td>
<td>6.01%</td>
<td>2.03%</td>
<td>10.46%</td>
<td>14.23%</td>
<td>17.75%</td>
</tr>
<tr>
<td>8</td>
<td>10.29%</td>
<td>6.00%</td>
<td>2.03%</td>
<td>10.62%</td>
<td>14.61%</td>
<td>17.83%</td>
</tr>
<tr>
<td>12</td>
<td>10.37%</td>
<td>5.99%</td>
<td>2.03%</td>
<td>10.68%</td>
<td>14.65%</td>
<td>17.83%</td>
</tr>
<tr>
<td>20</td>
<td>10.52%</td>
<td>5.94%</td>
<td>2.02%</td>
<td>10.73%</td>
<td>14.54%</td>
<td>17.83%</td>
</tr>
<tr>
<td>Average</td>
<td>10.39%</td>
<td>6.04%</td>
<td>2.04%</td>
<td>10.67%</td>
<td>14.43%</td>
<td>18.00%</td>
</tr>
<tr>
<td>Interest Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.76%</td>
<td>6.05%</td>
<td>1.91%</td>
<td>9.28%</td>
<td>9.26%</td>
<td>12.79%</td>
</tr>
<tr>
<td>4</td>
<td>13.18%</td>
<td>6.53%</td>
<td>1.97%</td>
<td>12.08%</td>
<td>11.33%</td>
<td>16.09%</td>
</tr>
<tr>
<td>8</td>
<td>13.38%</td>
<td>6.29%</td>
<td>1.92%</td>
<td>12.61%</td>
<td>12.08%</td>
<td>16.40%</td>
</tr>
<tr>
<td>12</td>
<td>13.38%</td>
<td>6.17%</td>
<td>1.91%</td>
<td>12.61%</td>
<td>12.19%</td>
<td>16.41%</td>
</tr>
<tr>
<td>20</td>
<td>13.39%</td>
<td>6.02%</td>
<td>1.90%</td>
<td>12.63%</td>
<td>12.19%</td>
<td>16.42%</td>
</tr>
<tr>
<td>Average</td>
<td>12.82%</td>
<td>6.21%</td>
<td>1.92%</td>
<td>11.84%</td>
<td>11.41%</td>
<td>15.62%</td>
</tr>
</tbody>
</table>

2.4 Robustness

To determine whether the results from the generalised impulse response functions presented above are robust, we consider two different prior specifications and extend the sample period. First, we consider a TVP-VAR-SV system under two different priors specifications. Second, we extend the sample period through to the end of 2015Q4.

2.4.1 Prior Sensitivity

In this section we consider two different priors specifications

Prior 1: \( \beta_1 \sim N(0,10I_{b \times b}), \ h_1 \sim N(0,10I_{n \times n}) \) and \( a_1 \sim N(0,10I_{m \times m}). \)

Prior 2: \( \omega^2_{a_1} \sim IG(5,0.01), \ \omega^2_{\beta_1} \sim IG(20,0.01) \) and \( \omega^2_{h_1} \sim IG(2,0.01). \)
These prior specifications are from Nakajima, Kasuya and Watanabe (2011). Both Figure 2.13 and 2.14 report the generalised impulse responses functions for real GDP, inflation and NEER growths to a 1 per cent increase in the interest rate under these two prior specifications respectively. It is clear that the generalised impulse response functions for these two prior specifications do not differ very much from the responses from our baseline TVP-VAR-SV model. However, for prior specification 2, there appears to be less oscillations in the generalised impulse response functions for real GDP growth and this could be due to the increase in the tightness of the hyperparameter of the variances of the time-varying $\beta$ parameters. Except for this case, majority of the impulse responses only differ in the magnitude and the same conclusion discussed in section 3.3 can be made from these two prior specifications too.

2.4.2 Extension of the Sample Period

In this section, we extend the model's sample period through to the end of 2015. However, due to data unavailability we were unable to include the capital controls index as an exogenous variable in the model. Figure 2.15 plots the corresponding generalised impulse response functions for all three endogenous variables to a 1 per cent increase in the interest rate from this extension. The impulse responses generated appear to be very similar to the baseline model. This shows that the exclusion of the capital controls index as an exogenous variable and the extension of the sample period do not significantly alter or impact on the above results. Therefore, this extension and two prior specifications above show that the results discussed in section 3.3 are robust. In Figure 2.16 we also plot all three endogenous variables generalised impulse response functions to a 1 per cent increase in interest rate for the period of 2012Q1, 2013Q4 and 2015Q4 for this extension. For real GDP growth, the impulse responses for all the periods exhibit similar oscillating behaviour as the baseline model. But it differs in that the impulse responses are more negative. In regards to inflation growth, the impulse responses for the period of 2013Q4 and 2015Q4 also exhibit oscillating behaviour whereas in the baseline model the impulse responses display a hump shaped
pattern. Although the patterns are different, the evolution of the effect of a monetary policy shock appears to be similar in both models. Lastly, for the NEER growth, the impulse responses for all the periods display a very similar pattern to the baseline model's impulse responses. In summary, from the plots of Figure 2.16 we can conclude that a monetary policy shock for the period of 2012Q1, 2013Q4 and 2015Q4, had a more persistent negative effect on real GDP growth and displayed similar results and features to the baseline model for both inflation and NEER growths.

2.5 Conclusion

The aim of this paper is to determine whether the propagation and transmission mechanism of Malaysian monetary policy differed during the Asian Financial Crisis of 1997/98 and the Global Financial Crisis of 2007/08. The methodology employs a time-varying vector autoregression framework. The primary result is that despite having no evidence of time-variation within the propagation mechanism of Malaysian monetary policy the average contribution of a monetary policy shock to the variability of each macroeconomic variable: Real GDP, Inflation and the Nominal Effective Exchange Rate, differs between the two crises. This finding suggests that despite the propagation mechanism being relatively constant, Malaysia's monetary policy transmission mechanism evolves over time. The finding that the main mechanism driving the evolution of the transmission mechanism is the error variance-covariances matrix of the model, not the VAR coefficients, is consistent with Chan and Eisenstat (2016) and Primiceri (2005) who examine the US economy. To elicit this insight we then conducted a formal model comparison using the Bayesian DIC measure for four competing models: the TVP-VAR-SV, a VAR-SV, a TVP-VAR and a VAR. The results showed that the constant parameter VAR with stochastic volatility (VAR-SV) is the preferred model or the best in sample fit out of the four models. This result further supports our argument above that the main source of time-variation in our model is through the variance-covariance matrix of the shocks. From a practical standpoint, these results suggest that if Malaysian policymakers want to analyse the effects of monetary policy or forecast a particular macroeconomic
variable that incorporates monetary policy, they should estimate a model that incorporates time-variation within the error variance-covariances matrix. By estimating a time-invariant model, the policymaker will not accurately capture the true dynamics of the data and it will result in bias estimates.

In addition to these results, we also find some evidence that the implementation of capital controls reduced the influenceability of monetary policy on the Malaysian economy. This result contradicts the argument put forward by Athukorala and Jongwanich (2012) that the imposition of capital controls allowed the BNM to regain monetary policy autonomy and enable them to pursue expansionary policies to reflate the Malaysian economy. Instead, the results presented here support the view that Malaysian capital controls were largely ineffective. Proponents against the capital controls argue that at the time of the implementation of controls, a large amount of capital had already left the country and capital outflows within the East Asian region had already began to subside. Also, Malaysia recovered about the same time as the other IMF supported crisis hit East Asian nations too.

A question that is left unanswered in this study is Malaysia’s monetary policy rule in regards to unexpected shocks to real GDP, inflation and the NEER. To investigate this issue further, one must fully identify the impact/contemporaneous matrix. One potential avenue for this research agenda is to follow Ellis, Mumtaz and Zabczyk (2014) and utilise a Dynamic Stochastic General Equilibrium (DSGE) model, simulate the impulse responses, and use these responses as motivating restrictions for the impact/contemporaneous matrix. In order for this agenda to begin, further research first needs to be undertaken in regards to the deep parameters of the Malaysian economy.
2.6 Figures

Figure 2.1: Capital Controls Index for Outflows. 1 refers to restriction and 0 refers to liberalization

Figure 2.2: Trace plots of selected parameters: (a) $\beta_{20}$, (b) $\beta_{555}$, (c) $\beta_{1600}$, (d) $a_{21}$, (e) $a_{120}$, (f) $h_{55}$ and (g) $h_{200}$
Figure 2.3: Posterior mean (blue line), 16th (red line) and 84th (brown line) percentiles of the estimated standard deviations of the stochastic volatility for each variable.
Figure 2.4: The median generalised impulse responses of Real GDP growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals.

Figure 2.5: The median generalised impulse responses of Real GDP growth to a contractionary monetary policy shock.
Figure 2.6: Differences between impulse responses for Real GDP growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval.
Figure 2.7: The median generalised impulse responses of inflation growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals.

Figure 2.8: The median generalised impulse responses of inflation growth to a contractionary monetary policy shock.
Figure 2.9: Differences between impulse responses for inflation growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval.
Figure 2.10: The median generalised impulse responses of NEER growth to a contractionary monetary policy shock and the shaded areas indicate the 68% posterior credible intervals.

Figure 2.11: The median generalised impulse responses of NEER growth to a contractionary monetary policy shock.
Figure 2.12: Differences between impulse responses for NEER growth: (a) 1996Q1-1997Q3, (b) 1997Q3-1998Q4, (c) 2006Q1-2008Q4, (d) 2006Q1-2010Q1, (e) 2008Q4-1997Q3 and (f) 2008Q4-1998Q4, and the shaded areas indicate the 68% posterior credible interval.
Figure 2.13: The median generalised impulse responses to a contractionary monetary policy shock for Prior 1 specification.
Figure 2.14: The median generalised impulse responses to a contractionary monetary policy shock for Prior 2 specification.
Figure 2.15: The median generalised impulse responses, for period of 1996Q1, 1997Q3, 1998Q4, 2006Q1, 2008Q4, and 2010Q1, for all variables to a contractionary monetary policy shock for a TVP-VAR-SV model with a sample period of 1990Q1-2015Q4 and no capital controls index.
Figure 2.16: The median generalised impulse responses, for period of 2012Q1, 2013Q4 and 2015Q4, for all variables to a contractionary monetary policy shock for a TVP-VAR-SV model with a sample period of 1990Q1-2015Q4 and no capital controls index.
2.7 Appendix

This appendix is divided into three sections labelled A to C. Appendix A outlines the complete estimation details of the TVP-VAR-SV model. Appendix B outlines the procedure for computing the generalised impulse response functions and the implementation of sign restrictions. Lastly, Appendix C denotes the convergence statistics of the MCMC routine and the procedure for computing the DIC.

2.8 Appendix A

To simulate the posterior distribution, we use a six blocks Gibbs Sampler that sequentially draws from each full conditional posterior. The outline of the steps are:

1. Draw from $p(\beta_t^{(i)} | y, a_t^{(i-1)}, h_t^{(i-1)}, \Omega_{\beta}^{(i-1)}, \Omega_{\alpha}^{(i-1)}, \Omega_{h}^{(i-1)})$
2. Draw from $p(h_t^{(i)} | y, \beta_t^{(i)}, a_t^{(i-1)}, \Omega_{\beta}^{(i-1)}, \Omega_{\alpha}^{(i-1)}, \Omega_{h}^{(i-1)})$
3. Draw from $p(a_t^{(i)} | y, \beta_t^{(i)}, h_t^{(i)}, \Omega_{\alpha}^{(i-1)}, \Omega_{\alpha}^{(i-1)}, \Omega_{h}^{(i-1)})$
4. Draw from $p(\Omega_{\alpha}^{(i)} | y, \beta_t^{(i)}, h_t^{(i)}, \Omega_{\beta}^{(i-1)}, a_t^{(i)}, \Omega_{h}^{(i-1)})$
5. Draw from $p(\Omega_{h}^{(i)} | y, \beta_t^{(i)}, h_t^{(i)}, \Omega_{\beta}^{(i-1)}, \Omega_{\alpha}^{(i)}, a_t^{(i)})$
6. Draw from $p(\Omega_{\beta}^{(i)} | y, \beta_t^{(i)}, h_t^{(i)}, a_t^{(i)}, \Omega_{\alpha}^{(i)}, \Omega_{h}^{(i)})$
7. Repeat steps 1 to 6

where the superscript denotes the $i – th$ draw of the simulation. Primiceri (2005) uses standard Kalman filtering and smoothing techniques from Carter and Kohn (1994) to estimate the time-varying coefficients. However we adopt a different method, for the draws of Step 1 and 3 we use the algorithm derived from Chan and Jeliazkov (2009). For step 3 we use an auxiliary mixture sampler from Kim et al. (1998) and the estimation algorithm we used is from Chan and Hsiao (2014).
2.8.1 Step 1 Drawing $\beta$

The measurement equation of (2.2) in the paper can be rewritten into the form:

$$ y = X\beta + \epsilon, \quad \epsilon \sim N(0, \Sigma), \quad (2.13) $$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_T)'$, $y = (y_1, \ldots, y_T)'$, $\beta = (\beta_1, \ldots, \beta_T)'$, $\Sigma = diag(\Sigma_1, \ldots, \Sigma_T)$ and

$$ X = \begin{bmatrix} X_1 & 0 & \cdots & \cdots & 0 \\ 0 & X_2 & 0 & \cdots & 0 \\ 0 & \ddots & X_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & X_T \end{bmatrix}, $$

Next the transition equation of (2.3) in the paper can be rewritten into:

$$ H_\beta = \tilde{\alpha}_\beta + \nu, \quad \nu \sim N(0, S_\beta), \quad (2.14) $$

where $\tilde{\alpha}_\beta = (\beta'_0, 0, \ldots, 0)$, $S_\beta = diag(V_\beta, \Omega_\beta, \ldots, \Omega_\beta)$ and

$$ H = \begin{bmatrix} I_b & 0 & 0 & \cdots & 0 \\ -I_b & I_b & 0 & \cdots & 0 \\ 0 & -I_b & I_b & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -I_b & I_b \end{bmatrix}, $$

Thus equation (2.2) is $\beta \sim N(\alpha_\beta, (H'S_\beta^{-1}H)^{-1})$ and $\alpha_\beta = H^{-1}\tilde{\alpha}_\beta$.

The conditional posterior distribution is:

$$ p(\beta \mid y, a, h, \Omega_\beta, \Omega_a, \Omega_h) \propto p(y \mid \beta, a, h, \Omega_\beta, \Omega_a, \Omega_h)p(\beta), $$

$$ \propto |\Sigma|^{-\frac{1}{2}} \exp -\frac{1}{2}(y - X\beta)'\Sigma^{-1}(y - X\beta) \exp -\frac{1}{2}(\beta - \alpha_\beta)'H'S_\beta^{-1}H(\beta - \alpha_\beta), $$
\[\alpha \exp - \frac{1}{2} \beta' (X' \Sigma^{-1}X + H' S_{\beta}^{-1}H) \beta - 2\beta' (X' \Sigma^{-1}y + H' S_{\beta}^{-1}H \alpha_{\beta})],\]

Using the standard results from linear regression

\[p(\beta \mid y, a, h, \Omega_{\beta}, \Omega_{a}, \Omega_{h}) \sim N(\hat{\beta}, \Theta_{\beta}),\]

where

\[\Theta_{\beta} = (X' \Sigma^{-1}X + H' S_{\beta}^{-1}H)^{-1}, \quad \hat{\beta} = \Theta_{\beta} (X' \Sigma^{-1}y + H' S_{\beta}^{-1}H \tilde{\alpha}_{\beta}),\]

Note since for our priors we assumed \(\beta_{0} = 0\) then

\[\Theta_{\beta} = (X' \Sigma^{-1}X + H' S_{\beta}^{-1}H)^{-1}, \quad \hat{\beta} = \Theta_{\beta} (X' \Sigma^{-1}y),\]

To draw from \(N(\hat{\beta}, \Theta_{\beta})\), we use the algorithm from Chan and Jeliazkov (2009), that is we first take the Cholesky factor of \(\Theta_{\beta}\) which is \(\Theta_{\beta} = C_{\beta} C'_{\beta}\). Next we obtain \(Tk\) independent draws from a standard normal distribution \(N(0, 1)\) denoted as \(Z = (Z_1, \ldots, Z_{Tk})'\) and return \(\beta = \hat{\beta} + (C'_{\beta})^{-1}Z\). It is easy to check that the mean \(\beta\) is \(\hat{\beta}\) and its covariance matrix is

\[(C'_{\beta})^{-1}((C'_{\beta})^{-1}) = (C'_{\beta})^{-1}(C'_{\beta})^{-1} = (C_{\beta} C'_{\beta})^{-1} = \Theta_{\beta}^{-1}.\]

### 2.8.2 Step 2 Drawing \(h\)

To estimate the nonlinear stochastic volatility, we follow the methodology governed in Chan and Hsiao (2014) and rearrange equation (2.2) in the paper to be:

\[\tilde{y}_{t} = L_{t} \epsilon_{t},\]  \hspace{1cm} (2.15)

from this we know \(E[\tilde{y}_{t} \mid a_{t}, h_{t}, \beta_{t}] = 0\) and \(Var[\tilde{y}_{t} \mid a_{t}, h_{t}, \beta_{t}] = L_{t}(L'_{t} D_{t}^{-1}L_{t})^{-1}L'_{t} = D_{t}\). Therefore \((\tilde{y}_{it} \mid a, h, \beta) \sim N(0, e^{\frac{1}{2}h_{it}})\) where \(i = 1, \ldots, 4\) and each variable of stochastic volatility can be specified as:
\[
\bar{y}_{i,t} = e^{\frac{1}{2}h_{i,t}}\epsilon_{i,t}, \quad \epsilon_{i,t} \sim N(0, 1), \quad (2.16)
\]

Equation (2.4) is nonlinear model and as we are using mixture of linear Gaussian models to approximate it, we must first transform this measurement equation to become linear in the log-volatility of \( h_{i,t} \). Thus, we square both sides of (2.4) and take the logarithm:

\[
y^*_i,t = h_{i,t} + \epsilon^*_i,t, \quad (2.17)
\]

where \( y^*_i,t = \log(\bar{y}^2_{i,t}) \) and \( \epsilon^*_i,t = \log(\epsilon^2_{i,t}) \). In most cases \( y^*_i,t = \log(\bar{y}^2_{i,t} + c) \) for some small constant \( c \) and normally \( c = 10^{-4} \) to avoid numerical problems when \( y^*_i,t \) is close to zero. However \( \epsilon^*_i,t \) no longer follows a Gaussian distribution, it now follows a \( \log - \chi^2_1 \) distributions. According to Chan and Hsiao (2014), we can approximate the density of \( f(\epsilon^*_i,t) \) by a seven component Gaussian mixture such as:

\[
f(\epsilon^*_i,t) \approx \sum_{i=1}^{7} p_i \varphi(\epsilon^*_i,t; \mu_i - 1.2704, \sigma^2_i), \quad (2.18)
\]

where \( \varphi(\epsilon^*_i,t; \mu_i, \sigma^2_i) \) is the Gaussian density with \( \mu \) and variance \( \sigma^2 \) and \( p_i \) is the probability of the \( i \) – th mixture component for each point in time. The values of the parameters are given in Table 2.4. Chan and Hsiao (2014) emphasize that these values are fixed and do not depend on any unknown parameters. Equivalently (2.6) can be written in terms of an auxiliary random variables \( s_t \in \{1, \ldots, 7\} \) that serves as the mixture component indicator for each point at time such as

\[
(\epsilon^*_i,t \mid s_t = i) \sim N(\mu_i - 1.2704, \sigma^2_i), \quad (2.19)
\]

\[
\mathbb{P}(s_t = i) = p_i, \quad (2.20)
\]

Model (2.17) and the transition equation of (2.5) in the paper are now conditionally linear Gaussian given the component indicators \( s = (s_1, \ldots, s_T)' \). In terms of our study we can derive the joint distribution of \( p(h_i \mid y^*_i,s_i, \omega^2_{ni}) \) for each variable \( i \) by rewriting
(1) in matrix notation:

$$y^*_i = h^*_i + \epsilon^*_i,$$  \hspace{1cm} (2.21)

and

$$(\epsilon^*_i \mid s_i) \sim N(d_i, \Sigma_{y_i^*}),$$

where $d_i = (\mu_{s_1} - 1.2704, \ldots, \mu_{s_T} - 1.2704)'$, $\Sigma_{y_i^*} = diag(\sigma^2_{s_1}, \ldots, \sigma^2_{s_T})$ and the fixed parameters $\mu_1, \ldots, \mu_T$ and $\sigma^2_1, \ldots, \sigma^2_T$ are given in Table 2.4. By a simple change of variable, we have $(y^*_i \mid s_i, h_i) \sim N(h_i + d_i, \Sigma_{y_i^*})$ and the log likelihood is:

$$\log p(y^*_i \mid s, h_i) = -\frac{1}{2} (y^*_i - h_i - d_i)' \Sigma_{y_i^*}^{-1} (y^*_i - h_i - d_i) + c_1,$$  \hspace{1cm} (2.22)

Note the transition equation (2.5) in the paper can be rewritten for each variable $i$

$$h_{i,t} = h_{i,t-1} + \eta_{i,t}, \hspace{0.5cm} \eta_{i,t} \sim N(0, \omega^2_{h_{i,t}}),$$  \hspace{1cm} (2.23)

and we can rewrite (2.23) into matrix form

$$H_{h_{i,t}} h_i = \tilde{\alpha}_{h_{i,t}} + \eta_i, \hspace{0.5cm} \eta_i \sim N(0, \Phi),$$  \hspace{1cm} (2.24)

where $\tilde{\alpha}_{h_{i,t}} = (h_{0,i}, 0, \ldots, 0)$, $\Phi = diag(V_{h_{i}}, \omega^2_{h_{i}}, \ldots, \omega^2_{h_{i}})$ and

$$H_{h_{i,t}} = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
-1 & 1 & 0 & \cdots & \vdots \\
0 & -1 & 1 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & -1 & 1
\end{bmatrix},$$

Thus $(h_i \mid \Phi, \tilde{\alpha}_{h_{i}}) \sim N(\alpha_{h_{i}}, (H_{h_{i}}' \Phi^{-1} H_{h_{i}})^{-1})$, where $\alpha_{h_{i}} = H_{h_{i}}^{-1} \tilde{\alpha}_{h_{i}}$ and $|H_{h_{i}}| = 1$. Assuming $(H_{h_{i}}' \Phi^{-1} H_{h_{i}})^{-1} = \Sigma_{h_{i}}$, the log likelihood is:
\[
\log p(h_i | \Phi, \tilde{\alpha}_{h_i}) = \frac{1}{2} \log (2\pi) - \frac{1}{2} \log |\Sigma_{h_i}| - \frac{1}{2}(h_i - \alpha_{h_i})'\Sigma_{h_i}^{-1}(h_i - \alpha_{h_i}),
\]

\[
= \frac{-T}{2} \log(2\pi) - \frac{1}{2} \log V_{h_i} - \frac{T-1}{2} \log \omega_{h_i}^2 - \frac{1}{2}(h_i - \alpha_{h_i})'\Sigma_{h_i}^{-1}(h_i - \alpha_{h_i}), \tag{2.25}
\]

Using (2.23) and (2.25) we can derive the conditional posterior distribution

\[
p(h_i | y^*_i, s_i, \omega_{h_i}^2) \propto p(y^*_i | s_i, h_i)p(h_i | \Phi, \tilde{\alpha}_{h_i}),
\]

\[
= -\frac{1}{2}(y^*_i - h_i - d_i)'\Sigma_{y_i}^{-1}(y^*_i - h_i - d_i) - \frac{1}{2}(h_i - \alpha_{h_i})'\Sigma_{h_i}^{-1}(h_i - \alpha_{h_i}),
\]

\[
= -\frac{1}{2}[h_i'\Sigma_{h_i}^{-1} + \Sigma_{y_i}^{-1}]h_i - 2h_i'(\Sigma_{h_i}^{-1}\alpha_{h_i} + \Sigma_{y_i}^{-1}(y^*_i - d_i))],
\]

Since this log-density is quadratic in \(h_i\), it is Gaussian and therefore

\[
p(h_i | y^*_i, s_i, \omega_{h_i}^2) \sim N(\hat{h}_i, K_{h_i}^{-1}),
\]

where

\[
K_{h_i} = H_{h_i}'\Phi^{-1}H_{h_i} + \Sigma_{y_i}^{-1}, \quad \hat{h}_i = K_{h_i}^{-1}(H_{h_i}'\Phi^{-1}H_{h_i}\alpha_{h_i} + \Sigma_{y_i}^{-1}(y^*_i - d_i)),
\]

Note since for our priors we assumed \(h_0 = 0\) then

\[
K_{h_i} = H_{h_i}'\Phi^{-1}H_{h_i} + \Sigma_{y_i}^{-1}, \quad \hat{h}_i = K_{h_i}^{-1}(\Sigma_{y_i}^{-1}(y^*_i - d_i)),
\]

Since \(K_{h_i}\) is band matrix, \(\hat{h}_i\) can be easily obtained by solving the linear system

\[
K_{h_i}x = H_{h_i}'\Phi^{-1}H_{h_i}\alpha_{h_i} + \Sigma_{y_i}^{-1}(y^*_i - d_i),
\]

for \(x\), which avoids computing the inverse \(K_{h_i}^{-1}\). To draw from \(N(\hat{h}_i, K_{h_i}^{-1})\), we applied the same algorithm from Chan and Jeliazkov (2009) as in step 1.
Table 2.4: A Seven Component Gaussian Mixture for Approximating the \(\log - \chi^2_1\) distribution

<table>
<thead>
<tr>
<th>Component</th>
<th>(p_i)</th>
<th>(\mu_{i})</th>
<th>(\sigma_{i}^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00730</td>
<td>-10.12999</td>
<td>5.79596</td>
</tr>
<tr>
<td>2</td>
<td>0.10556</td>
<td>-3.97281</td>
<td>2.61369</td>
</tr>
<tr>
<td>3</td>
<td>0.00002</td>
<td>-8.56686</td>
<td>5.17950</td>
</tr>
<tr>
<td>4</td>
<td>0.04395</td>
<td>2.77786</td>
<td>0.16735</td>
</tr>
<tr>
<td>5</td>
<td>0.34001</td>
<td>0.61942</td>
<td>0.64009</td>
</tr>
<tr>
<td>6</td>
<td>0.24566</td>
<td>1.179518</td>
<td>0.34023</td>
</tr>
<tr>
<td>7</td>
<td>0.25750</td>
<td>-1.08819</td>
<td>1.26261</td>
</tr>
</tbody>
</table>

Source: Chan and Hsiao (2014)

2.8.3 Step 3 Drawing \(\alpha\)

To draw \(\alpha\), we can expand (2.15) and in the case where \(n = 4\) then (2.15) will be

\[
L_t \epsilon_t = \begin{bmatrix}
1 & 0 & 0 & 0 \\
\alpha_{21,t} & 1 & 0 & 0 \\
\alpha_{31,t} & \alpha_{32,t} & 1 & 0 \\
\alpha_{41,t} & \alpha_{42,t} & \alpha_{43,t} & 1
\end{bmatrix} \begin{bmatrix}
\epsilon_{1,t} \\
\epsilon_{2,t} \\
\epsilon_{3,t} \\
\epsilon_{4,t}
\end{bmatrix} = \begin{bmatrix}
\epsilon_{2,t} \\
\alpha_{21,t} \epsilon_{1,t} + \epsilon_{2,t} \\
\alpha_{31,t} \epsilon_{1,t} + \alpha_{32,t} \epsilon_{2,t} + \epsilon_{3,t} \\
\alpha_{41,t} \epsilon_{1,t} + \alpha_{42,t} \epsilon_{2,t} + \alpha_{43,t} \epsilon_{3,t} + \epsilon_{4,t}
\end{bmatrix},
\]

which can be rearranged into the form

\[
\begin{bmatrix}
\epsilon_{1,t} \\
\epsilon_{2,t} \\
\epsilon_{3,t} \\
\epsilon_{4,t}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
-\epsilon_{1,t} & 0 & 0 & 0 & 0 & 0 \\
0 & -\epsilon_{1,t} & -\epsilon_{2,t} & 0 & 0 & 0 \\
0 & 0 & 0 & -\epsilon_{1,t} & -\epsilon_{2,t} & -\epsilon_{3,t}
\end{bmatrix} \begin{bmatrix}
\alpha_{21,t} \\
\alpha_{31,t} \\
\alpha_{32,t} \\
\alpha_{41,t} \\
\alpha_{42,t} \\
\alpha_{43,t}
\end{bmatrix},
\]

Thus

\[
L_t \epsilon_t = \epsilon_t - E_t a_t, \quad L_t \epsilon_t \sim N(0, D_t),
\]

and the likelihood for (2.26) will be

\[
f(L_t \epsilon_t | a_t, \beta_t, h_t) \propto \prod_{t=1}^{T} |D_t|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \sum_{t=1}^{T} (L_t \epsilon_t)'D_t^{-1}(L_t \epsilon_t)\right\},
\]
\[
\propto (\prod_{t=1}^{T} |D_t|^{-\frac{1}{2}})^{\exp\left\{-\frac{1}{2} \sum_{t=1}^{T} (\epsilon_t - E_t a_t)' D_t^{-1} (\epsilon_t - E_t a_t)\right\}},
\]

which implies (2.26) is the same as the measurement equation

\[
\epsilon_t = E_t a_t + \gamma_t, \quad \gamma_t \sim N(0, D_t),
\]

(2.27)

With (2.27), we can now applied the same methodology as in step 1, that is in matrix notation

\[
\epsilon = E a + \gamma, \quad \gamma \sim N(0, D),
\]

(2.28)

where \( \epsilon = (\epsilon_1, \ldots, \epsilon_T)' \), \( a = (a_1, \ldots, a_T)' \), \( D = diag(D_1, \ldots, D_T) \) and

\[
E = \begin{bmatrix}
E_1 & 0 & \cdots & 0 \\
0 & E_2 & 0 & \cdots \\
0 & \cdots & E_T & \cdots \\
0 & \cdots & \cdots & 0 \\
0 & \cdots & 0 & E_T
\end{bmatrix},
\]

Next the transition equation of (2.4) in the paper can be rewritten into:

\[
H_a a = \tilde{\alpha}_a + \zeta, \quad \zeta \sim N(0, S_a),
\]

(2.29)

where \( \tilde{\alpha}_a = (a_0, 0, \ldots, 0) \), \( S_a = diag(V_a, \Omega_a, \ldots, \Omega_a) \) and

\[
H_a = \begin{bmatrix}
I_m & 0 & 0 & \cdots & 0 \\
-I_m & I_m & 0 & \cdots & 0 \\
0 & -I_m & I_m & \cdots & 0 \\
\vdots & \vdots & \cdots & \cdots & 0 \\
0 & \cdots & 0 & -I_m & I_m
\end{bmatrix},
\]

Thus equation (2.29) is \( a \sim N(\tilde{\alpha}_a, (H_a' S_a^{-1} H_a)^{-1}) \) and \( \alpha_a = H_a^{-1} \tilde{\alpha}_a \).

The conditional posterior distribution is:
\[ p(\mathbf{a} \mid \mathbf{y}, \mathbf{\beta}, \mathbf{h}, \Omega_\mathbf{\beta}, \Omega_\mathbf{a}, \Omega_\mathbf{h}) \propto p(\mathbf{y} \mid \mathbf{\beta}, \mathbf{a}, \mathbf{h}, \Omega_\mathbf{\beta}, \Omega_\mathbf{a}, \Omega_\mathbf{h}) p(\mathbf{a}), \]

\[ \propto |\mathbf{D}|^{-\frac{1}{2}} \exp -\frac{1}{2} (\mathbf{\epsilon} - \mathbf{Ea})' \mathbf{D}^{-1} (\mathbf{\epsilon} - \mathbf{Ea}) \exp -\frac{1}{2} (\mathbf{\alpha} - \mathbf{\alpha_a})' \mathbf{H_a}^{-1} \mathbf{H_a} (\mathbf{\alpha} - \mathbf{\alpha_a}), \]

\[ \propto \exp -\frac{1}{2} [\mathbf{a'}(\mathbf{E}^{-1} \mathbf{E} + \mathbf{H_a}^\prime \mathbf{S_a}^{-1} \mathbf{H_a}) \mathbf{a} - 2\mathbf{a'}(\mathbf{E}^{-1} \mathbf{\epsilon} + \mathbf{H_a}^\prime \mathbf{S_a}^{-1} \mathbf{H_a} \mathbf{\alpha}),], \]

Using the standard results from linear regression

\[ p(\mathbf{a} \mid \mathbf{y}, \mathbf{\beta}, \mathbf{h}, \Omega_\mathbf{\beta}, \Omega_\mathbf{a}, \Omega_\mathbf{h}) \sim N(\hat{\mathbf{a}}, \mathbf{\Xi_a}), \]

where

\[ \mathbf{\Xi_a} = (\mathbf{E}^{-1} \mathbf{E} + \mathbf{H_a}^\prime \mathbf{S_a}^{-1} \mathbf{H_a})^{-1}, \quad \hat{\mathbf{a}} = \mathbf{\Xi_a}(\mathbf{E}^{-1} \mathbf{\epsilon} + \mathbf{H_a}^\prime \mathbf{S_a}^{-1} \mathbf{H_a} \mathbf{\alpha}). \]

Note since for our priors we assumed \( \mathbf{a}_0 = \mathbf{0} \) then

\[ \mathbf{\Xi_a} = (\mathbf{E}^{-1} \mathbf{E} + \mathbf{H_a}^\prime \mathbf{S_a}^{-1} \mathbf{H_a})^{-1}, \quad \hat{\mathbf{a}} = \mathbf{\Xi_a}(\mathbf{E}^{-1} \mathbf{\epsilon}). \]

To draw from \( N(\hat{\mathbf{a}}, \mathbf{\Xi_a}) \) we apply the same algorithm from Chan and Jeliazkov (2009) as discussed in step 1 and 2.

**2.8.4 Step 4, 5, 6 Drawing \( \Omega_\mathbf{\beta}, \Omega_\mathbf{a}, \Omega_\mathbf{h} \)**

The diagonal elements of each \( \Omega_\mathbf{\beta}, \Omega_\mathbf{a}, \Omega_\mathbf{h} \) are conditionally independent given the data and the other parameters. Therefore

\[ (\omega^2_{\beta_i} \mid \mathbf{y}, \mathbf{h}, \mathbf{a}, \Omega_\mathbf{\beta}, \Omega_\mathbf{h}) \sim IG(10 + (T - 1)/2, 0.01 + \frac{1}{2} \sum_{t=2}^{T} (\beta_{i,t} - \beta_{i,t-1})^2) \quad i = 1, \ldots, b. \]

\[ (\omega^2_{\alpha_i} \mid \mathbf{y}, \mathbf{h}, \mathbf{\beta}, \Omega_\mathbf{\beta}, \Omega_\mathbf{h}) \sim IG(2 + (T - 1)/2, 0.01 + \frac{1}{2} \sum_{t=2}^{T} (\alpha_{i,t} - \alpha_{i,t-1})^2) \quad i = 1, \ldots, m. \]

\[ (\omega^2_{h_i} \mid \mathbf{y}, \mathbf{a}, \mathbf{\beta}, \Omega_\mathbf{\beta}, \Omega_\mathbf{a}) \sim IG(2 + (T - 1)/2, 0.01 + \frac{1}{2} \sum_{t=2}^{T} (h_{i,t} - h_{i,t-1})^2) \quad i = 1, \ldots, n. \]
2.9 Appendix B

2.9.1 Generalised Impulse Response Functions and Sign Restrictions

To implement sign restrictions we followed the methodology governed in Baumeister and Peersman (2013). To draw the candidate solutions for $A_t$ that satisfy the sign restrictions above, we first take the eigenvalue-eigenvector decomposition of the time-varying variance-covariance matrix $\Sigma_t = P_t \Lambda_t P_t'$ for time $t$. Next, we draw $n \times n$ matrix, denoted as $K$, from a standard normal distribution of $N(0, 1)$ and then the $QR$ decomposition of $K$ is taken (Rubio-Ramirez et al. 2010), which is denoted as $Q$. $Q$ is an orthonormal matrix, whereby the columns are orthonormal to each other. Thus, the impact matrix is computed as $A_t = P_t \Lambda_t^{\frac{1}{2}} Q'$ for time $t$. If the generated impact matrix $A_t$ satisfy the sign restrictions stated above, it is then used to compute the impulse response function. However, if this generated impact matrix $A_t$ does not satisfy the sign restrictions stated above, it is discarded and another candidate solutions for $A_t$ is drawn.

Since the TVP-VAR with stochastic volatility is a non-linear multivariate model, we must compute the generalised impulse response function in the spirit of Koop, Pesaran and Potter (1996). The generalised impulse response function is obtained from the difference between two conditional expectations with and without the exogenous shock:

$$IRF_{t+k} = \mathbb{E}[y_{t+k} \mid u_t, \gamma_t] - \mathbb{E}[y_{t+k} \mid \gamma_t], \quad (2.30)$$

where $y_{t+k}$ is the forecast of the endogenous variables at the horizon $k$, $\gamma_t$ represent the current information set and $u_t$ is the current structural disturbance terms. The current information set $\gamma_t$ contains the actual values of the lagged endogenous variables and a random draws of the model parameters and hyper parameters for each point of time. The computation of the generalised impulse response functions for a horizon $k$ can be summarised in 5 steps:
1. For each time $t$, we first draw $\beta_t$, $h_t$, and $a_t$ from the posterior distributions within the Gibb Sampler.

2. Next we draw the structural disturbances $u_t = (u_1, u_2, u_3)'$ from a standard normal distribution $N(0, 1)$. This allows us to derive the reduced form errors as $\epsilon_t = A_t u_t$.

3. We then generate two paths, one with the shock and the other without shock. For the latter case, we just compute $\epsilon_t = A_t u_t$ and then stochastically simulate a random path of length $k$ starting from the coefficients drawn from step 1. For the former case, we set $u_{i,t+1}$ to the corresponding shock that we are interested in. For example, say we are interested in the structural shock for the first variable, then the structural disturbance term will be $\tilde{u}_t = (u_1 + 1, u_2, u_3)'$. Thus, we compute $\tilde{\epsilon}_t = A_t \tilde{u}_t$ and then stochastically simulate another random path of length $k$.

4. Along the path $k$, we simulate the same reduced form shocks hitting both paths from $k + 1$ onwards and use the stochastically generated time-varying coefficients. The reason for this is to allow the system to be hit by other shocks along the time path.

5. To compute the impulse response function, we take difference between the two paths.

We repeat this procedure within the Gibb Sampler and 20,000 draws of impulse response functions for each shock and time are taken. Then the median is taken on these draws across each shock and time.
2.10 Appendix C

2.10.1 Markov Chain Monte Carlo Convergence

This appendix assesses the convergence of the MCMC algorithm described above. Primiceri (2005) states that in order to judge how well the chain mixes, it common practice to examine the autocorrelation function of the draws. Low autocorrelations suggest that the draws are almost independent, which increases the efficiency of the algorithm. Figure 2.17 plots the 20th order sample autocorrelation of the draws. Panel (a), (b), (c) and (d) corresponds to the $\beta$, $a$, $h$ and $\omega^2$ parameters respectively. For majority of the parameters, the sample autocorrelation are very small. The highest sample autocorrelation occurs in both the $\omega^2$ and $h$'s where it is around 0.7.

Table 2.5: Summary Distribution of Ineciency Factors for different set of parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>median</th>
<th>mean</th>
<th>min</th>
<th>max</th>
<th>10th percentile</th>
<th>90th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$'s</td>
<td>2.47</td>
<td>3.84</td>
<td>0.69</td>
<td>25.42</td>
<td>1.12</td>
<td>10.11</td>
</tr>
<tr>
<td>$a$'s</td>
<td>9.28</td>
<td>13.54</td>
<td>2.80</td>
<td>57.45</td>
<td>3.10</td>
<td>36.53</td>
</tr>
<tr>
<td>$h$'s</td>
<td>52.00</td>
<td>99.90</td>
<td>4.82</td>
<td>260.21</td>
<td>11.26</td>
<td>243.23</td>
</tr>
<tr>
<td>$\omega^2$'s</td>
<td>10.63</td>
<td>21.12</td>
<td>4.74</td>
<td>174.87</td>
<td>6.64</td>
<td>57.10</td>
</tr>
</tbody>
</table>

Another measure that is used to assess the the convergence of the MCMC algorithm is the inefficiency factors (IF). The IF is the inverse of the relative numerical efficiency measure of Geweke’s (1992) and it computed as $(1 + 2\sum_{k=1}^{\infty} \rho_k)$, where $\rho_k$ is the $k$ – th autocorrelation of the chain. We follow Primiceri (2005) and the IF estimates are performed using a 4 percent tapered window for the estimation of the spectral density at frequency zero. Normally, values of the IFs below or around twenty are regarded as satisfactory. Table 2.5 reports the summarised the distribution of the IFs for the posterior estimates of four sets of parameters. Except for the $h$'s parameters, it is clearly evident that on average the IFs for all the parameters are below or around 20. It appears that for certain parameters of $\omega^2$ the IFs are large. However, Primiceri (2005) noted that its not uncommon for these parameters to have IFs values between 4 and 75. In regards to the high IFs for the $h$'s parameters, this result is similar to Franta, Horvath and Rusnak (2014) and they note this potential inefficiency does not pose an issue unless the the impulse responses are normalised, which we do in our study.
2.10.2 DIC Estimation

To compute the DIC we need to evaluate the integrated likelihood which is

\[
p(y|\Omega_\beta, \Omega_a, \Omega_h, h_0, a_0) = \int p(y|\beta, a, \Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)p(\beta, a, h|\Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)d(\beta, a, h),
\]

\[
= \int p(y|h, \Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)p(h|\Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)dh,
\]

(2.31)

Chan and Eisenstat (2015) uses an importance sampling estimator to estimate the integrated likelihood above

\[
p(y|\Omega_\beta, \Omega_a, \Omega_h, h_0, a_0) = \frac{1}{R} \sum_{r=1}^{R} \frac{p(y|h^r, \Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)p(h^r|\Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)}{g(h^r; \Omega_\beta, \Omega_a, \Omega_h, h_0, a_0)},
\]

(2.32)

where \(h^1, \ldots, h^R\) are draws from the importance sampling density \(g\) that might depend on the parameters. Therefore, the DIC can then be obtained by simply averaging the
integrated likelihood (2.32) over the posterior draws. Please see Chan and Eisenstat (2016) for further information on the estimation algorithm of the integrated likelihood.
3 Chapter 3

Forecasting Structural Change and Fat-Tailed Events in Australian Macroeconomic Variables

This paper is published in the Journal Economic Modelling

This paper is coauthored with Jamie Cross and we both contributed 50 per cent each to the paper.

______________________________
Collaborating Author declaration
3.1 Introduction

Structural change refers to variation in the fundamental behavior of macroeconomic time series. Causes of structural change range from unanticipated events such as financial crisis (Hamilton and Lin, 1996; Hamilton, 2005) to man made changes in macroeconomic policy (Primiceri, 2005; Sims and Zha, 2006; Kudrna, Tran and Woodland, 2015). Figure 3.1 shows that key Australian macroeconomic variables: real GDP growth, CPI inflation and a short-term interest rate - the 90 day Bank Accepted Bills/Negotiable Certificates of Deposit - have undergone significant structural changes since the 1970's. Inflation was particularly high during the mid to late 1970's and 1980's and low in the last decade with interesting variations in and around the 2007/08 global financial crisis (GFC). Next, whilst actual real GDP doubled over the past decade, business cycle fluctuations have substantially moderated in the last 20 years. Finally, the adoption of inflation targeting by the Reserve Bank of Australia (RBA) in 1992/93 has seen a dramatic decline in short-term interest rate volatility over the sample period.

In addition to being subject to endogenous structural change, the modern market economy is also exposed to ubiquitous and diverse macroeconomic ‘shocks’. Broadly speaking, these shocks can be categorized into two types: anticipated shocks and unanticipated shocks. Anticipated shocks, such as seasonal changes in tastes and preferences, do not significantly alter the pattern of macroeconomic activities and can be factored into policy decisions. Unanticipated shocks, such as unanticipated tax cuts, can have temporary or permanent effects on real economic activity (Mertens and Ravn, 2011). Although such shocks are a natural driver of the ebbs and flows of the business cycle, outlier or fat-tailed shocks have varying and often significant macroeconomic implications. For instance large unanticipated shocks, such as the oil price shocks of the 1970's, or the 2007/08 Global Financial Crisis (GFC), are difficult to forecast and may result in temporary or permanent structural changes within the economy making the policy responses difficult (see e.g., Hamilton (1983) for the former and Mian and Sufi (2010) for the latter).
In this paper, we investigate whether the incorporation of time variation and fat-tails into traditionally Gaussian, fixed coefficients multivariate and univariate autoregressive models leads to enhanced forecast performance of key Australian macroeconomic variables: real GDP growth, CPI inflation and a short-term interest rate. As discussed in D’Agostino, Gambetti and Giannone (2013), the answer to this question is far from trivial. On the one hand, it seems obvious that if the economy is subject to structural change then any forecasting model that can account for such changes would be better suited, thus increasing forecast accuracy. On the other hand, a richer modelling structure implies a higher number of parameters, thus increasing the risk of estimation errors and possibly reducing forecast accuracy.

The class of univariate autoregressive (AR) and multivariate vector autoregressive (VAR) models includes the following specifications: constant parameter, constant parameter with stochastic volatility, time varying parameter and time varying parameter with stochastic volatility. Set in this manner, we allow for time variation through two sources: (1) in the models coefficients and (2) in the variance of the shocks. For the multivariate models we follow Primiceri (2005) and consider a third source of time variation via the covariance terms. In addition to accounting for time variation within the coefficients and volatilities, all models are estimated under both Gaussian and Student-t error distributions. A consequence of this modelling feature is that it leads to faster adaptation to large fluctuations, making it more appropriate model during times of economic uncertainty. For instance, when considering financial spillovers in macroeconomic linkages amongst developed countries throughout the GFC period, Ciccarelli et al (2016) provide evidence that a panel VAR model with Student’s-t distributed errors enhances the in sample fit of a panel VAR with Gaussian errors. In addition to this class of models we also consider the forecast performance of non-linear regime switching as well as rolling-window ARs and VARs. The former class of models have been shown to generate a good description of the evolution of monetary policy and inflation dynamics in the US economy (Sims and Zha, 2006), whilst the latter class of models are simpler, implying that any forecast improvements would have significant practical implications.
Our paper is related to the growing literature on modelling structural instabilities as well as the reviving literature on the modelling of fat tailed events. In the first line of literature Cogley and Sargent (2001, 2005) and Primiceri (2005) pioneered the work on the time-varying parameter vector autoregression with stochastic volatility in the variance covariance matrix (TVP-VAR-SV). The TVP-VAR-SV model has since been a catalyst in the literature on the identification of structural instabilities within the monetary policy transmission mechanisms of various economies (see e.g. Benati (2008), Nakajima et al. (2011), Cross (2016) or Poon (2016)). Important for this study, Cross (2016) shows that stochastic volatility is an important modelling feature when examining the in-sample properties of Australian macroeconomic data. Despite this growing literature a major criticism of economic modelling has been the inability to predict the 2007-08 Global Financial Crisis (GFC) (see, for instance Ng and Wright (2013)). Since then, researchers have began investigating whether the class of aforementioned autoregressive models can enhance the forecastability of financial and macroeconomic variables (see, for instance: D’Agostino, Gambetti and Giannone (2013), Barnett, Mumtaz and Theodoridis (2014), Bekiros (2014), Baxa, Plašil and Vašiček (2015) or Charfeddine (2016)). For instance D’Agostino et al (2013) and Barnett et al (2014) utilize the TVP-VAR-SV to respectively forecast US and UK macroeconomic indicators. Both studies conclude that the TVP-VAR-SV model produces superior forecasts as compared to a traditional fixed coefficients VAR model, however they lack a systematic comparison of the various nested VAR models listed above. The next line of research revives the earlier work of Geweke (1993, 1994) and Ni and Sun (2005), by incorporating Student’s-t errors (Student, 1908) into macroeconomic models to allow for the possibility of fat-tailed events. For instance Chib and Ramamurthy (2014) show that incorporating fat-tails improves the in-sample fit of a traditional US calibrated DSGE model with Gaussian errors. In addition, Chiu et al (2015) suggests that incorporating both fat-tails and stochastic volatility is fruitful in forecasting US macroeconomic and financial data.

Methodologically our paper is most similar to the recent study by Chiu et al (2015) who investigate the importance of fat-tails and stochastic volatility in forecasting US
data. We highlight that our study differs from Chui et al (2015) in three ways. First, rather than solely focusing on VARs we forecast with both multivariate and univariate autoregressive models. This is important for at least two reasons. First, a well known feature of macroeconomic forecasting is that multivariate models have struggled to out-predict univariate models (see, for instance; Nelson (1972), Atkeson and Ohanian (2001), Stock and Watson (2007), D’Agostino and Surico (2012) or Chauvet and Potter (2013)). Second, as shown by Clark and Ravazzolo (2015), when considering US data the AR and VAR models including stochastic volatility produce comparable forecast results, with the AR outperforming the VAR in inflation forecasts and the VAR providing superior interest rate forecasts. The next difference from our study and that of Chui et al (2015) is that we provide a more rigorous and systematic comparison of models. To be specific, in their paper Chui et al (2015) compare the forecast performance of their time-varying parameter VAR with stochastic volatility and fat-tails to three alternative specifications: (1) a time-invariant parameter VAR with stochastic volatility and Gaussian errors, (2) a time-invariant VAR without stochastic volatility and Gaussian errors and (3) a time-invariant VAR without stochastic volatility and fat-tailed errors. In our paper we allow for all possible combinations of models with and without time-varying parameters and stochastic volatility under both Gaussian and Student’s-t error distributions. This comparison is critical in establishing a clear distinction between the forecast contributions made by each element of the respective models. In addition, we also offer a more complete model comparison in that we consider the forecast performance of regime switching and rolling-window VAR models. Finally, as opposed to traditional Kalman filter estimation methods, our estimation utilizes efficient precision sampler techniques adopted from Chan and Jeliazkov (2009) and Chan and Hsiao (2014).

The full sample consists of quarterly data between 1969Q4 and 2014Q3. To allow for comparability of all models at various forecast horizons the main forecast period runs from 1992Q1 to 2011Q3. Set in this manner we replicate the Reserve Bank of Australia’s forecasting responsibilities since adopting inflation targeting. Forecasts are conducted
over one quarter, one year, two years and three years using a pseudo out-of-sample methodology. Density forecasts are constructed via the predictive density and point forecasts are taken to be the mean of the predictive density. Point forecast accuracy is measured by the mean squared forecast error (MSFE), whilst the performance of density forecast is measured by the log of the predictive likelihoods (LPL).

The results yield four important findings. First, fat-tailed models consistently outperform their Gaussian counterparts. Second, adding time varying parameters and stochastic volatility improves forecast performance across all variables given a constant benchmark. Third, Student-t distributed stochastic volatility models are found to generate more accurate density forecasts as compared to all Gaussian counterparts. Taken together these results suggest that both structural instabilities and fat-tail events are important features in modelling Australian macroeconomic variables. Finally, when comparing the forecast accuracy of univariate and multivariate models we obtain the striking result that a simple rolling window autoregression with fat-tails produces the most accurate real GDP growth forecasts, however the time varying vector autoregression with stochastic volatility and fat-tails produces the best interest and inflation forecasts. From an inflation targeting perspective, this means that multivariate inflation forecasting models which are able to take into account potential structural instabilities and fat-tail events provide important information for central bankers policy decisions.

The rest of the paper is structured as follows. Section 2 presents the forecasting models. Section 3 presents the data and forecast metrics. Section 4 presents the full sample results, section 5 presents the intertemporal forecasting results and section 6 concludes.

3.2 Models

In this section we present the forecasting models used in this study. In order to distinguish between structural changes and fat tailed events we employ a range of ARs and VARs. Because in AR models are simplifications of corresponding VAR models in which the number of variables is equal to one, we save space by only presenting the...
VAR specifications. Moreover, since many of our model variants are nested versions of a more complete model specification, we only present the most complex model in each class. For instance, the traditional time invariant VAR put forth by Sims (1980) is a nested version of the time varying parameter VAR with stochastic volatility developed by Primiceri (2005) in which there is no time variation within the coefficients or the covariance matrix. All of the models are estimated via Bayesian methods with priors and estimation algorithms for each model provided in the Appendix.

3.2.1 Regime Switching Models

Following Barnett et al (2014) we examine the possibility of structural shifts by employing a regime switching VAR of the following form:

\[ y_t = c_{S_t} + \sum_{j=1}^{p} A_{j,S_t} y_{t-j} + \epsilon_t, \quad \epsilon_t \sim N(0, \Omega_{S_t}) \]  

(3.1)

where \( y_t \) is a \( T \times n \) data matrix, \( c_{S_t}, A_{j,S_t} \) and \( \Omega_{S_t} \) are regime dependent intercepts, autoregressive coefficients and variance-covariance matrices respectively. Following Chib (1998) the break dates are modeled via the latent variables \( S_t \) for the VAR coefficients and \( H_t \) for the error covariance matrix. In other words, the specification allows for \( M \) structural breaks at unspecified dates. For estimation purposes (1) can be written in the form of a seemingly unrelated regression (SUR) model:

\[ Y_t = X_t \beta_{S_t} + \epsilon_t, \]  

(3.2)

where \( X_t = I_N \otimes (1, Y_{t-1}', \ldots, Y_{t-p}') \) and \( \beta_{S_t} = vec([c_{S_t}, A_{1,S_t}, \ldots, A_{p,S_t}])' \). In the estimation we impose a standard Normal-Wishart prior:

\[ \beta_{S_t} \sim N(\beta_0, V_{\beta}), \]  
\[ \Omega_{S_t} \sim IW(\tau, \Sigma). \]  

(3.3)

where \( IW(\tau, \Sigma) \) is the Inverse Wishart distribution with degree of freedom parameter \( \tau \geq p \) and positive definite scale matrix \( \Sigma \).
In the most general form, the state variables are assumed to evolve independently with their transitions governed by first-order Markov chains with $M + 1$ regimes, restricted transition probabilities $p_{ij} = p(S_t = j|S_{t-1} = i)$ and $q_{ij} = p(H_t = j|H_{t-1} = i)$. The transition probability matrices are defined as:

$$p_{ij}, q_{ij} = \begin{cases} 
> 0 & \text{if } i = j \\
> 0 & \text{if } j = i + 1 \\
= 1 & \text{if } i = j = M \\
= 0 & \text{otherwise} 
\end{cases} \quad (3.4)$$

For instance if $M = 3$ then the transition matrices are defined as:

$$\tilde{P} = \begin{bmatrix}
p_{11} & 0 & 0 \\
1 - p_{11} & p_{22} & 0 \\
0 & 1 - p_{22} & 1 
\end{bmatrix},$$

$$\tilde{Q} = \begin{bmatrix}
q_{11} & 0 & 0 \\
1 - q_{11} & q_{22} & 0 \\
0 & 1 - q_{22} & 1 
\end{bmatrix}.$$
of regimes becomes larger. To distinguish between breaks in mean and variance we estimate two versions of the proposed model:

1. The joint switching model as set out above which allows for independent breaks in the VAR coefficients and error covariances (JSRS-VAR); and
2. An independent switching model in which the breaks in VAR coefficients and the covariance matrix are restricted to occur jointly (ISRS-VAR).

Set in this manner specification 2 is able to gauge the forecast performance of allowing for different timing in variance and coefficient breaks. In each case we allow for up to three breaks or four regimes. For notation purposes a model with 2 regimes is denoted $RS(q) - VAR$ where $q = 2, 3, 4$. The optimal numbers of regimes for each model are chosen at each date in the sample by maximizing the marginal likelihood. The computation of the marginal likelihood via the Chib (1995) method. A detailed description of the calculation of the marginal likelihood for change point models can be found in Bauwens and Rombouts (2012). Estimation details are provided in Appendix B. We highlight the fact that whilst we employ a normal inverse Wishart prior on the VAR parameters in each regime, as described in Appendix B, the tightness parameters are set to large values, hence rendering the prior distributions non-informative.

3.2.2 Time-varying Models

Following Primiceri (2005) the general time varying parameter vector autoregression with stochastic volatility (TVP-VAR-SV) model with $n$ variables and $p$ lags is given by:

$$y_t = b_t + \sum_{i=1}^{p} B_{i,t} y_{t-i} + u_t, \ u_t \sim N(0, \Sigma_t), \ (3.5)$$

where $y_t$ is an $n \times 1$ vector of variables of interest, $b_t$ is an $n \times 1$ vector of time varying intercepts, $B_{i,t}, i = 1, \ldots, p$, are $n \times n$ matrices of time varying VAR coefficients and
\( \Sigma_t \) is an \( n \times n \) time varying error-covariance matrix. For estimation purposes (5) can be written in the form of a seemingly unrelated regression (SUR) model:

\[
y_t = X_t \beta_t + u_t,
\]

(3.6)

where \( X_t = I_n \otimes \begin{bmatrix} 1 & y'_{t-1} & \ldots & y'_{t-p} \end{bmatrix} \) and \( \beta_t = vec \left( \begin{bmatrix} b_t & B_{1,t} & \ldots & B_{p,t} \end{bmatrix}' \right) \). Note that \( \otimes \) denotes the Kronecker Product and \( vec(\cdot) \) is a vectorization operation that takes the intercept and the VAR coefficients and stacks them into a \( k \times 1 \) vector equation by equation where \( k = n (np + 1) \).

To model the time varying error covariance matrix \( \Sigma_t \), it is common to decompose it into two matrices \( L_t \) and \( D_t \) in which \( L_t \) is a lower triangular matrix with ones along the main diagonal and the contemporaneous interactions amongst the endogenous variables as the off diagonal elements, and \( D_t \) is a diagonal matrix that contains the exogenous disturbances. Following Primiceri (2005) this is completed using an LDL decomposition:

\[
\Sigma_t = (L_t' D_t^{-1} L_t)^{-1}.
\]

(3.7)

For instance, for \( n = 3 \):

\[
L_t = \begin{bmatrix} 1 & 0 & 0 \\ a_{21,t} & 1 & 0 \\ a_{31,t} & a_{32,t} & 1 \end{bmatrix}, \quad D_t = \begin{bmatrix} e^{h_{1,t}} & 0 & 0 \\ 0 & e^{h_{2,t}} & 0 \\ 0 & 0 & e^{h_{3,t}} \end{bmatrix}.
\]

For notational convenience let \( \mathbf{h}_{t} = (h_{1,t}, h_{2,t}, h_{3,t}, \ldots, h_{n,t})' \) and \( \mathbf{h}_{i,\bullet} = (h_{i,1}, \ldots, h_{i,T})' \). That is, \( \mathbf{h}_{t} \) is an \( n \times 1 \) vector obtained by stacking \( h_{i,t} \) by the first subscript whilst \( \mathbf{h}_{i,\bullet} \) is the \( T \times 1 \) vector obtained by stacking the second subscript. Next, let \( \mathbf{a}_t \) denote the vector of covariance terms collected row wise from \( L_t \) i.e. \( \mathbf{a}_t = [a_{21}, a_{31}, a_{32}, \ldots, a_{n(n-1)}]' \) so that \( \mathbf{a}_t \) is an \( m \times 1 \) vector of parameters where \( m = n (n-1) / 2 \). Then, the state equations for the time varying coefficients are given by:
\[ \beta_t = \beta_{t-1} + \nu_t, \nu_t \sim N(0, \Omega_\beta), \quad (3.8) \]
\[ a_t = a_{t-1} + \psi_t, \psi_t \sim N(0, \Omega_a), \quad (3.9) \]
\[ h_{*,t} = h_{*,t-1} + \eta_t, \eta_t \sim N(0, \Omega_h), \quad (3.10) \]

for \( t = 2, \ldots, T \), where \( \Omega_\beta = \text{diag}(\omega_{\beta_1}^2, \ldots, \omega_{\beta_k}^2) \), \( \Omega_a = \text{diag}(\omega_{a1}^2, \ldots, \omega_{am}^2) \) and \( \Omega_h = \text{diag}(\omega_{h1}^2, \ldots, \omega_{hn}^2) \), where all elements are assumed to follow independent Inverse Gamma distributions. The states are initialized as follows:

\[ \beta_1 \sim N(\beta_0, V_\beta), \quad a_1 \sim N(a_0, V_a), \quad h_1 \sim N(h_0, V_h), \quad (3.11) \]

where \( \beta_0, a_0, h_0, V_\beta, V_a \) and \( V_h \) are all assumed to be known. Estimation details are provided in Appendix C.

In order to distinguish between the importance of allowing for time variation in both the coefficients and the volatility of exogenous shocks we estimate three alternative models. They are:

1. A VAR with constant coefficients and constant covariance-variance matrix (CVAR);
2. A VAR with time varying coefficients and constant covariance-variance matrix (TVP-VAR); and
3. A VAR with constant coefficients and time varying covariance-variance matrix (CVAR-SV).

All of the above models are nested in (1) and can be estimated using the framework described in Appendix B. To be clear, the TVP-VAR is a nested version of the TVP-VAR-SV model with the only difference being that the covariance-variance matrix is constant i.e. \( \Sigma = \Sigma_1 = \cdots = \Sigma_T \). In this case we follow standard conventions and set \( \Sigma \sim IW(\nu_\Sigma, S_\Sigma) \). Next, the CVAR is a nested version of the TVP-VAR with the
only difference being that the parameters are not time varying i.e. $\beta = \beta_1 = \cdots = \beta_T$.

In this case we set $\beta \sim N(\bar{\beta}_0, \bar{\Sigma}_\beta)$. Finally, the CVAR-SV model is a nested version of the TVP-VAR-SV model with the only difference being that the parameters are not time varying. In this case we set the same prior for $\beta$ as in the case of the CVAR and the same prior for $\Sigma_t$ as in the TVP-VAR-SV model. As for lag length, allowing for a possible lag length of $p = 1, \ldots, 10$ we estimate a time invariant VAR model along with the data described in Section 3.1 and find that the Bayesian information criterion (BIC) selects two quarters as the optimal lag length. To facilitate a direct comparison of the models forecast performance we consequently estimate all other models using a lag length of two quarters. This specification also allows for direct comparison with studies by Barnett et al. (2014), Chui et al. (2015) and D’Agostino et al. (2013) on the Euro Area and US economies, which also use a lag length of two quarters in specifying their autoregressive models.

3.2.3 Rolling-window Models

The final VAR model is the rolling-window VAR:

$$y_t = b + \sum_{i=1}^{p} B_i y_{t-i} + u_t, \ u_t \sim N(0, \Sigma), \tag{3.12}$$

where $y_t$ is an $n \times 1$ vector of variables of interest, $b$ is an $n \times 1$ intercept vector, $B_i, i = 1, \ldots, p$, are $n \times n$ matrices of VAR coefficients and $\Sigma$ is an $n \times n$ covariance matrix. The rolling VAR model uses a 10-year rolling window to estimate the model parameters. Clearly the specification in (11) is much simpler than those in (1) and (5). Consequently, any finding that this model forecasts relatively well as compared to the more sophisticated alternatives has significant practical importance. We highlight that estimation of this model is identical to the constant VAR model nested in (5). To ensure comparability with the earlier specified autoregressive models, in the estimation process we use a lag length of two.
3.2.4 Stochastic Volatility under Student’s-t Distributed Errors

In this section we show how to model stochastic volatility with Student’s-t distributed errors for the basic VAR model. From a methodological perspective, the Student’s-t distribution is a robustification of a Gaussian distribution which places more weight on tail events. In fact, it is easy to show that Student’s-t distribution is a simple mixture of a Gaussian and Inverse Gamma distribution of the form:

\[
\begin{align*}
\mathbf{u}_t|\lambda_{i,t} & \sim N(0, \mathbf{D}_t), \\
\lambda_{i,t}|\nu & \sim IG\left(\frac{\nu_i}{2}, \frac{\nu_i}{2}\right),
\end{align*}
\]

(3.13) (3.14)

where the diagonal matrix \( \mathbf{D}_t = diag(\lambda_{1,t}e^{h_{1,t}}, \ldots, \lambda_{n,t}e^{h_{n,t}}) \) is complementary with the specification in Equation (7) and \( \nu_i \) denotes the degrees of freedom parameter from the Student’s-t distribution which follows a uniform distribution:

\[
\nu_i \sim U(0, \bar{\nu}).
\]

(3.15)

In theory \( \bar{\nu} \) can be set to any positive real number. In our empirical analysis we set \( \bar{\nu} = 50 \). This seems reasonable given the plots in Section 3.3. We highlight the non-informative nature of the uniform prior on the degrees of freedom parameter. Also note that modelling time varying Student’s-t distributed errors for associated AR models is equivalent to modelling a single \( \lambda_i \). Estimation details are provided in Appendix D.

3.3 Data and Forecast Metrics

3.3.1 Data

The full sample consists of quarterly data between 1969Q4 and 2014Q3. In line with the macroeconomics forecasting literature the variables of interest are real GDP growth, inflation and a short-term interest rate taken to be the 90 day Bank Accepted Bills/Negotiable Certificates of Deposit (here on simply referred to as the interest rate). Inflation and real GDP data is sourced from the Australian Bureau of Statistics and are taken to be
the Consumer Price Index (all groups) and seasonally adjusted real GDP respectively.
Interest rate data is sourced from the Reserve Bank of Australia. We note that the short term interest rate is taken to be the quarterly average of the monthly data whilst GDP and CPI are converted to annualized growth rates.

3.3.2 Degrees of Freedom

It is well known that the probability density function (pdf) of the Student’s t distribution converges to the pdf of a Gaussian distribution with zero mean and unit variance as the degree of freedom parameter goes to infinity (see, for instance: Kroese and Chan (2014, p.50)). For this reason it is useful to plot the degree of freedom parameter for each model prior to forecasting. Both the univariate and multivariate inflation results are in Figure 3.2, whilst the output and interest rate results are in Figure 3.3 and Figure 3.4 respectively.

It is immediately clear that there exists a substantial degree of time variation in the degrees of freedom parameter across all variables. Figure 3.2 shows that models with both stochastic volatility and Student t errors are able to capture the structural break in inflation following the introduction of the goods and services tax (GST) in the year 2000. Figure 3.4 shows that a similar case exists for the interest rate. In that case models with both stochastic volatility and Student t errors are able to capture the structural break in the interest rate in the 2007-08 GFC. Finally, whilst Figure 3.3 shows that the degree of freedom parameters for GDP are declining over time, the models with both stochastic volatility and Student t errors show little evidence of time-variation in the degree of freedom parameter. This suggests that the decline in the degree of freedom parameter for non-stochastic volatility models is noise entering the system due to changes in volatility of the error term.

Before proceeding to the full sample results, an aggregate measure of the degrees of freedom across all eight models with a Student’s t distribution is provided in Figure 3.5. The mode of the degrees of freedom parameter for inflation, GDP and interest rate models is three, twenty-nine and three respectively. Since these parameter values are
quite small, we have a prior belief that models with Student-t errors will provide more accurate forecasts as compared to their Gaussian counterparts.

3.3.3 Forecast Metrics

In this section we discuss the forecast metrics along with a brief discussion of how to implement the recursive out-of-sample forecasting methodology. To this end, let $y_{1:t}$ denote the data from the initial time period up until time $t$ and $\hat{y}_{t+h}$ represent the vector of $h$-steps-ahead forecasts with $h = 1, 4, 8$ and $12$. Each of the models produce both point and density forecasts. Density forecasts are obtained by the predictive density: $f(\hat{y}_{t+h}|y_{1:t})$, and point forecasts are taken to be the mean of the predictive density: $E[y_{t+h}|y_{1:t}]$.

To conduct the forecasting exercise we utilize predictive simulation. This begins by estimating the model parameters using data between 1978:Q1 and 1992:Q1. We then forecast observations between 1992:Q1 to 2011:Q3. The reason for choosing this period is that it replicates the central banks forecasting responsibilities since formally adopting inflation targeting. To produce a $h$-step ahead forecast let $t_0$ denote 1992:Q1. Next, conditioning upon the model parameters up to time $t_0$, use the MCMC draws along with the relevant transition equations to simulate the future states up to time $t_0 + h - 1$. For instance, in simulating the log-volatility: $h_s$, we use the relative state equation provided by equation (6) and draw $\eta_s \sim N(0, \Omega_h)$ conditional upon $h_{s-1}$ for $s = t_0 + 1, \ldots, t_0 + h - 1$. These forecasts are then averaged over all the posterior draws to produce estimates for $E[y_{t+h}|y_{1:t}]$ and $f(y_{t+h}|y_{1:t})$. The exercise is then repeated using data up to time $t_0 + 1$ and so on.

We now discuss the forecast metrics for both point and density forecasts. To this end, let $y^o_{t+h}$ denote the observed value of the data at time $t + h$. The metric used to

\footnote{As discussed in Cross (2015) the exact date that Australia adopted inflation targeting is blurred. The formal announcement of an inflation target was made in 1996 in the Statement on the Conduct of Monetary Policy (Reserve Bank of Australia, 1996), however reference of such a target was made in speeches by then Governor of the RBA Bernie Fraser as early as 1992-93 (Fraser, 1992;1993(a),1993(b)). Without loss of generality we commence forecasts from 1992Q1.}
evaluate the accuracy of the point forecasts is the mean squared forecast error (MSFE) which is defined by:

$$MSFE = \frac{1}{T - h - t_0 + 1} \sum_{t=t_0}^{T-h} (y_{t+h}^o - y_{t+h|t})^2.$$  \hspace{1cm} (3.16)

In order to facilitate an easier comparison we then compute the relative mean squared forecast errors (RMSFE) subject to a CVAR benchmark. The RMSFE is defined as the ratio between the MSFE of a specific model and the MSFE of the CVAR. Mathematically the RMSFE is defined by:

$$RMSFE_i = \frac{MSFE_i}{MSFE_{CVAR}},$$ \hspace{1cm} (3.17)

where $i$ denotes the model of interest. A RMSFE of less than one indicates that the specific model outperforms the CVAR whilst a relative MSFE of greater than one indicates inferior forecast performance.

The metric used to evaluate the density forecasts is the the predictive likelihood: $f(y_{t+h} = y_{t+h|y_{1:t}})$, which is the predictive density of $y_{t+h}$ evaluated at the observed value $y_{t+h}^o$. We evaluate the density forecasts using the mean score of the log of the predictive likelihoods:

$$LS = \frac{1}{T - h - t_0 + 1} \sum_{t=t_0}^{T-h} \log f(y_{t+h} = y_{t+h|y_{1:t}}).$$ \hspace{1cm} (3.18)

If the actual outcome $y_{t+h}^o$ is unlikely under the density forecast then the value of the predictive likelihood will be small, and vice-verse. When interpreting this metric a larger value indicates better forecast performance (for a more detailed discussion of the predictive likelihood see Geweke and Amisano (2011) ). Forecast comparison is then completed using relative sum of the log of the predictive likelihoods (RLPL) subject to a CVAR benchmark. The RLPL is defined as the difference between the log score of the i-th model and the CVAR. Mathematically the RLPL for model $i$ is defined by:

$$RLPL_i = LS_i - LS_{CVAR}.$$ \hspace{1cm} (3.19)
Set in this manner, a model with a positive RLPL outperforms the CVAR benchmark whereas a model with a negative RLPL fails to outperform the CVAR benchmark.

3.4 Full Sample Results

In this section we present the point and density forecast results over the entire sample. Univariate and multivariate point forecast results are presented in Tables 3.1 and Table 3.2 respectively whilst Table 3.3 compares the best univariate and multivariate point forecasting models. Similarly, the univariate and multivariate density forecast results are presented in Tables 3.4 and Table 3.5 respectively whilst a comparison of the best univariate and multivariate point forecasting models is in Tables 3.6.

3.4.1 Point Forecast Results

The results in Table 3.1 suggest that the TVP-AR-SV model improves upon the average forecast performance of the standard AR model across all three variables. Adding stochastic volatility is particularly useful for forecasting both interest and inflation rates, however GDP is relatively harder to forecast. This is seen by the similar average forecast performance of the AR and AR-SV models. Interestingly, accounting for instability in the AR coefficients of the GDP equation is more promising. By comparing the average forecast performance of the TVP-AR and AR models it is clear that the TVP-AR model is preferred. The results also show that fat-tails enhances forecast performance across all variables. This is seen by the fact that the TVP-AR-SVt model is the best forecasting model inflation whilst the TVP-AR-t and AR-SVt models respectively provide the most accurate GDP and interest rate forecasts. It’s also worth noting the poor performance of regime switching models as compared to models with SV.

The results in Table 3.2 point to similar findings for multivariate point forecasts. In line with the univariate forecast results, when comparing the VAR and CVAR-SV results, adding SV is shown to enhance forecast accuracy across all three variables. Similar to the univariate case, accounting for instability in the VAR coefficients enhances both
inflation and output forecasts whilst accounting for instability in the shocks improves interest rate forecasts. Also in line with the univariate results, fat-tails are shown to enhance the forecast performance of all variables. This is seen by the fact that models that allow for fat-tailed error distributions provide the best forecasts across all variables. We note that the case for inflation is less clear with the TVP-VAR-SV and TVP-VAR-SVt models producing similar forecast results. This being said, when considering inflation forecastability the CVAR-SVt model clearly outperforms the CVAR-SV model. Also consistent with the univariate case, we note that relatively poor performance of regime switching models as compared to models with SV.

Finally, Table 3.3 presents the results for the best point forecasting model for each variable. These models represent our preferred models if we were to produce a best guess of a future interest, GDP or inflation rate. Interestingly, we find that the univariate AR-SVt and TVP-AR-t models produce the best interest rate and GDP forecasts respectively. Conversely, the multivariate TVP-VAR-t model provides the most accurate inflation forecasts. In summary the point forecast results suggest that modelling of both time variation and fat-tails using both univariate and multivariate models is important in the modelling of Australian CPI, GDP and interest rates.

3.4.2 Density Forecast Results

Unlike point forecasts which produce a single best guess estimate of the future, density forecasts are able to account for uncertainty by providing a range for possible future values of of GDP, inflation and interest rates. The results in Tables 3.3 and 3.4 indicate that accounting for time variation in the model parameters and stochastic volatility along with fat-tails enhances the forecast accuracy of all variables relative to a constant, Gaussian benchmark specification. In each case the unanimity of model selection is quite remarkable. For instance, when viewing the multivariate modelling results the TVP-VAR-SVt produces the best interest and inflation forecasts at all forecast horizons, whilst the CVAR-SV model dominates the GDP forecasts. Interestingly, when viewing the univariate results, whilst the TVP-AR-SVt model provides the best inter-
est rate forecasts, the simple rolling window AR-t model is shown to produce the most accurate GDP and inflation forecasts. Finally, in contrast to Sims and Zha (2006) who find that regime switching models have good forecasting properties when considering US macroeconomic variables, we find that when considering Australian variables such models fail to outperform those with fat-tails and stochastic volatility.

In similar fashion with the point forecast results Table 3.6 presents a comparison of the best density forecasting model for each variable. Since density forecasts encompass a wider range of possible outcomes as compared to point estimates, the best density forecasting model represents our preferred modelling choice if we were to produce a probabilistic based best guess of future interest, GDP growth or inflation rates. The results show that the TVP-VAR-SVt model provides the best interest and inflation forecasts, whilst the simple rolling window AR-t model provides the most accurate GDP forecasts. This suggests that whilst time variation and information from other macroeconomic variables play a key role in interest rate and inflation decisions, these features play less of a role in accurately predicting in real GDP growth. Nonetheless, if we had to choose a “best” model to forecast Australian macroeconomic variables, then the results suggest that the TVP-VAR-SVt would be the correct choice.

3.5 Intertemporal Forecast Results of Autoregressive Models

As mentioned earlier, a growing body of literature has revealed that forecast performance is often not stable over time (see, e.g. Stock and Watson (2007, 2010), Chan et al (2012), Chan (2013), D’Agostino et al (2013), Clark and Ravazzolo (2014), Chan (2015) or Chiu et al (2015)). With this literature in mind, we investigate the intertemporal forecast performance of time varying AR and VAR models under both Gaussian and Student's-t distributions over time by plotting the cumulative sums of log predictive likelihoods. Since they do not provide the best forecasts of any variables over any time horizons we exclude the Markov switching models from this intertemporal analysis. The univariate and multivariate inflation results are in Figures 3.6 and 3.7, the output results are in Figures 3.8 and 3.9 and the interest rate results are in Figures 3.10 and
Overall, from a holistic macroeconomic modelling perspective, it can be seen that the TVP-VAR-SVt and the simple rolling window AR-t models respectively provide the most accurate multivariate and univariate forecasts. More generally, a few patterns in the forecast performance of all series are worth discussing. First, when comparing Gaussian and fat-tail models, with but one exception in the interest rate forecasts, the fat-tail models produce superior forecasts across all variables. This shows that models with fat-tails produce better forecasts as compared to their Gaussian counterparts. Next, when comparing models with and without stochastic volatility, the models with stochastic volatility produce superior forecasts across all variables. This shows that models with stochastic volatility produce better forecasts as compared to their fixed counterparts. Finally, when comparing models with and without time varying parameters, the TVP-AR and TVP-VAR models consistently produce superior forecasts across all variables. This shows that models with stochastic volatility and fat-tails produce better forecasts as compared to their fixed Gaussian counterparts.

It is also worth discussing some interesting features of the forecast performance of individual variables. First, when looking at the inflation results, it’s noticeable that before the year 2000 the Gaussian and fat-tail models produce similar forecasts. After 2000 there is a divergence with fat-tail models clearly outperforming the Gaussian counterparts. This break is likely due to the introduction of the goods and services tax (GST). A different pattern emerges in the RGDP forecasts results. Specifically, rather than a divergence in forecast performance following 2000 there is almost no evidence of a break with difference between the fat-tailed model and the Gaussian model remaining relatively consistent over the majority of the sample period. A noticeable break does occur in 2006 however, when comparing the multivariate TVP-VAR-SVt and TVP-VAR-SV models. A similar result is found in the multivariate interest rate forecast results in which accounting for fat-tails improves the forecastability of interest rates after the 2007/08 GFC period.
3.6 Conclusion

We assess whether modelling structural change and fat-tailed events can improve the forecast accuracy of key Australian macroeconomic variables: real GDP growth, CPI inflation and a short-term interest rate taken to be the 90 day Bank Accepted Bills/Negotiable Certificates of Deposit. Methodologically, we incorporate time variation and fat-tails into traditionally Gaussian, fixed coefficients multivariate and univariate autoregressive models. The class of univariate autoregressive (AR) and multivariate vector autoregressive (VAR) models allow for time variation via two sources: (1) in the models coefficients, (2) in the variance of the shocks. For the multivariate models we consider a third source of time variation via the covariance terms. In addition to accounting for time variation within the coefficients and volatilities, all models are estimated under both Gaussian and Student-t error distributions. Adding fat-tails to various models allows increases the likelihood of extreme events and may lead to faster adaptation to expansions and/or recessions. For completeness, we also consider the forecast performance of non-linear regime switching as well as rolling-window ARs and VARs.

The results yield four important findings. First, fat-tailed models consistently outperform their Gaussian counterparts. Second, time varying parameters and stochastic volatility improves forecast performance across all variables relative to a constant parameter benchmark. Third, stochastic volatility models under a Student’s-t distribution are found to generate more accurate density forecasts as compared to the same models under a Gaussian specification. Taken together these results suggest that both structural instabilities and fat-tail events are important features in modelling Australian macroeconomic variables. Finally, when comparing the forecast accuracy of univariate and multivariate models the simple rolling window autoregression with fat-tails produces the most accurate output growth forecasts, whilst the time varying parameter vector autoregression with stochastic volatility and fat-tails produces the best interest and inflation forecasts. Nonetheless, from a holistic macroeconomic modelling perspective, the vector autoregression with the proposed modelling features provides important information for central bankers policy decisions.
We note that we have only provided an out of sample study of the proposed modelling features. For future research it would be useful analyze the in sample fit by incorporating structural instabilities and fat-tails into general equilibrium models of the Australian economy. For instance, the New Keynesian model of Australia developed by Jääskelä and Nimark (2011) could be extended by allowing for time varying Student’s-t distributed disturbances within both aggregate demand and supply shocks.
3.7 Appendix

3.7.1 Appendix A Tables and Charts

Figure 3.1: Australian Macroeconomic Time Series
Table 3.1: Full sample univariate point forecast for interest, GDP growth, and inflation: relative mean square forecast errors (RMSFE) subject to an AR benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>AR</th>
<th>ARt</th>
<th>AR-SV</th>
<th>AR-SVt</th>
<th>TVP-AR</th>
<th>TVP-ARt</th>
<th>TVP-AR-SV</th>
<th>TVP-AR-SVt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>i</td>
<td>1.000</td>
<td>0.742</td>
<td>0.755</td>
<td>0.717</td>
<td>0.915</td>
<td>0.843</td>
<td>0.699</td>
<td>0.755</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.000</td>
<td>1.000</td>
<td>1.003</td>
<td>1.005</td>
<td>0.991</td>
<td>0.995</td>
<td>1.007</td>
<td>1.010</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>1.000</td>
<td>0.953</td>
<td>0.948</td>
<td>0.952</td>
<td>0.933</td>
<td>0.900</td>
<td>0.906</td>
<td>0.904</td>
</tr>
<tr>
<td>4</td>
<td>i</td>
<td>1.000</td>
<td>0.810</td>
<td>0.792</td>
<td>0.762</td>
<td>0.791</td>
<td>0.783</td>
<td>0.749</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.000</td>
<td>1.001</td>
<td>0.998</td>
<td>0.999</td>
<td>0.984</td>
<td>0.975</td>
<td>0.982</td>
<td>0.981</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>1.000</td>
<td>0.881</td>
<td>0.857</td>
<td>0.858</td>
<td>0.798</td>
<td>0.741</td>
<td>0.749</td>
<td>0.736</td>
</tr>
<tr>
<td>8</td>
<td>i</td>
<td>1.000</td>
<td>0.760</td>
<td>0.697</td>
<td>0.675</td>
<td>0.660</td>
<td>0.646</td>
<td>0.686</td>
<td>0.669</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.000</td>
<td>1.001</td>
<td>1.001</td>
<td>1.002</td>
<td>0.985</td>
<td>0.979</td>
<td>0.981</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>1.000</td>
<td>0.844</td>
<td>0.807</td>
<td>0.803</td>
<td>0.711</td>
<td>0.650</td>
<td>0.654</td>
<td>0.640</td>
</tr>
<tr>
<td>12</td>
<td>i</td>
<td>1.000</td>
<td>0.712</td>
<td>0.572</td>
<td>0.579</td>
<td>0.602</td>
<td>0.552</td>
<td>0.647</td>
<td>0.593</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.000</td>
<td>1.001</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>0.981</td>
<td>0.987</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>1.000</td>
<td>0.823</td>
<td>0.790</td>
<td>0.783</td>
<td>0.689</td>
<td>0.607</td>
<td>0.619</td>
<td>0.606</td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>1.000</td>
<td>0.756</td>
<td>0.704</td>
<td>0.683</td>
<td>0.742</td>
<td>0.706</td>
<td>0.686</td>
<td>0.697</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.000</td>
<td>1.001</td>
<td>1.000</td>
<td>1.001</td>
<td>0.990</td>
<td>0.983</td>
<td>0.989</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>1.000</td>
<td>0.875</td>
<td>0.851</td>
<td>0.849</td>
<td>0.783</td>
<td>0.724</td>
<td>0.732</td>
<td>0.721</td>
</tr>
<tr>
<td>Forecast Horizon</td>
<td>Variable</td>
<td>AR (Roll)</td>
<td>AR-t (Roll)</td>
<td>JSRS(2)-AR</td>
<td>JSRS(3)-AR</td>
<td>JSRS(4)-AR</td>
<td>ISRS(2)-AR</td>
<td>ISRS(3)-AR</td>
<td>ISRS(4)-AR</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>i</td>
<td>0.769</td>
<td>0.752</td>
<td>1.424</td>
<td>5.466</td>
<td>5.968</td>
<td>4.617</td>
<td>4.204</td>
<td>3.829</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.053</td>
<td>1.056</td>
<td>1.642</td>
<td>1.511</td>
<td>1.556</td>
<td>1.589</td>
<td>1.471</td>
<td>1.323</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.945</td>
<td>0.907</td>
<td>1.199</td>
<td>1.048</td>
<td>1.067</td>
<td>0.970</td>
<td>0.897</td>
<td>0.853</td>
</tr>
<tr>
<td>4</td>
<td>i</td>
<td>0.908</td>
<td>0.903</td>
<td>1.016</td>
<td>2.062</td>
<td>2.180</td>
<td>1.924</td>
<td>1.785</td>
<td>1.682</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.006</td>
<td>1.011</td>
<td>1.702</td>
<td>1.611</td>
<td>1.623</td>
<td>1.688</td>
<td>1.612</td>
<td>1.468</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.751</td>
<td>0.778</td>
<td>1.010</td>
<td>0.896</td>
<td>0.895</td>
<td>0.874</td>
<td>0.778</td>
<td>0.755</td>
</tr>
<tr>
<td>8</td>
<td>i</td>
<td>0.780</td>
<td>0.776</td>
<td>0.948</td>
<td>1.562</td>
<td>1.611</td>
<td>1.507</td>
<td>1.366</td>
<td>1.281</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.002</td>
<td>1.003</td>
<td>1.707</td>
<td>1.633</td>
<td>1.635</td>
<td>1.701</td>
<td>1.638</td>
<td>1.515</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.665</td>
<td>0.707</td>
<td>0.858</td>
<td>0.793</td>
<td>0.783</td>
<td>0.776</td>
<td>0.656</td>
<td>0.687</td>
</tr>
<tr>
<td>12</td>
<td>i</td>
<td>0.586</td>
<td>0.610</td>
<td>0.885</td>
<td>1.398</td>
<td>1.425</td>
<td>1.357</td>
<td>1.203</td>
<td>1.111</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.002</td>
<td>0.996</td>
<td>1.710</td>
<td>1.634</td>
<td>1.643</td>
<td>1.702</td>
<td>1.632</td>
<td>1.486</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.678</td>
<td>0.676</td>
<td>0.813</td>
<td>0.755</td>
<td>0.741</td>
<td>0.743</td>
<td>1.015</td>
<td>5.431</td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>0.761</td>
<td>0.760</td>
<td>1.068</td>
<td>2.622</td>
<td>2.796</td>
<td>2.351</td>
<td>2.140</td>
<td>1.976</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>1.016</td>
<td>1.017</td>
<td>1.690</td>
<td>1.597</td>
<td>1.614</td>
<td>1.670</td>
<td>1.588</td>
<td>1.448</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.760</td>
<td>0.767</td>
<td>0.970</td>
<td>0.873</td>
<td>0.872</td>
<td>0.841</td>
<td>0.836</td>
<td>1.942</td>
</tr>
</tbody>
</table>
Table 3.2: Full sample multivariate point forecast for interest, GDP growth, and inflation, relative mean square forecast errors (RMSFE) subject to a VAR benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>CVAR</th>
<th>CVARt</th>
<th>CVAR-SV</th>
<th>CVAR-SVt</th>
<th>TVP-VAR</th>
<th>TVP-VARt</th>
<th>TVP-VAR-SV</th>
<th>TVP-VAR-SVt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( i )</td>
<td>1.000</td>
<td>0.867</td>
<td>0.880</td>
<td>0.851</td>
<td>1.148</td>
<td>1.067</td>
<td>0.963</td>
<td>1.003</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.000</td>
<td>0.974</td>
<td>0.999</td>
<td>0.999</td>
<td>0.995</td>
<td>0.982</td>
<td>1.008</td>
<td>1.006</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.000</td>
<td>0.957</td>
<td>0.940</td>
<td>0.948</td>
<td>0.891</td>
<td>0.894</td>
<td>0.884</td>
<td>0.892</td>
</tr>
<tr>
<td>4</td>
<td>( i )</td>
<td>1.000</td>
<td>0.927</td>
<td>0.944</td>
<td>0.925</td>
<td>1.129</td>
<td>1.045</td>
<td>1.036</td>
<td>1.050</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.000</td>
<td>0.991</td>
<td>0.993</td>
<td>0.989</td>
<td>0.991</td>
<td>0.989</td>
<td>0.993</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.000</td>
<td>0.868</td>
<td>0.860</td>
<td>0.849</td>
<td>0.781</td>
<td>0.766</td>
<td>0.784</td>
<td>0.780</td>
</tr>
<tr>
<td>8</td>
<td>( i )</td>
<td>1.000</td>
<td>0.818</td>
<td>0.830</td>
<td>0.798</td>
<td>1.011</td>
<td>0.834</td>
<td>0.857</td>
<td>0.840</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.000</td>
<td>0.997</td>
<td>1.001</td>
<td>0.997</td>
<td>0.996</td>
<td>0.995</td>
<td>1.015</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.000</td>
<td>0.796</td>
<td>0.794</td>
<td>0.775</td>
<td>0.694</td>
<td>0.686</td>
<td>0.695</td>
<td>0.689</td>
</tr>
<tr>
<td>12</td>
<td>( i )</td>
<td>1.000</td>
<td>0.742</td>
<td>0.722</td>
<td>0.692</td>
<td>1.008</td>
<td>0.663</td>
<td>0.674</td>
<td>0.649</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.000</td>
<td>0.955</td>
<td>1.009</td>
<td>0.998</td>
<td>0.998</td>
<td>0.989</td>
<td>1.002</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.000</td>
<td>0.759</td>
<td>0.762</td>
<td>0.731</td>
<td>0.655</td>
<td>0.628</td>
<td>0.632</td>
<td>0.633</td>
</tr>
<tr>
<td>Average</td>
<td>( i )</td>
<td>1.000</td>
<td>0.839</td>
<td>0.844</td>
<td>\textbf{0.817}</td>
<td>1.074</td>
<td>0.902</td>
<td>0.883</td>
<td>0.886</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.000</td>
<td>0.9893</td>
<td>1.000</td>
<td>0.996</td>
<td>0.995</td>
<td>\textbf{0.9886}</td>
<td>1.004</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.000</td>
<td>0.845</td>
<td>0.839</td>
<td>0.826</td>
<td>0.755</td>
<td>\textbf{0.743}</td>
<td>0.749</td>
<td>0.749</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>VAR (Roll)</th>
<th>VAR-t (Roll)</th>
<th>JSRS(2)-VAR</th>
<th>JSRS(3)-VAR</th>
<th>JSRS(4)-VAR</th>
<th>ISRS(2)-VAR</th>
<th>ISRS(3)-VAR</th>
<th>ISRS(4)-VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( i )</td>
<td>0.985</td>
<td>0.978</td>
<td>8.370</td>
<td>7.397</td>
<td>8.114</td>
<td>3.820</td>
<td>4.827</td>
<td>5.059</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.173</td>
<td>1.163</td>
<td>1.615</td>
<td>1.547</td>
<td>1.557</td>
<td>1.552</td>
<td>1.454</td>
<td>1.420</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.910</td>
<td>0.880</td>
<td>1.145</td>
<td>1.097</td>
<td>1.169</td>
<td>0.932</td>
<td>0.975</td>
<td>1.009</td>
</tr>
<tr>
<td>4</td>
<td>( i )</td>
<td>1.140</td>
<td>1.250</td>
<td>3.165</td>
<td>2.938</td>
<td>3.102</td>
<td>2.143</td>
<td>2.390</td>
<td>2.400</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.002</td>
<td>0.992</td>
<td>1.677</td>
<td>1.621</td>
<td>1.639</td>
<td>1.603</td>
<td>1.551</td>
<td>1.526</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.802</td>
<td>0.804</td>
<td>1.032</td>
<td>0.962</td>
<td>1.012</td>
<td>0.838</td>
<td>0.873</td>
<td>0.893</td>
</tr>
<tr>
<td>8</td>
<td>( i )</td>
<td>1.116</td>
<td>1.349</td>
<td>2.183</td>
<td>2.111</td>
<td>2.166</td>
<td>1.677</td>
<td>1.745</td>
<td>1.739</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>0.987</td>
<td>1.040</td>
<td>1.704</td>
<td>1.649</td>
<td>1.669</td>
<td>1.636</td>
<td>1.594</td>
<td>1.579</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.742</td>
<td>0.752</td>
<td>0.925</td>
<td>0.870</td>
<td>0.896</td>
<td>0.770</td>
<td>0.787</td>
<td>0.793</td>
</tr>
<tr>
<td>12</td>
<td>( i )</td>
<td>2.413</td>
<td>2.324</td>
<td>1.780</td>
<td>1.761</td>
<td>1.784</td>
<td>1.459</td>
<td>1.458</td>
<td>1.431</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.044</td>
<td>1.356</td>
<td>1.686</td>
<td>1.637</td>
<td>1.652</td>
<td>1.631</td>
<td>1.576</td>
<td>1.553</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>1.220</td>
<td>0.858</td>
<td>0.882</td>
<td>0.848</td>
<td>0.864</td>
<td>0.754</td>
<td>0.756</td>
<td>0.760</td>
</tr>
<tr>
<td>Average</td>
<td>( i )</td>
<td>1.413</td>
<td>1.475</td>
<td>3.874</td>
<td>3.552</td>
<td>3.792</td>
<td>2.275</td>
<td>2.595</td>
<td>2.657</td>
</tr>
<tr>
<td></td>
<td>( y )</td>
<td>1.052</td>
<td>1.138</td>
<td>1.671</td>
<td>1.613</td>
<td>1.629</td>
<td>1.605</td>
<td>1.544</td>
<td>1.520</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.918</td>
<td>0.823</td>
<td>1.001</td>
<td>0.944</td>
<td>0.985</td>
<td>0.823</td>
<td>0.848</td>
<td>0.864</td>
</tr>
</tbody>
</table>
Table 3.3: Best point forecast model for interest, $i$, GDP growth, $y$, and inflation, $\pi$, over the full sample: relative mean square forecast errors (RMSFE) with multivariate model as benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>CVAR-SVt</th>
<th>AR-SVt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$i$</td>
<td>1.00</td>
<td>0.993</td>
</tr>
<tr>
<td>4</td>
<td>$i$</td>
<td>1.00</td>
<td>0.999</td>
</tr>
<tr>
<td>8</td>
<td>$i$</td>
<td>1.00</td>
<td>0.986</td>
</tr>
<tr>
<td>12</td>
<td>$i$</td>
<td>1.00</td>
<td>0.931</td>
</tr>
<tr>
<td>Average</td>
<td>$i$</td>
<td>1.00</td>
<td>0.977</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>TVP-VAR-t</th>
<th>TVP-ARt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$y$</td>
<td>1.00</td>
<td>0.996</td>
</tr>
<tr>
<td>4</td>
<td>$y$</td>
<td>1.00</td>
<td>0.987</td>
</tr>
<tr>
<td>8</td>
<td>$y$</td>
<td>1.00</td>
<td>0.991</td>
</tr>
<tr>
<td>12</td>
<td>$y$</td>
<td>1.00</td>
<td>0.986</td>
</tr>
<tr>
<td>Average</td>
<td>$y$</td>
<td>1.00</td>
<td>0.990</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>TVP-VAR-t</th>
<th>TVP-AR-SVt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\pi$</td>
<td>1.00</td>
<td>1.081</td>
</tr>
<tr>
<td>4</td>
<td>$\pi$</td>
<td>1.00</td>
<td>1.050</td>
</tr>
<tr>
<td>8</td>
<td>$\pi$</td>
<td>1.00</td>
<td>1.032</td>
</tr>
<tr>
<td>12</td>
<td>$\pi$</td>
<td>1.00</td>
<td>1.070</td>
</tr>
<tr>
<td>Average</td>
<td>$\pi$</td>
<td>1.00</td>
<td>1.058</td>
</tr>
</tbody>
</table>
Table 3.4: Full sample univariate density forecast for interest, GDP growth, and inflation, subject to an AR benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>AR</th>
<th>ARt</th>
<th>AR-SV</th>
<th>AR-Svt</th>
<th>TVP-AR</th>
<th>TVP-ARt</th>
<th>TVP-AR-SV</th>
<th>TVP-AR-Svt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.000</td>
<td>0.667</td>
<td>1.058</td>
<td>1.115</td>
<td>0.100</td>
<td>0.833</td>
<td>1.128</td>
<td>1.184</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.000</td>
<td>0.118</td>
<td>0.296</td>
<td>0.314</td>
<td>0.019</td>
<td>0.144</td>
<td>0.288</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.319</td>
<td>0.228</td>
<td>0.350</td>
<td>0.038</td>
<td>0.306</td>
<td>0.291</td>
<td>0.418</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.000</td>
<td>0.172</td>
<td>0.326</td>
<td>0.433</td>
<td>0.263</td>
<td>0.485</td>
<td>0.287</td>
<td>0.407</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.000</td>
<td>0.089</td>
<td>0.315</td>
<td>0.331</td>
<td>0.010</td>
<td>0.104</td>
<td>0.303</td>
<td>0.319</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.000</td>
<td>0.187</td>
<td>0.251</td>
<td>0.297</td>
<td>0.148</td>
<td>0.426</td>
<td>0.410</td>
<td>0.497</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.092</td>
<td>0.247</td>
<td>0.316</td>
<td>0.465</td>
<td>0.609</td>
<td>0.556</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.000</td>
<td>0.155</td>
<td>0.270</td>
<td>0.293</td>
<td>0.228</td>
<td>0.509</td>
<td>0.503</td>
<td>0.575</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.000</td>
<td>0.095</td>
<td>0.390</td>
<td>0.405</td>
<td>0.578</td>
<td>0.765</td>
<td>0.754</td>
<td>0.776</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.000</td>
<td>0.131</td>
<td>0.334</td>
<td>0.353</td>
<td>0.000</td>
<td>0.101</td>
<td>0.311</td>
<td>0.330</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.161</td>
<td>0.270</td>
<td>0.293</td>
<td>0.257</td>
<td>0.555</td>
<td>0.541</td>
<td>0.627</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.000</td>
<td>0.270</td>
<td>0.505</td>
<td>0.567</td>
<td>0.352</td>
<td>0.681</td>
<td>0.733</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>0.000</td>
<td>0.098</td>
<td>0.319</td>
<td>0.336</td>
<td>0.009</td>
<td>0.11</td>
<td>0.304</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.206</td>
<td>0.255</td>
<td>0.308</td>
<td>0.168</td>
<td>0.47</td>
<td>0.436</td>
<td>0.529</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.000</td>
<td>0.206</td>
<td>0.255</td>
<td>0.308</td>
<td>0.168</td>
<td>0.47</td>
<td>0.436</td>
<td>0.529</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.876</td>
<td>0.667</td>
<td>-0.574</td>
<td>-0.950</td>
<td>-1.050</td>
<td>-0.541</td>
<td>-0.012</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.324</td>
<td>0.362</td>
<td>-8.946</td>
<td>-2.942</td>
<td>-0.432</td>
<td>-0.155</td>
<td>-0.221</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.245</td>
<td>0.319</td>
<td>-3.553</td>
<td>0.034</td>
<td>-0.001</td>
<td>0.124</td>
<td>0.354</td>
<td>0.414</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.251</td>
<td>0.172</td>
<td>-5.021</td>
<td>-0.672</td>
<td>-0.513</td>
<td>-0.193</td>
<td>-0.088</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.363</td>
<td>0.429</td>
<td>-6.471</td>
<td>-0.322</td>
<td>-0.434</td>
<td>-0.181</td>
<td>-0.098</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.400</td>
<td>0.187</td>
<td>-6.178</td>
<td>0.100</td>
<td>0.071</td>
<td>0.165</td>
<td>0.393</td>
<td>0.464</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.172</td>
<td>0.111</td>
<td>-8.867</td>
<td>-0.465</td>
<td>-0.514</td>
<td>-0.452</td>
<td>-0.226</td>
<td>-0.085</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.370</td>
<td>0.451</td>
<td>-8.941</td>
<td>-0.248</td>
<td>-0.367</td>
<td>-0.165</td>
<td>-0.090</td>
<td>0.023</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.502</td>
<td>0.155</td>
<td>-3.415</td>
<td>0.204</td>
<td>0.187</td>
<td>0.229</td>
<td>0.466</td>
<td>0.543</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>0.314</td>
<td>0.131</td>
<td>-16.291</td>
<td>-0.376</td>
<td>-0.413</td>
<td>-0.408</td>
<td>-0.108</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.381</td>
<td>0.465</td>
<td>-6.463</td>
<td>-0.253</td>
<td>-0.436</td>
<td>-0.151</td>
<td>-0.065</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.551</td>
<td>0.161</td>
<td>-7.485</td>
<td>0.206</td>
<td>0.177</td>
<td>0.247</td>
<td>0.453</td>
<td>0.507</td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>0.403</td>
<td>0.270</td>
<td>-7.688</td>
<td>-0.600</td>
<td>-0.657</td>
<td>-0.479</td>
<td>-0.135</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.360</td>
<td>0.427</td>
<td>-7.705</td>
<td>-0.279</td>
<td>-0.430</td>
<td>-0.162</td>
<td>-0.067</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>π</td>
<td>0.425</td>
<td>0.206</td>
<td>-5.158</td>
<td>0.136</td>
<td>0.109</td>
<td>0.191</td>
<td>0.417</td>
<td>0.482</td>
</tr>
</tbody>
</table>
Table 3.5: Full sample multivariate density forecast for interest, i, GDP growth, y, and inflation, \( \pi \): relative log predictive likelihood (RLPL) subject to a VAR benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>CVAR</th>
<th>CVARt</th>
<th>CVAR-SV</th>
<th>CVAR-SVt</th>
<th>TVP-VAR</th>
<th>TVP-VARt</th>
<th>TVP-VAR-SV</th>
<th>TVP-VAR-SVt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>i</td>
<td>0.000</td>
<td>0.501</td>
<td>0.988</td>
<td>1.059</td>
<td>0.347</td>
<td>0.557</td>
<td>0.989</td>
<td>1.061</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.136</td>
<td>0.307</td>
<td>0.318</td>
<td>0.015</td>
<td>0.135</td>
<td>0.233</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.000</td>
<td>0.349</td>
<td>0.297</td>
<td>0.407</td>
<td>0.088</td>
<td>0.401</td>
<td>0.341</td>
<td>0.466</td>
</tr>
<tr>
<td>4</td>
<td>i</td>
<td>0.000</td>
<td>0.077</td>
<td>0.274</td>
<td>0.376</td>
<td>0.141</td>
<td>0.264</td>
<td>0.118</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.097</td>
<td>0.336</td>
<td>0.347</td>
<td>0.012</td>
<td>0.102</td>
<td>0.333</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.000</td>
<td>0.219</td>
<td>0.317</td>
<td>0.389</td>
<td>0.168</td>
<td>0.388</td>
<td>0.405</td>
<td>0.508</td>
</tr>
<tr>
<td>8</td>
<td>i</td>
<td>0.000</td>
<td>0.034</td>
<td>0.256</td>
<td>0.315</td>
<td>0.202</td>
<td>0.386</td>
<td>0.426</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.094</td>
<td>0.351</td>
<td>0.363</td>
<td>0.002</td>
<td>0.097</td>
<td>0.332</td>
<td>0.357</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.000</td>
<td>0.194</td>
<td>0.359</td>
<td>0.420</td>
<td>0.222</td>
<td>0.438</td>
<td>0.484</td>
<td>0.572</td>
</tr>
<tr>
<td>12</td>
<td>i</td>
<td>0.000</td>
<td>0.060</td>
<td>0.344</td>
<td>0.364</td>
<td>0.288</td>
<td>0.512</td>
<td>0.629</td>
<td>0.645</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.000</td>
<td>0.092</td>
<td>0.345</td>
<td>0.360</td>
<td>-0.011</td>
<td>0.093</td>
<td>0.334</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.000</td>
<td>0.197</td>
<td>0.394</td>
<td>0.458</td>
<td>0.237</td>
<td>0.473</td>
<td>0.541</td>
<td>0.625</td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>0.000</td>
<td>0.17</td>
<td>0.466</td>
<td>0.529</td>
<td>0.245</td>
<td>0.43</td>
<td>0.541</td>
<td>0.635</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.000</td>
<td>0.10</td>
<td>0.335</td>
<td>0.347</td>
<td>0.004</td>
<td>0.11</td>
<td>0.313</td>
<td>0.342</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>VAR (Roll)</th>
<th>VAR-t (Roll)</th>
<th>JSRS(2)-VAR</th>
<th>JSRS(3)-VAR</th>
<th>JSRS(4)-VAR</th>
<th>ISRS(2)-VAR</th>
<th>ISRS(3)-VAR</th>
<th>ISRS(4)-VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>i</td>
<td>0.553</td>
<td>0.684</td>
<td>-1.304</td>
<td>-1.112</td>
<td>-1.212</td>
<td>0.027</td>
<td>-0.246</td>
<td>-0.256</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.247</td>
<td>0.261</td>
<td>-0.306</td>
<td>-0.185</td>
<td>-0.176</td>
<td>-0.109</td>
<td>-0.011</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.243</td>
<td>0.427</td>
<td>-0.133</td>
<td>0.033</td>
<td>0.025</td>
<td>0.316</td>
<td>0.318</td>
<td>0.322</td>
</tr>
<tr>
<td>4</td>
<td>i</td>
<td>0.171</td>
<td>0.203</td>
<td>-0.901</td>
<td>-0.774</td>
<td>-0.852</td>
<td>-0.404</td>
<td>-0.429</td>
<td>-0.352</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.279</td>
<td>0.306</td>
<td>-0.346</td>
<td>-0.214</td>
<td>-0.224</td>
<td>-0.090</td>
<td>-0.075</td>
<td>-0.087</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.294</td>
<td>0.434</td>
<td>-0.222</td>
<td>0.077</td>
<td>0.069</td>
<td>0.401</td>
<td>0.332</td>
<td>0.346</td>
</tr>
<tr>
<td>8</td>
<td>i</td>
<td>0.137</td>
<td>0.120</td>
<td>-0.662</td>
<td>-0.388</td>
<td>-0.067</td>
<td>-0.000</td>
<td>-0.409</td>
<td>-0.393</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.255</td>
<td>0.286</td>
<td>-0.365</td>
<td>-0.219</td>
<td>-0.224</td>
<td>-0.107</td>
<td>-0.097</td>
<td>-0.068</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.298</td>
<td>0.431</td>
<td>-0.178</td>
<td>0.139</td>
<td>0.137</td>
<td>0.435</td>
<td>0.385</td>
<td>0.383</td>
</tr>
<tr>
<td>12</td>
<td>i</td>
<td>0.199</td>
<td>0.183</td>
<td>-0.487</td>
<td>-0.446</td>
<td>-0.473</td>
<td>-0.590</td>
<td>-0.408</td>
<td>-0.305</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.230</td>
<td>0.260</td>
<td>-0.208</td>
<td>-0.130</td>
<td>-0.124</td>
<td>-0.046</td>
<td>-0.034</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.309</td>
<td>0.461</td>
<td>-0.086</td>
<td>0.191</td>
<td>0.206</td>
<td>0.460</td>
<td>0.403</td>
<td>0.405</td>
</tr>
<tr>
<td>Average</td>
<td>i</td>
<td>0.265</td>
<td>0.298</td>
<td>-0.809</td>
<td>-0.730</td>
<td>-0.793</td>
<td>-0.392</td>
<td>-0.388</td>
<td>-0.334</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>0.233</td>
<td>0.278</td>
<td>-0.306</td>
<td>-0.187</td>
<td>-0.187</td>
<td>-0.088</td>
<td>-0.064</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>( \pi )</td>
<td>0.286</td>
<td>0.438</td>
<td>-0.155</td>
<td>0.110</td>
<td>0.109</td>
<td>0.403</td>
<td>0.365</td>
<td>0.364</td>
</tr>
</tbody>
</table>
Table 3.6: Best density forecast model for interest, GDP growth, and inflation, over the full sample: relative log predictive likelihood (RLPL) with multivariate model as benchmark.

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>TVP-VAR-SVT</th>
<th>TVP-AR-SVT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( i )</td>
<td>0.00</td>
<td>-1.855</td>
</tr>
<tr>
<td>4</td>
<td>( i )</td>
<td>0.00</td>
<td>-0.393</td>
</tr>
<tr>
<td>8</td>
<td>( i )</td>
<td>0.00</td>
<td>-0.364</td>
</tr>
<tr>
<td>12</td>
<td>( i )</td>
<td>0.00</td>
<td>-0.371</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.00</td>
<td>-0.796</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>TVP-VAR-SVT</th>
<th>VAR-SVT</th>
<th>AR-( t ) (roll)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( y )</td>
<td>0.00</td>
<td>0.051</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( y )</td>
<td>0.00</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>( y )</td>
<td>0.00</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>( y )</td>
<td>0.00</td>
<td>0.127</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.00</td>
<td>0.096</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Variable</th>
<th>TVP-VAR-SVT</th>
<th>TVP-AR-SVT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \pi )</td>
<td>0.00</td>
<td>-0.055</td>
</tr>
<tr>
<td>4</td>
<td>( \pi )</td>
<td>0.00</td>
<td>-0.025</td>
</tr>
<tr>
<td>8</td>
<td>( \pi )</td>
<td>0.00</td>
<td>-0.018</td>
</tr>
<tr>
<td>12</td>
<td>( \pi )</td>
<td>0.00</td>
<td>-0.017</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.00</td>
<td>-0.029</td>
</tr>
</tbody>
</table>
Figure 3.2: Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for CPI Inflation

Figure 3.3: Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for real GDP growth
Figure 3.4: Posterior mean of the degree of freedom parameter estimation for univariate and multivariate models for the interest rate.

Figure 3.5: Aggregate posterior mean of the degree of freedom parameter for the interest rate, real GDP growth and the inflation rate.
Figure 3.6: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; CPI inflation.

Figure 3.7: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; CPI inflation.
Figure 3.8: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; RGDP Growth.

Figure 3.9: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; RGDP Growth.
Figure 3.10: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the autoregressive (AR) model; Interest.

Figure 3.11: Cumulative sums of log predictive likelihoods for one-quarter-ahead forecasts relative to the vector autoregressive (VAR) model; Interest.
3.7.2 Appendix B - Regime Switching VAR

Posterior draws for the most complicated Markov switching VAR which allows for independent breaks in the VAR coefficients and error covariances model are obtained through a four block Gibbs sampler that cycles though:

1. $p(S_t|y, \beta_{S_t}, \Omega_{S_t}, p)$
2. $p(\beta_{S_t}|y, S_t, \Omega_{S_t}, p)$
3. $p(\Omega_{S_t}|y, S_t, \beta_{S_t}, p)$
4. $p(p|y, S_t, \beta_{S_t}, \Omega_{S_t})$

where $p$ is a vector of transition probabilities. Consistent with Barnett et al (2014), we set $\beta_{S_t} \sim N(0, 4I_n)$, and $\Omega_{S_t} \sim IW(n+3, I_n)$. Whilst sampling from block 1 requires the use of a standard two-pass procedure, sampling blocks 2 and 3 can be efficiently completed via precision sampler techniques developed by Chan and Jeliazkov (2009).

We now describe how to sample each state in turn:

1. To draw from $p(S_t|y, \beta_{S_t}, \Omega_{S_t}, p)$ we follow the two pass procedure set in Kim and Nelson (1999, Chapter 9). Specifically, the Markov property of the state variable implies that:

   $$f(s|\tilde{y}_T, \theta) = f(S_T|\tilde{y}_T) \prod_{t=1}^{T-1} f(S_t|S_{t+1}, \tilde{y}_t),$$

   (3.20)

   where $\tilde{y}_t = (y_t, \ldots, y_{-(k-1)})$ denote the series of observations available up to time $t$ and $\theta = (\beta_{S_t=1}, \ldots, \beta_{S_t=M}, \Omega_{S_t=1}, \ldots, \Omega_{S_t=M})$ denote the collection of parameters in each state with $s = (S_1, \ldots, S_T)$. Sampling from (19) can be done in two steps:

   (a) Calculate $f(S_T|\tilde{y}_T)$: Following Hamilton (1989), we perform a forward filter for $f(S_t|\tilde{y}_t)$ where $t = 1, \ldots, T$. Initialization is done by setting $P(S_0 = i|\tilde{y}_0, \theta)$
equal to the unconditional probability \( P(S_0 = i) \).

(b) Calculate \( f(S_t|S_{t+1}, \tilde{y}_t) \): Following Kim and Nelson (1999) simulate \( f(S_t|S_{t+1}, \tilde{y}_t) \) backward from \( t = T - 1, T - 2, \ldots, 2, 1 \) using the relationship:

\[
f(S_t|S_{t+1}, \tilde{y}_t) \propto f(S_{t+1}|S_t) f(S_t|\tilde{y}_t),
\]

where \( f(S_{t+1}|S_t) \) is the transition probability and \( f(S_t|\tilde{y}_t) \) can be derived through the Hamilton (1989) filter. Kim and Nelson (1999, p. 214) show how to sample \( S_t \) from (20).

2. To draw \( p(\beta_s|y, S_t, \Omega_s, \mathbf{p}) \) note that the likelihood function is given by:

\[
(y_t|\beta_s, \Omega_s, s) = (2\pi)^{-\frac{T}{2}} \prod_{t=1}^{T} |\Omega_s|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (Y_t - X_t \beta_s)' \Omega_s^{-1} (Y_t - X_t \beta_s) \right\}.
\]

Combining the likelihood with the prior distribution in (3) gives the conditional posterior for \( \beta_s \):

\[
(\beta_s|\tilde{y}_t, \Omega_s, s) \propto \exp \left\{ -\frac{1}{2} (Y_t - X_t \beta_s)' \Omega_s^{-1} (Y_t - X_t \beta_s) \right\} \exp \left\{ -\frac{1}{2} (\beta_s - \beta_0)' V_\beta^{-1} (\beta_s - \beta_0) \right\}
\]

(3.23)

If we assume \( S_t = i \), then (22) can be simplified to give the conditional posterior for \( \beta_{s_t=i} \):

\[
(\beta_{s_t=i}|\tilde{y}_t, \Omega_s, s) \propto \exp \left\{ -\frac{1}{2} \left[ \beta_{s_t=i}' \left( V_\beta^{-1} + \sum_{t=1}^{T} 1(S_t = i) X_t' \Omega_{s_t=i}^{-1} X_t \right) \right] \beta_{s_t=i}
+ \beta_{s_t=i}' \left( V_\beta^{-1} \beta_0 + \sum_{t=1}^{T} 1(S_t = i) X_t' \Omega_{s_t=i}^{-1} Y_t \right) \right\},
\]

(3.24)
where \( \cdot \) denotes the indicator function. Thus \( p(\beta_{S_t} | y, S_t, \Omega_{S_t}, p) \sim N(\hat{\beta}_{S_t=i}, D_{\beta_{S_t=i}}) \) and we can use the precision sampler in Chan and Jeliazkov (2009), where:

\[
\hat{\beta}_{S_t=i} = D_{\beta_{S_t=i}} \left( V^{-1}_\beta \beta_0 + \sum_{t=1}^{T} 1(S_t = i) X'_t \Omega_{S_t=i}^{-1} Y_t \right) \tag{3.25}
\]

\[
D_{\beta_{S_t=i}} = \left( V^{-1}_\beta + \sum_{t=1}^{T} 1(S_t = i) X'_t \Omega_{S_t=i}^{-1} X_t \right)^{-1} \tag{3.26}
\]

where \( y = [ y_1 \ldots y_T ]', X = \text{diag} [ X_1 X_2 \ldots X_T ], \Sigma = \text{diag} [ \Sigma_1 \Sigma_2 \ldots \Sigma_T ] \) and \( H_\beta \) is a \( Tk \times Tk \) first difference matrix.

3. Following the same steps as above it is easy to show that \( p(\Omega_{S_t=i} | y, S_t, \beta_{S_t}, p) \sim IW(\tau_{\Omega_{S_t=i}}, \Sigma_{\Omega_{S_t=i}}) \), thus sampling is as in Chan and Jeliazkov (2009), where:

\[
\tau_{\Omega_{S_t=i}} = \tau + \sum_{t=1}^{T} 1(S_t = i), \tag{3.27}
\]

\[
\Sigma_{\Omega_{S_t=i}} = \Sigma + \sum_{t=1}^{T} 1(S_t = i) (Y_t - X_t \beta_{S_t=i}) (Y_t - X_t \beta_{S_t=i})' \tag{3.28}
\]

4. Following Barnett, Mumtaz and Theodoridis (2012), we set a Dirichlet distributed prior for the transition matrix:

\[
p^0_{ij} = D(u_{ij}) \tag{3.29}
\]

where \( D(\cdot) \) is a Dirichlet distribution and \( u_{ij} = 15 \) and \( u_{ij} = 1 \) if \( i \neq j \). This choice of \( u_{ij} \) implies that the regimes are fairly persistent. It is then straightforward to show that:

\[
p_{ij} = D(u_{ij} + \eta_{ij}) \tag{3.30}
\]

where \( \eta_{ij} \) denotes the number times regime \( i \) is followed by regime \( j \).
Independent Switching Case

It is worth clarifying that the Gibbs Sampler for the independent switching case in which $\beta_{st}$ and $\Omega_{ht}$ follow distinct Markov processes is essentially the same as the joint switching case. The major distinction is that the conditional density $f(Y_t|S_t, \tilde{y}_{t-1}, \theta, H_t)$ will be different. Specifically, for generic $H_t$, if $S_t = i$ then the conditional density is:

$$f(Y_t|S_t = i, \tilde{y}_{t-1}, \theta, H_t) = (2\pi)^{-\frac{N}{2}} |\Omega_{H_t}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (Y_t - X_t\beta_{st=i})' \Omega_{H_t=j}^{-1} (Y_t - X_t\beta_{st=i}) \right\}$$

(3.31)

Similarly, for generic $S_t$, if $H_t = j$ then the conditional density is:

$$f(Y_t|S_t, \tilde{y}_{t-1}, \theta, H_t = j) = (2\pi)^{-\frac{N}{2}} |\Omega_{H_t=j}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (Y_t - X_t\beta_{st})' \Omega_{H_t=j}^{-1} (Y_t - X_t\beta_{st}) \right\} ,$$

(3.32)

Once the conditional density $f(Y_t|S_t, \tilde{y}_{t-1}, \theta, H_t)$ for both $S_t$ and $H_t$ are determined, we follow the same steps as the joint switching case. Specifically if we assume $S_t = i$ and $H_t = j$ then one can draw from $p(\beta_{st}|y, S_t, \Omega_{Ht}, \theta) \sim N(\hat{\beta}_{st=i}, D_{\beta_{st=i}})$ and $p(\Omega_{Ht=j}|y, S_t, \beta_{st}, \theta) \sim IW(\tau_{\Omega_{Ht=j}}, \Sigma_{\Omega_{Ht=j}})$ using the precision sampler where:

$$D_{\beta_{st=i}} = \left( V_{\beta}^{-1} + \sum_{t=1}^{T} 1(S_t = i) X_t' \Omega_{H_t}^{-1} X_t \right)^{-1} ,$$

(3.33)

$$\hat{\beta}_{st} = D_{\beta_{st=i}} \left( V_{\beta}^{-1} \beta_0 + \sum_{t=1}^{T} 1(S_t = i) X_t' \Omega_{H_t}^{-1} Y_t \right) ,$$

(3.34)

$$\tau_{\Omega_{st=i}} = \tau + \sum_{t=1}^{T} 1(S_t = i) ,$$

(3.35)

$$\Sigma_{\Omega_{Ht=j}} = \Sigma + \sum_{t=1}^{T} 1(H_t = j) (Y_t - X_t\beta_{st})(Y_t - X_t\beta_{st})' .$$

(3.36)

Note that we also consider a regime switching model in which switching only occurs in the parameters. In this case, the same simulation methods as in the joint switching case apply to parameters, however for the variance, the steps are simplified with the conditional posterior of $\Omega$ being standard Inverse-Wishart distribution as in the time
invariant VAR.

3.8 Appendix C - Time Varying VAR with Stochastic Volatility

Posterior draws for the most complicated TVP-VAR-SVt model are obtained through a six block Gibbs sampler that cycles through:

1. \( p(\beta|y, h, a, \Omega_\beta, \Omega_h, \Omega_a) \)
2. \( p(h|y, \beta, a, \Omega_\beta, \Omega_h, \Omega_a) \)
3. \( p(a|y, \beta, h, \Omega_\beta, \Omega_h, \Omega_a) \)
4. \( p(\Omega_\beta|y, \beta, h, a, \Omega_h, \Omega_a) \)
5. \( p(\Omega_h|y, \beta, h, a, \Omega_\beta, \Omega_a) \)
6. \( p(\Omega_a|y, \beta, h, a, \Omega_\beta, \Omega_h) \)

All other models are nested versions and can easily be formulated by setting the variance of a given block equal to zero. Following Primiceri (2005), sampling can be conducted using Kalman Filter based algorithms as in Carter and Khon (1994) and Fruhwirth-Schnatter (1994). Here we make use of an efficient estimation algorithms which exploit the fact that the precision matrices (the inverse of the variance matrices) are sparse (that is, they have few non-zero elements). Specifically, sampling from blocks 1,3,4,5 and 6 is completed via precision sampler techniques developed by Chan and Jeliazkov (2009), whilst block 2 makes use of the auxiliary mixture sampler created by Kim, Shepherd and Chib (1998) along with a sparse algorithm put forth by Chan and Hsiao (2014).

We obtain 25,000 posterior draws, discarding the first 5000 draws to allow for convergence of the Markov chain. The choice of priors and initial conditions follows the recent studies of Chan and Eisenstat (2015) and Cross (2015) which employ Bayesian estimation of TVP-VAR-SV models using Australian data. To this end, we let the initial conditions of the state equations take the following forms: \( \beta_1 \sim N(0, 10 \cdot I_k) \), \( a_1 \sim \)
\[ N(0, 10 \cdot I_m), h_1 \sim N(0, 10 \cdot I_n). \]

Next, the priors for the i-th diagonals of the error covariances for the state equations are taken to be:

\[
\omega_{\beta_i} \sim IG(\nu_{\beta_i}, S_{\beta_i}), \quad i = 1, \ldots, k, \\
\omega_{a_j} \sim IG(\nu_{a_j}, S_{a_j}), \quad j = 1, \ldots, m, \\
\omega_{h_l} \sim IG(\nu_{h_l}, S_{h_l}), \quad l = 1, \ldots, n,
\]

where \( IG(\nu, S) \) denotes the Inverse-Gamma distribution with degree of freedom parameter \( \nu > 0 \) and scale parameter \( S \). Specifically, we set \( \nu_{\beta_i} = \nu_{a_j} = \nu_{h_l} = 5 \). Next, the scale parameter is set so that the prior means are 0.1^2, however we distinguish between VAR coefficients and intercepts by setting the prior mean to 0.01^2 for the former. Finally, when considering the constant variance-covariance matrix we set \( \nu_{\Sigma} = 5 \) and set the scale parameter \( S_{\Sigma} = I_n \). Similarly, for the constant parameters we set \( \bar{\beta}_0 = 0 \) and \( \bar{V}_\beta = 10 \cdot I_k \).

The full conditional distributions for each block of the Gibbs sampler are as follows:

1. Draw from \((\beta|y, h, a, S_{\beta}, S_{h}, S_{a}) \sim N(\hat{\beta}, D_\beta)\), using the precision sampler in Chan and Jeliazkov (2009), where:

\[
\hat{\beta} = D_\beta (X'\Sigma^{-1}y) \\
D_\beta^{-1} = H_\beta S_\beta^{-1}H_\beta + X'\Sigma^{-1}X
\]

where \( y = \begin{bmatrix} y_1 & \ldots & y_T \end{bmatrix}' \), \( X = diag\begin{bmatrix} X_1 & X_2 & \ldots & X_T \end{bmatrix} \), \( \Sigma = diag\begin{bmatrix} \Sigma_1 & \Sigma_2 & \ldots & \Sigma_T \end{bmatrix} \)

and \( H_\beta \) is a \( Tk \times Tk \) first difference matrix.

2. Draw from \((h_i|y, \beta, a, S_{\beta}, S_{h}, S_{a}) \sim N(h_i, D_{h_i})\), using the sampling techniques in
where $y^*_i$, $d_i$, and $\Sigma y^*$ are defined as in Chan and Hsiao (2014) and $H_h$ is a $T \times T$ first difference matrix.

3. Draw from $(a | y, \beta, h, S_\beta, S_h, S_a) \sim N(\hat{a}, D_a^{-1})$, where:

$$\hat{a} = D_a (E'D^{-1}\epsilon)$$
$$D_a^{-1} = H'_a S_a^{-1} H_a + E'D^{-1} E$$

where $\epsilon = [\epsilon_1, \ldots, \epsilon_T]'$, $E = diag[E_1, \ldots, E_T]$, $D = diag[D_1, \ldots, D_T]$ and $H_a$ is a $Tn \times Tn$ first difference matrix. Note that in the $n = 3$ variable case, $E_t$ is defined by:

$$E_t = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
-\epsilon_{1,t} & 0 & 0 & 0 & 0 & 0 \\
0 & -\epsilon_{1,t} & -\epsilon_{2,t} & 0 & 0 & 0 \\
0 & 0 & 0 & -\epsilon_{1,t} & -\epsilon_{2,t} & -\epsilon_{2,t}
\end{bmatrix}$$

4. Draw from $(\omega_\beta | y, h, a, S_\beta, S_h, S_a) \sim IG\left(10 + \frac{T-1}{2}, S_\beta + \frac{1}{2} \sum_{t=1}^{T} (\beta_{i,t} - \beta_{i,t-1})^2 \right)$, where $i = 1, \ldots, Tk$.

5. Draw from $(\omega_h | y, \beta, a, S_\beta, S_h, S_a) \sim IG\left(2 + \frac{T-1}{2}, S_h + \frac{1}{2} \sum_{t=2}^{T} (h_{l,t} - h_{l,t-1})^2 \right)$, where $l = 1, \ldots, n$. 

108
6. Draw from \( (\omega_{nj}|y, \beta, h, S_\beta, S_h) \) \( p \sim IG \left( 2 + \frac{T-1}{2}, S_{a_j} + \frac{1}{2} \sum_{t=2}^{T} (a_{j,t} - a_{j,t-1})^2 \right) \), where \( j = 1, \ldots, n \) \((n-1)\).

### 3.8.1 Appendix D - Stochastic Volatility with Student’s-t Error Distribution

Following Chan and Hsiao (2014) posterior draws for the CVAR-SVt model are obtained through a three block Gibbs sampler that cycles though\(^3\):

1. \( p(h|y, \lambda, \nu) \)
2. \( p(\lambda|y, h, \nu) \)
3. \( p(\nu|y, h, \lambda) \)

Using the sampling techniques in Chan and Hsiao (2014) the full conditional distributions for the \( i \)-th variable in the set containing interest, GDP growth and inflation in each Gibbs step are as follows:

1. **Draw from** \( p(h|y, \lambda, \nu) \sim N \left( \hat{h}_i, D_{h_i} \right) \) **where:**

   \[
   \hat{h}_i = D_{h_i} \left( \Sigma_{y_i}^{-1} (y^*_i - d_i) \right) \tag{3.45}
   \]

   \[
   D_{h_i}^{-1} = H_h \Sigma_{h_i}^{-1} H_h + \Sigma_{y_i}^{-1} \tag{3.46}
   \]

   where \( y^*_i, d_i \) and \( \Sigma_{y_i} \) are defined as in Chan and Hsiao (2014) and \( H_h \) is a \( T \times T \) first difference matrix. The key difference between this step and that in the TVP-VAR-SV model rests in the definitions of the variables.

\(^3\)Note that for estimation of the models with Student-t errors and no stochastic volatility we only require blocks two and three. Specifically, let: \( \sigma_i^2 \) denote the time-invariant variance of each endogenous variable, then it follows a standard Inverse-Gamma prior distribution: \( \sigma_i^2 \sim IG(\xi, \Xi) \). We set the degree of freedom hyperparameter \( \xi = 5 \) and the scale parameter \( \Xi \) to have a prior mean of 1. Since the variance terms are independent, in the univariate case the same structure is followed.
2. To draw from $p(\lambda_{i,t}|y, h_{i,t}, \nu_i)$ note that since $\lambda_{i,1}, \ldots, \lambda_{i,T}$ are conditionally independent of the model parameters and the data, we can sample each of them sequentially. A simple application of Bayes Theorem shows that: 

$$p(\lambda_{i,t}|y, h_{i,t}, \nu_i) \sim IG\left(\frac{1+\nu_i}{2}, \frac{\nu_i + e^{-h_{i,t}((\tilde{y}_{i,t})^2)}}{2}\right).$$

3. To draw from $p(\nu|y, h, \lambda)$ again note that the degree freedom parameters $\nu_i$ associated with $\lambda_i$ are also conditionally independent. Following Chan and Hsiao (2014) we maximize the log-density:

$$log(\nu_i|\lambda) = \frac{T\nu_i}{2} log(\nu_i/2) - T log\Gamma(\nu_i/2) - \left(\frac{\nu_i}{2} + 1\right) \sum_{t=1}^{T} log\lambda_{i,t} - 0.5 \sum_{t=1}^{T} \lambda_{i,t}^{-1}$$

where $c$ is normalizing constant and $\nu_i \in (0, \bar{\nu})$. The first and second derivatives are:

$$\frac{dlog(\nu_i|\lambda)}{d\nu_i} = \frac{T}{2} log(\nu_i/2) - \frac{T}{2} - \frac{T}{2} \Psi(\nu_i/2) - 0.5 \sum_{t=1}^{T} log\lambda_{i,t} - 0.5 \sum_{t=1}^{T} \lambda_{i,t}^{-1}$$

$$\frac{d^2log(\nu_i|\lambda)}{d\nu_i^2} = \frac{T}{2\nu_i} - \frac{T}{4} \Psi'(\nu_i/2)$$

(3.49)

where $\Psi(x) = \frac{d}{dx}$ and $\Psi'(x) = \frac{d}{dx} \Psi(x)$ are respectively the digamma and trigamma functions. Since the first and second derivatives can be evaluated quickly, we maximize the log $p(\nu_i|\lambda)$ using Newton-Raphson method and obtain the mode and the negative hessian evaluated at the mode denoted $\hat{\nu}_i$ and $K_{\nu_i}$ respectively. We then implement an independence chain Metropolis-Hastings step with a proposal distribution given by $N(\hat{\nu}_i, K_{\nu_i})$. The only restriction we place is that draws from the Metropolis-Hastings step be greater than two (i.e. $\nu > 2$). This technical restriction is necessary to ensure a finite variance.
Chapter 4

Assessing the Synchronicity and Nature of Australian State Business Cycles

This paper is currently under review in *Oxford Bulletin of Economics and Statistics*
4.1 Introduction

The onset of the mining boom in the mid-2000s has seen the economic performances of Western Australia and Queensland far exceed their counterpart States and Territories (Garton 2008). Throughout this time, the high level of investment directed towards the mining sector has seen capital and labour extracted away from non-mining States; thus reducing their rates of economic growth (Garton 2008). In addition to this hit, Norman and Walker (2007) provide evidence that intense global competition, especially from China, has furthered the economic slowdown of manufacturing States such as Victoria and New South Wales. This asymmetrical economic performance across States has been popularly characterised as a ’two speed economy’ (Garton 2008). Given that aggregate macroeconomic activity is determined by cumulative activity of each of the nation states, it is important that both federal and state governments are aware of state specific economic fluctuations. With this policy issue in mind, the objective of this paper is to assess the synchronicity and nature of business cycles in Australian states and territories; New South Wales (NSW), Victoria (VIC), Queensland (QLD), South Australia (SA), Western Australia (WA), Tasmania (TAS), the Australian Capital Territory (ACT) and Northern Territory (NT).

The current empirical literature on Australian State business cycles is relatively scarce. In fact, Norman and Walker (2007) is the only known study that attempts to examine the degree of co-movement among Australian State business cycles. To elicit this insight, the Norman and Walker (2007) methodology begins by conducting a correlation analysis of Gross State Product; a State counterpart to Gross Domestic Product, to document statistically significant evidence of co-movement amongst the economically larger states of NSW, VIC, QLD and WA. This result is consistent with the findings of Dixon and Shepherd (2001) who conducted a related study examining co-movements in State unemployment, as opposed to economic output. Having provided statistically significant evidence of co-movement within the economically larger State business cycles, Norman and Walker (2007) then employ an unobserved components model through which they find that the main source of fluctuations in state specific economic activity is
driven by a common shock. This finding then motivates their conclusion that common shocks are more important than there State specific counterparts in explaining State business cycle fluctuations.

One concern in utilizing an unobserved components model to assess the synchronicity and nature of Australian State business cycles is that it fails to account for the autoregressive nature of macroeconomic time series. To account for this important feature of the data, in this paper we propose an augmented version of an otherwise standard time-varying parameter Panel Bayesian vector-autoregression (BVAR) model. Originally put forth by Canova et al (2007), the time-varying parameter Panel BVAR model has been employed in various studies which assess the similarities and co-movement of business cycles among different countries. For instance, Canova et al (2007) and Canova and Ciccarelli (2012) employ this methodology to respectively assess the similarities and convergence of business cycles of the G-7 and Mediterranean economies. More recently, Ciccarelli, Ortega and Valderrama (2016) employed the model to investigate the evolution and heterogeneity in macro-financial linkages and international spillover effects across developed economies.

Compared to a traditional BVAR model, the Panel BVAR has two distinct advantages when seeking to examine the synchronicity of business cycles. In the first instance, the Panel BVAR is designed for large scale dynamic models that display unit specific dynamics and cross country lagged inter-dependencies; as opposed to a traditional BVAR which suffers from the curse of dimensionality. Next, the parsimonious parameterisation proposed in Canova and Ciccarelli (2009) allows the researcher to introduce time-varying coefficients through which indicators of common and country specific cycles are recursively constructed and dynamically span across country interdependencies.

One shortcoming of the Panel BVAR model when it comes to addressing our research question is that it assumes that the size and frequency of macroeconomic shocks are constant over time. More precisely, the aforementioned mining boom signifies that allowing for structural instabilities in shocks is also an important feature of Australian
macroeconomic time series. To accommodate this feature of the data, we propose a simple extension of the time-varying parameter Panel BVAR by incorporating the common stochastic volatility factor into the error covariance structure. Originally proposed by Carriero, Clark and Marcellino (2016), this common stochastic volatility factor structure significantly improves model fit and forecast accuracy of large BVARs compared their constant volatility counterparts.\textsuperscript{4}

In addition to this simple methodological extension of the time-varying parameter Panel BVAR model, we highlight that we also contribute towards the efficient estimation of the model both with and without the additional common stochastic volatility factor. More precisely, following Chan and Jeliazkov (2009) and Chan and Hsiao (2014), we implement precision based algorithms in the estimation of the model's time-varying parameters and the common stochastic volatility factor respectively. In all the previous studies discussed above, standard Kalman filtering and smoothing techniques are used to estimate the time-varying parameter Panel BVAR. The main reason we adopt this precision sampler technique is due to its computation efficiency advantage over the Kalman filtering and smoothing techniques.

Taken together, this study can been viewed primarily as an extension of the study by Norman and Walker (2007) with important secondary methodological contributions of the time-varying parameter Panel BVAR model. The results of the analysis reveal several key insights. Firstly, from a methodological perspective we show that the inclusion of the common stochastic volatility factor to the model significantly improves the in-sample goodness of fit. This result confirms our aforementioned hypothesis that stochastic volatility is an important feature when modelling the Australian economy. In addition to being statistically important, the common stochastic volatility factor reveals that the degree of volatility in the Australian economy was more pronounced during the Asian Financial Crisis rather than the recent Global Financial Crisis (GFC); a latent feature to the models constant volatility counterpart. Secondly, in addition to

\textsuperscript{4}For related literature on the use of stochastic volatility in improving model fit and forecastability, we refer the reader to, for example: Clark (2014) and Clark and Ravazzolo (2015) for out-of-sample point and density forecasting, or Chan and Eisenstat (2016) for in-sample analysis.
other interesting features, the common indicator from the time-varying Panel BVAR is able to capture the early 1990s’ recession along with the GFC. This modelling feature supports the hypothesis of strong common co-movement across each Australian State during times of economic contraction. Thirdly, we found that, on average, the common indicator is able to explain about 39 per cent of fluctuations across each of the State indicators and about 25, 8, 9 and 111 per cent of fluctuations in consumption, employment, retail turnover and investment indices respectively. Lastly, we found the common indicator fluctuations closely follow the trend line of the Organisation for Economic Co-operation and Development (OECD) composite leading indicators (CLI) for Australia, especially during the 2000s. This suggests that the common indicator captures the majority of fluctuations in economic activity for our sample period.

Finally, in regards to Australian State business cycles, we found that the average degree of synchronisation across Australian States cycles has decreased to about half, in terms of correlation from the 1990s to 2000s. It was also found that the fall in consumption growth was the main factor in driving the negative effects of the GFC across the majority of the states. However, for the SA and NT economies, we found that State-specific idiosyncratic factors were the dominant feature in driving this crisis. Turning to the period of 2013 to 2015, all four common variable type indices had minimal impact in the contribution of the downturn experienced within the WA, SA, VIC, ACT, and NT economies.

With these results in mind, our analysis has important implications for policymakers at both a state and federal level. Given that aggregate macroeconomic activity is determined by cumulative activity of each of the nation states, the results suggests that federal governments should award state governments greater autonomy in handling state specific cyclical fluctuations.

The rest of this paper is organised as follows. Section 2 explains the empirical methodology by illustrating the estimation procedure for the time-varying parameter Panel BVAR model, the data and the priors for the model. Section 3 describes empirical results from the time-varying parameter Panel BVAR model, and Section 4 concludes
and discusses the results.

4.2 Econometric Methodology

The econometric model that we use in this paper is a time-varying parameter Panel BVAR from Canova et al (2007) and Canova and Ciccarelli (2009). The econometric model can be written as:

\[
y_{it} = c_{it} + A_{1, it}Y_{t-1} + \ldots + A_{p, it}Y_{t-p} + u_{it},
\]

where \(t = 1, \ldots, T\) denotes time, \(i = 1, \ldots, N\) denotes the number of Australian States in the model and \(A_{p, it}\) are \(G \times NG\) time-varying matrices of the coefficients for each lag \(j = 1, \ldots p\). The vector \(y_{it}\) is \(G \times 1\) of observed endogenous variables that consist of consumption growth, employment growth, retail turnover growth and investment growth for each State \(i\) and \(Y_t = (y'_{1t}, \ldots, y'_{Nt})'\). Note that in (4.1), the model displays cross-unit lagged interdependencies, where the endogenous variables for each Australian State depends on the lags of the endogenous variables for every Australian State. Lastly, \(c_{it}\) and \(u_{it}\) are \(G \times 1\) vectors of intercepts and random disturbances respectively.

The econometric model (4.1) exhibits three important features in our study. Firstly, the coefficients are allowed to be time-varying. Without time variation, it would be impossible to study the evolution of business cycle characteristics over time. Secondly, dynamic relationships are allowed to be State specific. Without this feature, heterogeneity bias may be present and economic conclusions can become easily distorted. Lastly, the cross-unit lagged inter-dependencies, which are captured by the coefficients matrix \(A_{p, it}\) in the model, are likely to be important in explaining the dynamics of multi-region (Australian States) data. Canova et al (2007) notes that these three factors are essential when one wants to study the similarities, propagation and time variations in the structure of business cycles across regions (Australian States).

Model (4.1) can be re-written into standard linear regression matrix form:
\[ Y_t = X_t \beta_t + u_t, \quad (4.2) \]

where \( X_t = I_{NG} \otimes (Y_{t-1}', \ldots, Y_{t-p}') \) is a \( NG \times NGk \) matrix (where \( k = NGp + 1 \)), \( \beta_t = vec([A_t, c_t']) \), \( A_t = [A_{1t}', \ldots, A_{Nt}'] \), \( A_{it} = [A_{1it}, \ldots, A_{pit}] \), \( c_t = (c_{1t}', \ldots, c_{Nt}') \) and \( u_t = (u_{1t}', \ldots, u_{Nt}') \). Note that \( \beta_t \) is an \( NGk \times 1 \) vector that denotes the number of parameters in each time period. However, it is difficult to estimate the econometric model (4.2) using classical methods due to the sheer dimensionality of the model. To overcome this dimensionality problem, Canova et al (2007) and Canova and Ciccarelli (2009) assumes \( \beta_t \) follows a factor structure:

\[ \beta_t = \Xi_1 \theta_{1,t} + \Xi_2 \theta_{2,t} + \Xi_3 \theta_{3,t}, \quad (4.3) \]

where \( \Xi_1, \Xi_2, \Xi_3 \) are matrices of dimensions \( NGk \times 1, NGk \times N, NGk \times G \) respectively and \( \theta_{1t}, \theta_{2t}, \theta_{3t} \) are mutually orthogonal. \( \theta_{1t} \) is a scalar that captures components in the coefficient vector that are common across States and variables. \( \theta_{2t} \) is an \( N \times 1 \) vector that captures movements in the coefficient vector which are common within the States. \( \theta_{3t} \) is a \( G \times 1 \) vector that captures movements in the coefficient vector which are variable specific. By factoring \( \beta_t \) in (4.3), it transforms the over-parameterised panel VAR into a parsimonious SUR model, where the regressors are averages of certain right-hand side VAR variables. Instead of estimating \( NGk \times 1 \) (\( \beta_t \)) coefficients, only \( 1 + N + G \) (\( \theta_t \)) coefficients are estimated in the model in each period of time. Let \( \theta_t = (\theta_{1t}, \theta_{2t}, \theta_{3t})' \), Canova et al (2007) and Canova and Ciccarelli (2009) assume \( \theta_t \) evolves over time according to a random walk:

\[ \theta_t = \theta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \Omega), \quad (4.4) \]

where \( \Omega = diag(\omega_1^2, \ldots, \omega_m^2) \) and \( m = 1 + N + G \) denotes the dimension of \( \theta_t \). The
random walk assumption helps focus on permanent shifts and reduces the number of parameters in the estimation procedure.

In sum, we can substitute (4.3) into (4.2):

\[ Y_t = Z_t \theta_t + u_t, \quad (4.5) \]

where \( Z_t = X_t \Xi, \) \( \Xi = [\Xi_1, \Xi_2, \Xi_3] \). Economically, the decomposition in (5) allows us to measure the relative importance of common, State and variable specific influences in explaining fluctuations in \( Y_t \). In fact, \( X_t \Xi_1 \theta_{1t} \) represents a common indicator for \( Y_t \), while \( X_t \Xi_2 \theta_{2t} \) represents the vector of State specific indicators, and \( X_t \Xi_3 \theta_{3t} \) represents a vector of variable specific indicators. By construction, all these indicators correlate with each other, that is \( X_t \) enters in all of them. But as the number of States and variables becomes large the correlation will tend towards zero. In the appendix below we illustrate a simple example of the model.

In the model described above, Canova et al (2007) and Canova and Ciccarelli (2009) assume a time-invariant variance-covariance matrix structure of the shocks. We extend this model by allowing for a common stochastic volatility factor process (from Carriero, Clark and Marcellino (2016)) within the error structure of the model. Thus, our model includes (4.5) with:

\[ u_t \sim N(0, e^{h_t} \Sigma_u), \quad (4.6) \]

\[ h_t = \rho h_{t-1} + \zeta_t, \quad \zeta_t \sim N(0, \sigma^2_h), \quad (4.7) \]

where \( |\rho| < 1 \). Carriero, Clark and Marcellino (2016) commented that this common stochastic volatility factor structure is ideal for models with a large data set. From their results, Carriero, Clark and Marcellino (2016) found that large BVAR models...
with a common stochastic volatility factor error structure significantly improves model fit and forecast accuracy, when compared to a standard conventional large BVAR. The adoption of a common stochastic volatility factor is very important since there have been many studies undertaken in the literature that have documented the importance of stochastic volatility in improving model fit and forecastability (for instance see Clark (2014), Clark and Ravazzolo (2015), and Chan and Eisenstat (2016)). The inclusion of this time-varying volatility error specification allows us to capture any common structural shifts or breaks which are commonly found in the majority of macroeconomic data.

4.2.1 Data

In our study, we employed four business cycles variables commonly used within the literature, which are: consumption, employment, retail turnover and investment. The data frequency is quarterly, and the sample period covers dates between 1988Q4 to 2015Q1. Due to the relatively small sample period, and the risk of over-parameterisation, we only impose one lag length on the model. All data variables were gathered from the Australian Bureau of Statistics (ABS). Consumption is final household consumption expenditure and investment is private gross fixed capital formation. Following Ciccarelli, Ortega and Valderrama (2016), all the data are annualised, deseasonalised, deflated and standardised growth rates.

4.2.2 Priors

To calculate the posterior distribution for the model’s parameters we implement the prior distributions of:

\[ \Sigma_u \sim IW(z_1, Q_1), \]
\[ \sigma^2_h \sim G(w_1, S_1), \quad (4.8) \]

\[ \omega^2_i \sim IG(w_0, S_0), \quad i = 1, \ldots, m., \]

\[ \rho \sim N(\mu, V_\rho)1(|\rho| < 1), \]

where \(1(.)\) denotes an indicator function and we initialised

\[ \theta_1 \sim N(\theta_0, V_\theta), \quad h_1 \sim N(h_0, \frac{\sigma^2_h}{(1-\rho^2)}), \quad (4.9) \]

Note: here we impose the stationarity condition \(|\rho| < 1\) through the prior distribution of \(\rho\). The hyperparameters are either obtained from the data to tune the prior to the specific application, selected a-priori to produce relatively loose priors or initialised with a training sample. Since our sample period is relatively short, there is no training sample available to tune the priors. Therefore, we impose \(z_1 = NG + 5\), \(Q_1 = 5I_{NG}\), \(\theta_0 = 0\), \(V_\theta = 10I_{NG}\), \(\mu = 0\), \(h_0 = 0\), \(V_\rho = 1\), \(w_0 = w_1 = 5\), \(S_0 = (.01)^2 \times (w_0 - 1)\) and \(S_1 = .01\).

### 4.2.3 Estimation

The time-varying parameter Panel BVAR with a common stochastic volatility factor is estimated through a standard Markov Chain Monte Carlo (MCMC) method. Canova and Ciccarelli (2009) follow a standard Kalman filtering and smoothing techniques from Chib and Greenberg (1995) to estimate the model’s parameters. However, we adopt a different estimation technique, using the precision sampler from Chan and
Jeliazkov (2009) and Chan and Hsiao (2014) to estimate the model’s parameters and the common stochastic volatility factor respectively. The main reason we adopt this precision sampler technique is due to its computation efficiency advantage over the Kalman filtering and smoothing techniques. We stored 35,000 draws after the initial 15,000 draws were discarded. Further details on the Gibbs Sampler can be found in the appendix below. An important issue when using a Gibbs Sampler is the convergence of the limiting distribution of the sample to the posterior distribution. In theory, as the number of draws reaches infinity, the sampler should converge. However, in any applied work, determining how many draws that it will take to make the sample converge is very difficult. To assess whether our sample has converged, we compute a convergence diagnostic from Geweke (1992). A Geweke (1992) convergence diagnostic is calculated by taking the difference between the means $\bar{g}_a = \frac{1}{n_a} \sum_{i=1}^{n_a} \theta(i)$, based on the first $n_a$ draws and $\bar{g}_b = \frac{1}{n_b} \sum_{i=1}^{n_b} \theta(i)$, based on the last $n_b$ draws and dividing by the asymptotic standard errors of the difference $\sqrt{\frac{\sigma^2_a}{n_a} + \frac{\sigma^2_b}{n_b}}$.

Geweke (1992) suggests that $n_a$ and $n_b$ should be the first 10 percent and last 50 percent of the total draws respectively. Thus, in terms of our estimation, $n_a$ is the first 2,000 draws and $n_b$ is the last 10,000 draws. If the sequence of the MCMC sampling is stationary, then by the central limit theorem, the distribution of this diagnostic converges to a standard normal. Table 1 shows the posterior means, standard deviations and the convergence diagnostics for selected parameter estimates. Notice that for all the parameter estimates, the convergence diagnostics are all less than the 5 per cent significance level. This implies that the null hypothesis of the convergence to the posterior distribution is not rejected. We also report the trace plots of these selected parameters in Figure 4.1. For each parameter the chain appears to be stable. Therefore, both the Geweke convergence diagnostics and Figure 4.1 indicate that the parameters and state variables are efficiently drawn from the posterior distributions.
Table 4.1: Geweke Convergence Diagnostics Statistic

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean ((n_a))</th>
<th>Stdev. ((n_a))</th>
<th>Mean ((n_b))</th>
<th>Stdev. ((n_b))</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_{10})</td>
<td>0.188</td>
<td>0.005</td>
<td>0.183</td>
<td>0.002</td>
<td>1.000</td>
</tr>
<tr>
<td>(\theta_{550})</td>
<td>0.115</td>
<td>0.004</td>
<td>0.109</td>
<td>0.002</td>
<td>1.019</td>
</tr>
<tr>
<td>(\theta_{1600})</td>
<td>-0.132</td>
<td>0.003</td>
<td>-0.130</td>
<td>0.001</td>
<td>-1.036</td>
</tr>
<tr>
<td>(\omega_3)</td>
<td>0.002</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.995</td>
</tr>
<tr>
<td>(\omega_7)</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>1.085</td>
</tr>
<tr>
<td>((\Sigma_n)_{20})</td>
<td>-0.021</td>
<td>0.001</td>
<td>-0.020</td>
<td>0.000</td>
<td>-1.070</td>
</tr>
<tr>
<td>(h_{60})</td>
<td>1.294</td>
<td>0.002</td>
<td>1.297</td>
<td>0.001</td>
<td>0.996</td>
</tr>
</tbody>
</table>

4.3 Empirical Results

We estimated three components or indices of the econometric model: a common indicator for all series, one State-specific indicator for each Australian State and four variable-specific indices. For the first sub-section, we undertake a model comparison exercise. In the second sub-section, we examined the common factor stochastic volatility across all the variables. For the third sub-section, we determined whether there was significant common movement in the four variables across each Australian State. In regards to the fourth sub-section, we assessed the synchronicity of each Australian State-specific indicator. For the last sub-section, we tried to determine the relative weight of each of the four variable-type indices in explaining the GFC across each Australian State.

4.4 Model Comparison

To determine whether the proposed new methodological feature of the model is favoured by the data, we undertake a model comparison exercise in which we compare the time-varying parameter Panel BVAR with a common stochastic volatility factor (TVP-PVAR-CSV) against the time-varying parameter Panel BVAR with a constant variance (TVP-PVAR). To this end, the marginal likelihood for each of the models is computed as the product of the one-step-ahead predictive likelihood of Geweke and Amisano (2011). The reason for using the one-step-ahead predictive likelihood as compared to
the harmonic mean estimator as in Canova and Ciccarelli (2012) is that recent work has shown that this approach can be extremely inaccurate. More precisely, Chan and Grant (2015) show that the 14 marginal likelihood estimates computed using 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.

Following Geweke and Amisano (2011) the marginal likelihood for the $i^{th}$ model is:

$$p(Y|M_i) = p(Y_1|M_i) \prod_{t=1}^{T-1} p(Y_{t+1}|Y_t, \ldots, Y_1, M_i),$$

(4.10)

where $p(Y_{t+1}|Y_t, \ldots, Y_1, M_i)$ is the one-step-ahead predictive likelihood given the data up to time $t$. The marginal likelihood results are reported in Table 2. The results clearly show that the TVP-PVAR-CSV is clearly the better model. This means that the addition of the common stochastic volatility factor in the econometric methodology is a key feature of Australian State level data.

<table>
<thead>
<tr>
<th></th>
<th>TVP-PVAR</th>
<th>TVP-PVAR-CSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log marginal likelihood</td>
<td>-7828.50</td>
<td>-6138.70</td>
</tr>
</tbody>
</table>

4.4.1 Common Stochastic Volatility

Figure 4.2 plots the posterior median of the common stochastic volatility factor, expressed as standard deviation, from the model. It is clear that there is significant time variation within the common stochastic volatility factor. In general, the common stochastic volatility factor trend appears to be declining over time. This declining trend implies that the Australian business cycle is less susceptible to large fluctuations or shocks over time. However, the common volatility does exhibit a significant increase during the late 1990s to early 2000s. This sharp increase in common volatility could be attributed either to the introduction of the Goods and Services Tax, or international factors, mostly likely the Asian Financial Crisis of 1997-98. Also, there appears to be no pronounced jump in volatility during the recent Global Financial Crisis, which sug-
gests that the Global Financial Crisis had less of an impact on the Australian economy. This result is plausible since, technically, the Australian economy did not experience a recession during this crisis period. Another issue is that Australia has a close proximity to Asia and most of Australia’s major trading partners are from this region. Therefore, the Australian economy will be highly influenced by the economic performances of countries in this region compared to the US economy. These two crises highlights the importance of stochastic volatility in the error structure since each crisis (or shocks) have differing impacts on the economy.

To see whether this result is robust, we also plotted the posterior median estimates of the stochastic volatility variable for each of the aggregated variables from an univariate one lag autoregressive model with stochastic volatility (AR-SV(1)) in Figure 4.3. The stochastic volatility for total consumption growth has the same declining trend as in the common stochastic volatility factor in Figure 4.2. However, for total employment, investment and retail turnover growths, the stochastic volatility for all these variables only exhibits a declining trend after 2000. Notice that the stochastic volatility for both investment and retail growths peaked around the late 1990s and early 2000s. This could provide explanation towards the sharp increase in the common stochastic volatility factor at the same time. Therefore, the common stochastic volatility factor appears to capture the declining volatility trend in all these four variables.

4.4.2 Commonality

Figure 4.4 shows the evolution of the common indicator for all series, expressed as the standard deviations from the historical average annual growth rates. The common indicator is very volatile and the majority of the 68 per cent posterior credible interval consistently includes zero over time. This implies there is a large degree of parameter uncertainty associated with the common indicator. However, the common indicator does appears to capture the early 1990s’ recession and the recent slowdown that the Australian economy experienced due to the GFC. Similarly, in the study by Canova and Ciccarelli (2012), they also found (see Table 2: Percentage of variance explained by
Table 4.3: Percentage of variance explained by the common indicator

<table>
<thead>
<tr>
<th></th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSW</td>
<td>26.60%</td>
</tr>
<tr>
<td>VIC</td>
<td>11.21%</td>
</tr>
<tr>
<td>QLD</td>
<td>20.31%</td>
</tr>
<tr>
<td>SA</td>
<td>45.75%</td>
</tr>
<tr>
<td>WA</td>
<td>19.27%</td>
</tr>
<tr>
<td>TAS</td>
<td>21.90%</td>
</tr>
<tr>
<td>NT</td>
<td>83.78%</td>
</tr>
<tr>
<td>ACT</td>
<td>86.81%</td>
</tr>
<tr>
<td>Average</td>
<td>39.45%</td>
</tr>
</tbody>
</table>

the common indicator) that the size of the co-movement across variables and countries appears to be more similar in contraction than in expansion. To measure the contribution of the common indicator on each State, we follow Canova and Ciccarelli (2012) and compute a simple numerical measure that explains how much of the variance of each of the State indicator is explained by the common indicator. Table 2 shows the results of this measure for each Australian State. We find on average the common indicator explains about 39 per cent of the fluctuations across each of the State indicators. For comparison, the study by Canova, Ciccarelli and Ortega (2007) found that the common indicator explains about 30 per cent fluctuations across each of the G7 countries. The common indicator appears to have the largest influence in SA, NT and ACT business cycles or State-specific indicators compared to the other Australian States. Note that, from Table 2, if we exclude both ACT and NT from the calculation, we find that, on average, the common indicator explains only about 24 per cent of the fluctuations across each of the State indicators.

Although there is a large degree of parameter uncertainty associated with common indicator, the results from Table 1 appear to indicate that the common indicator does have significant influence across each State’s business cycle. To further investigate whether the common indicator is robust, in Figure 4.5 we plot the posterior median of the common indicator together with the OECD CLIs for Australia. The OECD CLIs are indices that measure fluctuations of economic activity or business cycles of a particular country. Figure 4.5 shows that our model common indicator captures the majority of the upturn and downturn displayed in the OECD CLI. It appears that since
2000, our common indicator trend or fluctuations closely resemble the trend line of the OECD CLI. In addition, our common indicator captures the recent slowdown due to the GFC earlier than the OECD CLI. Therefore, based on the visual comparison with the OECD CLI, it appears that the common indicator captures the majority of the fluctuations in economic activity present during this sample period.

Figure 4.6 shows the variable specific indices for all the Australian States. Apart from the employment indicator, the majority of the other variables specific indices 68 per cent posterior credible interval includes zero for most dates. Similarly, this implies that consumption, retail turnover and investment variable components/indices exhibit a large degree of parameter uncertainty. This means that each of these three variables do not feature a significant common movement across the States. However, the consumption indicator does appear to be statistically significant from 2006 onward, as the 68 per cent posterior credible interval lies above and below zero. Furthermore, we find that the common indicator explains about 25, 8, 9 and 111 per cent of fluctuations in consumption, employment, retail turnover and investment indices respectively. This result is consistent with theory since both consumption and investment are very volatile macroeconomic variables, and we would expect these two variables to be the main driver of fluctuations across the State's business cycle.

The early 1990s' recession appeared to have a significant negative impact on employment growth across the States. Both consumption and investment growth were also affected by this recession, yet their fall was considerably less than that of employment growth. However, irrespective of the aforementioned, this recession only had a minimal impact on total State retail turnover growth. The introduction of the Goods and Services Tax during the early 2000s appears to have only a negative impact on retail turnover growth. For the period before the GFC, the posterior median for both consumption and employment growths were significantly positive which implies that the Australian economy was in a period of successful growth before the GFC hit. The GFC had an immediate, negative impact on consumption, employment and retail turnover growths across all the States. Investment growth appeared to be unaffected by this
crisis. Employment and consumption growths were the worst affected by this crisis in comparison to the two other variables. Both consumption and employment growths during 2013 to 2015 have remained sluggish and they have not recovered back to its pre-GFC levels. In comparison, retail turnover growth became close to positive around late 2014 or early 2015. Therefore, from variable-specific indices, the fall in consumption and employment growths, across all the States, can be attributed to the slowdown that the Australian economy experienced during the GFC.

4.4.3 Convergence or Divergence

To assess the similarities and synchronicity of each Australian State business cycles, we plot each Australian State-specific indicator or component in Figure 4.7. It is clear that the majority of the State’s 68 per cent posterior credible interval includes zero for most dates. This implies that their explanatory power for domestic fluctuations over the sample period is small. For example, excluding TAS, during the period between the early 2000s to the mid 2000s, cyclical fluctuations across the States were very minor. However, upon visual examination of Figure 4.7, each States cyclical fluctuations differ in intensity, timing and duration. For instance, cyclical fluctuations in QLD and TAS are relatively more persistent compared to the other States. The early 1990s’ recession had a negative effect on the majority of States. It appears that the both VIC and WA economies were the worst affected by this recession, in comparison to other States. In regards to the GFC, the QLD economy was the worst affected out of all the States; the QLD State indicator has been falling since 2008. Both the WA and TAS economies were also affected by this crisis, and similarly their recovery appears to also be sluggish. The GFC seemed to have a minimal impact on both the NSW and VIC economies. However, the NSW economy appears to experience a downturn during late 2014, or early 2015. For the Territories, cyclical fluctuations are very weak across both the ACT and NT economies. Based on these descriptive findings, each State’s cycles has unique a pattern, and there appear to be no common similarities amongst the States.

To further assess the synchronicity of each Australian State’s business cycles, we
followed a similar strategy to Canova and Ciccarelli (2012) and plot, in Figures 4.8, 4.9 and 4.10, the pairwise rolling correlations between each State indicator. The rolling correlations are computed using 10 years of data ending at the date listed on the horizontal axis. Due to our small sample size, we were only able to compute the rolling pairwise correlations for the 2000’s period. According to Canova and Ciccarelli (2012), if convergence (divergence) takes place we should see these correlations uniformly increase (decrease) with time.

The correlations display distinct periods of convergence and divergence across the States. Firstly, focusing on mining States, the correlation between QLD and WA has remained surprisingly low during the period of 2002 to 2009 when the mining boom was most prevalent. Apart from the periods around 2010, SA is the only non-mining State that has a consistently high correlation with QLD over the 2000s. WA mainly has a consistently negative correlation with all the other States, except for TAS and ACT. For the two largest State economies, NSW and VIC, their correlation with each other has also persistently remained negative over the 2000s. This is also the case for the Territories, where they appear to be less correlated with major Australian States. Figure 4.11 reports the average correlations between a State indicator with all the others States for the periods of the 1990s and 2000s. It appears that only WA and ACT have increased their degree of synchronisation with other States between the 1990s to 2000s. Overall, Figure 4.11 shows that the degree of synchronisation on average across the States, in terms of correlation, has decreased by about half from the 1990s to 2000s. In other words, State economic performance or cyclical fluctuation has clearly diverged during the 2000s period.

The above findings uncover four facts. Firstly, it appears the mining boom in QLD and WA were not synchronised; each of their booms were driven by idiosyncratic factors. An explanation for this non-synchronisation could be due to the differentiation in their exports of resources and minerals. For example, QLD is a large exporter of coal production, whilst for WA, iron ore dominates the exports. Secondly, the mining boom in QLD appears to have a positive effect on the SA economy. Thirdly, the mining
boom of WA appears to have no spillover effects to the other major States economies. Lastly, the economic performances of the Territories (ACT and NT) appear to be less synchronised with the major States. Therefore, these findings show that cyclical fluctuations were clearly different and became less synchronised across each State and Territories during the 200's. Thus, there is evidence of heterogeneity present within each State’s business cycle.

### 4.4.4 Historical Decomposition

To determine the relative weight of each of the four variable type indices in explaining the GFC across each Australian State, we compute the historical decomposition of each State’s final demand growth, in Figures 4.12 and 4.13, for the period between 2005Q3 to 2015Q1. The historical decomposition is based on an estimation of a State by State factor-augmented VAR for final demand growth, which consists of the aggregation of both investment and consumption (both private and public), and the four variable-specific indices estimated from our model. For the identification of these State VAR, we follow a standard recursive assumption and order the investment indicator first, followed by consumption, sales, employment indices and the State’s final demand growth last. All four variable-specific indices appear to have a significant influence on cyclical fluctuations across each State during this period.

Figures 4.12 and 4.13 show that the majority of States were hit by the GFC during late 2008 and early 2009. Falling consumption growth appears to be the dominant factor in driving the crisis in the TAS, VIC and QLD economies. This fall also had negative impacts on the NSW, ACT and WA economies, however, the impact or contribution is relatively small compared to the three aforementioned States. However, it should be noted that, this crisis can also be explained by a fall in employment growth in VIC, QLD and ACT economies. For the SA and NT economies, State-related idiosyncratic factors were also a dominant feature in driving the crisis. Turning to the period of 2013 to 2015, it is clear that the majority of States have not recovered from the GFC. For instance, falling consumption growth again appears to be contributing to the
downturn in both the QLD and TAS economies. Apart from the QLD, TAS and NSW economies, all four common variable-specific indices had a minimal contribution on the downturn experienced by the other States. Therefore, the results from the historical decomposition show that the fall in consumption growth was the main factor in driving the negative effects of the GFC across the majority of the States.

4.5 Conclusion

The objective of this paper is to assess the nature and synchronicity of Australian State business cycles. To this end, the econometric methodology proposes a simple extension of the time-varying parameter Panel BVAR in which the error structure is augmented to have an additional common stochastic volatility factor. Another contribution that this study makes is the implementation of precision based algorithms in the estimation of the model's time-varying coefficients and the common stochastic volatility factor. The proposed algorithm is important as it allows for greater computation efficiency over the traditional Kalman filtering and smoothing techniques used in all previous studies with a time-varying parameter Panel BVAR model. Taken together, this study can thus be viewed primarily as an extension of the study by Norman and Walker (2007) with important secondary methodological contributions of the time-varying parameter Panel BVAR model. This econometric framework provides several advantageous to those previously employed in this literature.

In the first instance, the constant error time-varying parameter Panel BVAR model provides a flexible structure which allows for multiple types of contemporaneous and lagged time varying co-movements within cyclical fluctuations across variables and States. Moreover, the parsimonious parameterisation of this model allows us to endogenously produce an index structure where indicators of common and State specific cycles are observable, recursively constructed and dynamically span across State interdependencies. By constructing these indexes, we can therefore assess, compare and contrast each State-specific indicator, and determine the synchronisation of these indices. The model also allows us to compare the relative weight of the four variable-type
indices in explaining the GFC across each Australian State.

Next, the inclusion of the common stochastic volatility factor is important because it allows us to capture any common structural shifts within the covariance structure of the macroeconomic variables used in the analysis. Importantly, a model comparison exercise revealed that the addition of the common stochastic volatility factor significantly improves the model’s in-sample-fit. This suggests that stochastic volatility is an important feature when modelling the Australian economy. More precisely, the common stochastic volatility factor reveals that the Australian economy exhibited a large jump in volatility during the Asian Financial Crisis of 1997-98, but no such jump was observed during the recent Global Financial Crisis. This result is plausible because the Australian economy technically did not experience a recession during the recent crisis period, however Australia’s heavily reliance of trading partners within the Asian region. Whilst our results provide no investigation into the significance of Australia’s trading partners on the nations economy, one possible explanation for this result is that the Australian economy is highly influenced by the economic performances of countries within the Asian region rather than the US economy. The fact that these two crises have such diverse impacts on the economy further highlights the importance of allowing for stochastic volatility in the models error structure.

The main empirical result from our analysis is that with the exception of economic contractions, there appears to be significant co-movement among the Australian States and Territories. That being said, over the past two decades, the degree of synchronisation across Australian States has decreased, on average, by about half. This result was supported by the State’s pairwise correlation and historical decomposition of the States final demand, which indicate strong heterogeneity among the Australian State’s business cycles. This implies that national or federal policymakers should monitor, or emphasise, each individual State’s economic performance instead of the combined performance of the all States when implementing a macroeconomic policy. Thus, policy measures that are designed based on national interest will be ineffective or even counter-productive to each State’s economy. In addition, the results indicate that the federal
government should implore each State government to pursue an active role in managing their own economy, since idiosyncratic factors are the main driver of the majority of the State’s cyclical fluctuations.

There are several questions that our paper has left unanswered. The pairwise correlations show that synchronicity varies among each State. It would be intriguing to discover an explanation towards why some States share a higher correlation whilst others have a lower correlation. Furthermore, it will be interesting to examine the relationship between Australia’s major trading partner business cycles and each States business cycle, and whether there is synchronicity and commonality between the States and their trading partners’ economic performances.
Figure 4.1: Trace plots of selected parameters: (a) $\theta_{10}$, (b) $\theta_{550}$, (c) $\theta_{1000}$, (d) $\omega_3^2$, (e) $\omega_7^2$, (f) $(\Sigma_u)_{20}$ and (g) $h_{60}$.
Figure 4.2: Posterior median (blue line), 16th (red line) and 84th (red line) percentiles of the common factor stochastic volatility (expressed as standard deviation).
Figure 4.3: Posterior median (blue line), 16th (red line) and 84th (red line) percentiles of the stochastic volatility (expressed as standard deviation) from AR-SV(1) model.
Posterior median (blue line) and 68% Bayesian credible interval (shaded area)

Figure 4.4: Plot of common indicator over time
Figure 4.5: Plot of the posterior median common indicator and the OECD composite leading indicator (CLI) over time
Figure 4.6: Plot of variable-specific indices over time

Posterior median (blue line) and 68% Bayesian credible interval (shaded area)
Posterior median (blue line) and 68% Bayesian credible interval (shaded area)

Figure 4.7: Plot of State Indices over time
Figure 4.8: Plot of pairwise rolling correlations between the State factors
Figure 4.9: Plot of pairwise rolling correlations between the State factors
Figure 4.10: Plot of pairwise rolling correlations between the State factors
Figure 4.11: Plot of average correlations between a State indicator with all others
Figure 4.12: Plot of Historical Decomposition
Figure 4.13: Plot of Historical Decomposition
4.7 Appendix

4.7.1 Model Example

In this section we illustrate a simple example of the structure of the matrices Ξ's.

Suppose there are \( G = 2 \) variables, \( N = 2 \) countries/States and \( p = 1 \). Then from (4.1) we have

\[
\begin{bmatrix}
  y_t^1 \\
  x_t^1 \\
  y_t^2 \\
  x_t^2
\end{bmatrix} =
\begin{bmatrix}
  a_{11,t} & a_{12,t} & a_{13,t} & a_{14,t} \\
  a_{21,t} & a_{22,t} & a_{23,t} & a_{24,t} \\
  a_{31,t} & a_{32,t} & a_{33,t} & a_{34,t} \\
  a_{41,t} & a_{42,t} & a_{43,t} & a_{44,t}
\end{bmatrix}
\begin{bmatrix}
  y_{t-1}^1 \\
  x_{t-1}^1 \\
  y_{t-1}^2 \\
  x_{t-1}^2
\end{bmatrix}
+ \begin{bmatrix}
  c_t^x \\
  c_t^y \\
  c_t^z \\
  c_t^c
\end{bmatrix} + u_t,
\]

(4.11)

Here \( \beta_t = [a_{11,t}, a_{12,t}, a_{13,t}, a_{14,t}, c_t^x, a_{21,t}, a_{22,t}, a_{23,t}, a_{24,t}, c_t^y, a_{31,t}, a_{32,t}, a_{33,t}, a_{34,t}, c_t^z, a_{41,t}, a_{42,t}, a_{43,t}, a_{44,t}] \) is \( 20 \times 1 \). Thus, the factorisation in (4.3) is

\[
\beta_t = \begin{bmatrix}
  1 \\
  1 \\
  \vdots \\
  1
\end{bmatrix} \theta_{1t} +
\begin{bmatrix}
  \nu_1 \\
  0
\end{bmatrix} \theta_{2t} +
\begin{bmatrix}
  \nu_3 \\
  \nu_4
\end{bmatrix} \theta_{3t},
\]

(4.12)

where \( \nu_1 = [1, 1, 0, 0]' \), \( \nu_2 = [0, 0, 1, 1]' \), \( \nu_3 = [1, 0, 1, 0]' \) and \( \nu_4 = [0, 1, 0, 1]' \).

Substituting (4.12) into (4.2) and we can rewrite (4.11) as

\[
\begin{bmatrix}
  y_t^1 \\
  x_t^1 \\
  y_t^2 \\
  x_t^2
\end{bmatrix} =
\begin{bmatrix}
  Z_{1,t} \\
  Z_{1,t} \\
  Z_{1,t} \\
  Z_{1,t}
\end{bmatrix}
\begin{bmatrix}
  \theta_{1t} +
  \begin{bmatrix}
  \nu_1 \\
  0
\end{bmatrix} \theta_{2t} +
  \begin{bmatrix}
  \nu_3 \\
  \nu_4
\end{bmatrix} \theta_{3t} + u_t,
\end{bmatrix}
\]

(4.13)

where \( Z_{1,t} = y_{t-1}^1 + x_{t-1}^1 + y_{t-1}^2 + x_{t-1}^2 + 1, Z_{2,t} = y_{t-1}^1 + x_{t-1}^1, Z_{3,t} = y_{t-1}^2 + x_{t-1}^2, Z_{1,t} = y_{t-1}^1 + y_{t-1}^2 \) and \( Z_{3,t} = x_{t-1}^1 + x_{t-1}^2 \). Ciccarelli, Ortega and Valderrama (2015) note there are several important differences between our model (4.5) and standard factor models. First, the indices derived in this model weight the information in all variables
equally while in the factor models the weights generally depend on the variability of the components. Second, these indices dynamically span the lagged interdependencies across units (countries/states) and variables. In contrast, standard factor models statistically span the variables of the system. Third, these indices are directly observable while in the factor models they are estimated. In addition, they are correlated by construction because the factorisation is applied on the coefficient vector rather than the variables. Lastly, the averaging approach in this model creates a moving average in terms of the order $p$ in the regressors of (4.5), even when $y_{it}$ are serially independent. This means the indices implicitly filter out from the right hand side variables of the VAR high frequency variability. Canova et al (2007) note the fact that the regressors of the SUR model captures the low frequencies movements in the variables of the VAR is important in forecasting in the medium term and in detecting turning points of GDP growth.

### 4.8 Gibbs Sampler

To simulate the posterior distribution, we use a six block Gibbs Sampler that sequentially draws from each conditional posterior distribution. The outline of the steps are:

1. Draw from $p(\theta^{(i)} \mid y, h^{(i)}, \Sigma_u^{(i-1)}, \Omega^{(i-1)}, \sigma_h^{2(i-1)}, \rho^{(i-1)})$
2. Draw from $p(\Sigma_u^{(i)} \mid y, h^{(i)}, \theta^{(i-1)}, \Omega^{(i-1)}, \sigma_h^{2(i-1)}, \rho^{(i-1)})$
3. Draw from $p(h^{(i)} \mid y, \Sigma_u^{(i-1)}, \theta^{(i-1)}, \sigma_h^{2(i-1)}, \Omega^{(i-1)}, \rho^{(i-1)})$
4. Draw from $p(\Omega^{(i)} \mid y, h^{(i)}, \Sigma_u^{(i-1)}, \theta^{(i-1)}, \sigma_h^{2(i-1)}, \rho^{(i-1)})$
5. Draw from $p(\rho^{(i)} \mid y, h^{(i)}, \Sigma_u^{(i-1)}, \theta^{(i-1)}, \Omega^{(i-1)}, \sigma_h^{2(i-1)})$
6. Draw from $p(\sigma_h^{2(i)} \mid y, h^{(i)}, \Sigma_u^{(i-1)}, \theta^{(i-1)}, \Omega^{(i-1)}, \rho^{(i-1)})$
7. Repeat step 1 to 6.
where the superscript denotes the $i$-th draw of the simulation. Canova and Ciccarelli (2009) use standard Kalman filtering and smoothing techniques from Chib and Greenberg (1995) to estimate the time-varying coefficients. However, we adopt a different method: for the draws of Step 1 we use the algorithm derived from Chan and Jeliazkov (2009).

### 4.8.1 Step 1 Drawing $\theta_t$

The measurement equation of (4.5) can be rewritten in the form:

$$y = Z\theta + u, \quad u \sim N(0, \Sigma), \quad (4.14)$$

where $u = (u_1, \ldots, u_T)'$, $y = (Y_1, \ldots, Y_T)'$, $\theta = (\theta_1, \ldots, \theta_T)'$, $\Sigma = \text{diag}(e^{h_1\Sigma_u}, \ldots, e^{h_T\Sigma_u})$ and

$$Z = \begin{bmatrix} Z_1 & 0 & \cdots & \cdots & 0 \\ 0 & Z_2 & 0 & \cdots & 0 \\ 0 & \ddots & Z_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & Z_T \end{bmatrix},$$

Next the transition equation of (4.4) can be rewritten:

$$H\theta = \tilde{\theta}_0 + \eta, \quad (4.15)$$

where $\tilde{\theta}_0 = (\theta_0', 0, \ldots, 0)$, $S_\theta = \text{diag}(V_\theta, \Omega, \ldots \Omega)$ and

$$H = \begin{bmatrix} I_m & 0 & 0 & \cdots & 0 \\ -I_m & I_m & 0 & \cdots & 0 \\ 0 & -I_m & I_m & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -I_m & I_m \end{bmatrix},$$

148
Therefore $\theta \sim N(\hat{\theta}, (H' S_{\theta}^{-1} H)^{-1})$, where $\hat{\theta} = H^{-1} \theta_0$. $H$ is a band matrix with a determinant of $|H| = 1$.

The conditional posterior distribution is:

$$p(\theta|y, \Sigma_u, \Omega, \sigma^2_h, \rho, h) \propto p(y|\theta, \Sigma_u, \Omega, \sigma^2_h, \rho, h)p(\theta),$$

(4.16)

$$\propto |\Sigma|^{-\frac{1}{2}} \exp -\frac{1}{2}[(y - Z\theta)' \Sigma^{-1} (y - Z\theta)] \exp -\frac{1}{2}[(\theta - \hat{\theta})' H' S_{\theta}^{-1} H (\theta - \hat{\theta})],$$

(4.17)

$$\propto \exp -\frac{1}{2} [\theta'(Z' \Sigma^{-1} Z + H' S_{\theta}^{-1} H) \theta - 2\theta'(Z' \Sigma^{-1} y + H' S_{\theta}^{-1} H \hat{\theta})],$$

(4.18)

Using the standard results from linear regression

$$p(\theta|y, \Sigma_u, \Omega, \sigma^2_h, \rho, h) \sim N(\bar{\theta}, \Theta^{-1} \theta),$$

(4.19)

where

$$\Theta_{\theta} = Z' \Sigma^{-1} Z + H' S_{\theta}^{-1} H, \quad \bar{\theta} = \Theta^{-1}_{\theta}(Z' \Sigma^{-1} y + H' S_{\theta}^{-1} H \hat{\theta}).$$

(4.20)

Note: since for our priors we assumed $\theta_0 = 0$ then

$$\Theta_{\theta} = Z' \Sigma^{-1} Z + H' S_{\theta}^{-1} H, \quad \bar{\theta} = \Theta^{-1}_{\theta}(Z' \Sigma^{-1} y).$$

(4.21)

Since $H$ is a band matrix, this implies that the precision matrix $\Theta_{\theta}$ is also a band matrix. Thus, this means that $\bar{\theta}$ can be drawn efficiently by solving the linear system

$$\Theta_{\theta} x = Z' \Sigma^{-1} y,$$

(4.22)

for $x$, which avoids computing the inverse $\Theta_{\theta}^{-1}$. To draw from $N(\bar{\theta}, \Theta_{\theta})$, we use the algorithm from Chan and Jeliazkov (2009), that is, we first take the Cholesky factor of $\Theta_{\theta}$ which is $\Theta_{\theta} = C_{\theta} C_{\theta}'$. Next we obtain $Tm$ independent draws from a standard nor-
mal distribution $N(0, 1)$ denoted as $N = (N_1, \ldots, N_{T_m})'$ and return $\theta = \bar{\theta} + (C\theta')^{-1}N$.

It is easy to check that the mean $\theta$ is $\bar{\theta}$ and its covariance matrix is

$$(C\theta)^{-1}((C\theta')^{-1})' = (C\theta)^{-1}(C\theta)^{-1} = (C\thetaC\theta')^{-1} = \Theta\theta^{-1}. \tag{4.23}$$

This precision sampler technique from Chan and Jeliazkov (2009) has a clear computation efficiency advantage over the Kalman filtering techniques.

4.8.2 Step 2 Drawing $\Sigma_u$

The conditional posterior is

$$p(\Sigma_u|y, h, \theta, \sigma_h^2, \Omega, \rho) \propto p(y|\theta, h, \Sigma_u, \Omega, \sigma_h^2, \rho)p(\Sigma_u), \tag{4.24}$$

$$\propto \prod_t e^{\frac{1}{2}h_t\sum_{u}|\sum_{u}^{-\frac{1}{2}2}e^{\frac{1}{2}h_t\sum_{u}^{-1}}(Y_t - Z_t\theta_t)' e^{\frac{1}{2}h_t\sum_{u}^{-1}}(Y_t - Z_t\theta_t)}\prod_t e^{\frac{1}{2}tr(Q_1\sum_{u}^{-1})}, \tag{4.25}$$

$$\propto |\sum_{u}|^{-\frac{T+z_1+p+1}{2}}\exp\left(-\frac{1}{2}tr(Q_1\sum_{u}^{-1}) + \sum_t (Y_t - Z_t\theta_t)'(Y_t - Z_t\theta_t)e^{\frac{1}{2}h_t\sum_{u}^{-1}}(Y_t - Z_t\theta_t)'\right), \tag{4.26}$$

$$p(\Sigma_u|y, h, \theta, \sigma_h^2, \Omega, \rho) \sim \text{IW}(z_1 + T, \sum_t (Y_t - Z_t\theta_t)'(Y_t - Z_t\theta_t)e^{\frac{1}{2}h_t\sum_{u}^{-1}}(Y_t - Z_t\theta_t)' + Q_1), \tag{4.27}$$

4.8.3 Step 3 Drawing $h$

First we rearrange (4.5) into

$$P^{-1}(Y_t - X_t\Xi\theta_t) = e^{\frac{1}{2}h_t}\epsilon_t, \epsilon_t \sim N(0, I_n), \tag{4.28}$$

Note $\Sigma_u = PP'$ and $P$ is a lower triangular matrix of the Cholesky factor of $\Sigma_u$. We can square both sides of (4.28) and take the logarithm:
\[ y_t^* = i_n h_t + \epsilon_t^*, \quad (4.29) \]

where, \( i_n = [1, 1, \ldots, 1]' \) is an \( n \times 1 \) vector, \( y_t^* = \log((P^{-1}(Y_t - X_t \Xi \theta_t))^2 + c) \), \( c \) is some small constant and \( \epsilon_t^* = [\log(\epsilon_{1,t}^2), \ldots, \log(\epsilon_{n,t}^2)]' \) follows a \( \log - \chi_1^2 \) distribution.

To draw the common stochastic volatility factor we implement the precision sampler technique by Chan and Hsiao (2014) and follow their procedure whereby they implement the Kim, Shepherd and Chib (1998) auxiliary mixture sampler in approximating the \( \log - \chi_1^2 \) distribution using a seven component Gaussian mixture density with fixed parameters as shown in Table 4.

Table 4.4: A Seven Component Gaussian Mixture for Approximating the \( \log - \chi_1^2 \) distribution

<table>
<thead>
<tr>
<th>Component</th>
<th>( p_j )</th>
<th>( \mu_j )</th>
<th>( \sigma_j^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00730</td>
<td>-10.12999</td>
<td>3.79596</td>
</tr>
<tr>
<td>2</td>
<td>0.10556</td>
<td>-3.97281</td>
<td>2.61369</td>
</tr>
<tr>
<td>3</td>
<td>0.0002</td>
<td>-8.56686</td>
<td>3.17950</td>
</tr>
<tr>
<td>4</td>
<td>0.04395</td>
<td>2.77786</td>
<td>0.16735</td>
</tr>
<tr>
<td>5</td>
<td>0.34001</td>
<td>0.61942</td>
<td>0.64009</td>
</tr>
<tr>
<td>6</td>
<td>0.24566</td>
<td>1.179518</td>
<td>0.34023</td>
</tr>
<tr>
<td>7</td>
<td>0.25750</td>
<td>-1.08819</td>
<td>1.26261</td>
</tr>
</tbody>
</table>

Source: Chan and Hsiao (2014)

More specifically, Chan and Hsiao (2014) noted that

\[ f(\epsilon_{i,t}^*) \approx \sum_{j=1}^{7} p_j \varphi(\epsilon_{j,t}^*; \mu_j - 1.2704, \sigma_j^2), \quad (4.30) \]

where \( \varphi(\epsilon_{j,t}^*; \mu_j, \sigma_j^2) \) is the Gaussian density with \( \mu_j \) and variance \( \sigma_j^2 \) and \( p_i \) is the probability of the \( j^{\text{th}} \) mixture component. Chan and Hsiao (2014) emphasise that these parameter values are fixed and do not depend on any unknown parameters. Equivalently (4.30) can be written in terms of an auxiliary random variables \( s_t \in \{1, \ldots, 7\} \) that serves as the mixture component indicator for each point at time such as

\[ (\epsilon_{j,t}^* | s_t = j) \sim N(\mu_j - 1.2704, \sigma_j^2), \quad (4.31) \]

\[ \mathbb{P}(s_t = j) = p_j. \quad (4.32) \]

\(^{\text{Note } n = NG \text{ is the total number of endogenous variables in the model.}}\)
Under this representation, the model is now a linear Gaussian model conditional on the component indicator $s_t$. Chan and Hsiao (2014) applied this auxiliary mixture sampling approach to a univariate case. However, in our study we have to apply the auxiliary mixture sampling approach for each endogenous variable. The procedure is outlined below:

1. We apply the auxiliary mixture sampler for each $i = 1, \ldots, n$ endogenous variable and we draw both $s^i = (s^i_1, \ldots, s^i_T)'$ and $\sigma^2_{s^i} = (\sigma^2_{s^i_1}, \ldots, \sigma^2_{s^i_T})'$ respectively.

2. Once we have applied the auxiliary mixture sampler for all the endogenous variables in the model, we can stack up (4.29) to time $T$.

$$y^* = X_h h + \epsilon^*,$$  \hspace{1cm} (4.33)

$$\epsilon^* \sim N(d_s, \Sigma^*_y),$$  \hspace{1cm} (4.34)

where both $y^* = (y^*_1, \ldots, y^*_T)'$ and $\epsilon^* = (\epsilon^*_1, \ldots, \epsilon^*_T)'$ are $Tn \times 1$ vectors, $h = (h_1, \ldots, h_T)'$ is a $T \times 1$ vector and $X_h = I_T \otimes i_n$ is a $Tn \times T$ matrix. $d_s = (\mu_{s^1_1} - 1.2704, \mu_{s^2_1} - 1.2704, \ldots, \mu_{s^1_T} - 1.2704, \mu_{s^2_T} - 1.2704, \ldots, \mu_{s^1_T} - 1.2704, \mu_{s^2_T} - 1.2704)'$ is a $Tn \times 1$ vector and $\Sigma^*_y = diag(\sigma^2_{s^1_1}, \ldots, \sigma^2_{s^2_1}, \ldots, \sigma^2_{s^1_T}, \ldots, \sigma^2_{s^2_T})$ is a $Tn \times Tn$ matrix.

3. Next, using the (4.33) and (4.34) we can derive the log likelihood

$$\log p(y^* \mid s, h) = -\frac{1}{2}[(y^* - X_h h - d_s)' \Sigma^{-1}_{y^*} (y^* - X_h h - d_s)] + c_1.$$

We can rewrite (4.7) into matrix form

$$H_h h = \tilde{\alpha}_h + \xi, \quad \xi \sim N(0, \Phi),$$  \hspace{1cm} (4.36)

where $\tilde{\alpha}_h = (h_0, 0, \ldots, 0)$, $\Phi = diag(\frac{\sigma^2_{h^2}}{1-p^2}, \sigma^2_{h}, \ldots, \sigma^2_{h})$ and
Thus \((h | \Phi, \tilde{\alpha}_h) \sim N(\alpha_h, (H'_h \Phi^{-1}H_h)^{-1})\), where \(\alpha_h = H_h^{-1}\tilde{\alpha}_h\). \(H_h\) is a band matrix with a determinant \(|H_h| = 1\) for all values of \(\rho\). From (4.35) and (4.36) we can derive the conditional posterior distribution of \(p(h|y, \theta, \Omega, \Sigma, \sigma^2_h, \rho)\)

\[
\alpha - \frac{1}{2}[(y^* - X_h h - d_s)' \Sigma^{-1}_{y^*} (y^* - X_h h - d_s)] - \frac{1}{2}[(h - \alpha_h)'(H'_h \Phi^{-1}H_h)(h - \alpha_h)],
\]

(4.37)

Therefore

\[
K_h = H'_h \Phi^{-1}H_h + X'_h \Sigma^{-1}_{y^*}X_h, \quad \hat{h} = K_h^{-1}(H'_h \Phi^{-1}H_h \alpha_h + X'_h \Sigma^{-1}_{y^*}(y^* - d_s)),
\]

(4.38)

\[
p(h|y, \theta, \Omega, \Sigma, \sigma^2_h, \rho) \sim N(\hat{h}, K_h^{-1}).
\]

(4.39)

Notice that here again the precision matrix \(K_h^{-1}\) is also a band matrix, which means we can apply the same precision sampler technique as discussed in step 1 to draw \(\hat{h}\).

**4.8.4 Step 4 Drawing \(\Omega\)**

The elements of \(\Omega\) are conditionally independent and follow an inverse-gamma distribution:

\[
(\omega^2_i) \sim IG(T/2 + \omega_0, \sum_{t=2}^{T} (\theta_i^t - \theta_i^{t-1})^2 + S_0), \quad \text{for } i = 1, \ldots, m.
\]

(4.40)
4.8.5 Step 5 Drawing $\rho$ 

\[ p(\rho | y, \theta, \Omega, h, \sigma_h^2, \Sigma_u) \propto p(\rho) g(\rho) \exp \left( -\frac{1}{2\sigma_h^2} \sum_{t=2}^{T} (h_t - \rho h_{t-1})^2 \right), \]  

(4.41)

where \( g(\rho) = (1 - \rho^2)^{\frac{1}{2}} \exp \left( -\frac{1}{2\sigma_h^2} (1 - \rho^2) (h_1 - h_0)^2 \right) \) and \( p(\rho) \) is the truncated normal given in (4.8). The conditional posterior density \( p(\rho | y, \theta, \Omega, h, \sigma_h^2, \Sigma_u) \) is non-standard, which means a Metropolis-Hastings step has to be undertaken to draw \( \rho \). We follow the methodology governed in Chan and Hsiao (2014) and implement an independence-chain Metropolis-Hasting step. Please see Chan and Hsiao (2014) for further details about the algorithm.

4.8.6 Step 6 Drawing $\sigma_h^2$

The conditional posterior for $\sigma_h^2$ follows an inverse-gamma distribution:

\[ p(\sigma_h^2 | y, \theta, \Omega, h, \rho, \Sigma_u) \sim IG(w_1 + \frac{T}{2}, \tilde{S}_1), \]  

(4.42)

where \( \tilde{S}_1 = S_1 + [(1 - \rho^2)(h_1)^2 + \sum_{t=2}^{T} (h_t - \rho h_{t-1})^2]/2. \)
5 Chapter 5

5.1 Conclusion

The main objective of this thesis is to examine three applications of different model specifications within the TVP-VAR framework. Firstly, in Chapter 2 we determine whether the propagation and transmission mechanism of Malaysian monetary policy differed during the Asian Financial Crisis of 1997/98 and the Global Financial Crisis of 2007/08 using a standard TVP-VAR with stochastic volatility model from Primiceri (2005). The main result we found is that despite having no evidence of time-variation within the propagation mechanism of Malaysian monetary policy the average contribution of a monetary policy shock to the variability of each macroeconomic variable: Real GDP, Inflation and the Nominal Effective Exchange Rate, differs between the two crises. This finding suggests that despite the propagation mechanism being relatively constant, Malaysia’s monetary policy transmission mechanism evolves over time. The finding that the main mechanism driving the evolution of the transmission mechanism is the error variance-covariances matrix of the model, not the VAR coefficients, is consistent with Chan and Eisenstat (2016) and Primiceri (2005) who examine the US economy. To elicit this insight we then conducted a formal model comparison using the Bayesian DIC measure for four competing models: the TVP-VAR-SV, a VAR-SV, a TVP-VAR and a VAR. The results showed that the constant parameter VAR with stochastic volatility (VAR-SV) is the preferred model or the best in sample fit out of the four models. This result further supports our argument above that the main source of time-variation in our model is through the variance-covariance matrix of the shocks. Also, we found some evidence that the implementation of capital controls reduced the influenceability of monetary policy on the Malaysian economy. This result contradicts the argument put forward by Athukorala and Jongwanich (2012) that the imposition of capital controls allowed the BNM to regain monetary policy autonomy and enable them to pursue expansionary policies to reflate the Malaysian economy.

Secondly, Chapter 3 investigates whether incorporating time variation and fat-tails into a class of popular univariate and multivariate Gaussian distributed models can im-
prove the forecast performance of key Australian macroeconomic variables: Real GDP growth, CPI Inflation and a short-term interest rate. We found four important results. First, fat-tailed models consistently outperform their Gaussian counterparts. Second, time varying parameters and stochastic volatility improves forecast performance across all variables relative to a constant parameter benchmark. Third, stochastic volatility models under a Student’s-t distribution are found to generate more accurate density forecasts as compared to the same models under a Gaussian specification. Taken together these results suggest that both structural instabilities and fat-tail events are important features in modeling Australian macroeconomic variables. Finally, when comparing the forecast accuracy of univariate and multivariate models the simple rolling window autoregression with fat-tails produces the most accurate output growth forecasts, whilst the time varying parameter vector autoregression with stochastic volatility and fat-tails produces the best interest and inflation forecasts.

Finally, Chapter 4 estimates a time-varying parameter Panel BVAR with a new feature, a common stochastic volatility factor in the error structure to assess the synchronicity and the nature of Australian State business cycles. The adoption of a common stochastic volatility factor is crucial since there have been many studies undertaken in the literature that have highlighted the importance of the addition of stochastic volatility to the error structure in improving model fit and forecastability (for instance see Clark (2014), Clark and Ravazzolo (2015), and Chan and Eisenstat (2016)). From our results, we show that the inclusion of the common stochastic volatility factor to the model is important since it shows that volatility or uncertainty on the Australian economy was more pronounced during the Asian Financial Crisis rather than the recent Global Financial Crisis. This result is plausible since technically the Australian economy did not experience a recession during the recent crisis period and Australia’s heavily reliance of trading partners within the Asian region. We also found that the common indicator reveals some interesting economic facts. It appears to capture the early 1990’s recession and slowdown that the Australian economy experienced during the the GFC, which suggests there is commonality across each Australian State during
a contraction. Also, we found that the fluctuations of the common indicator closely follows the trend line of the OECD CLIs for Australia, especially during the 2000’s period. This means that the common indicator appropriately captures majority of the fluctuations in economic activity for our sample period. In terms of the synchronicity of Australian State business cycles, we found on average that the degree of synchronisation across the States has decreased to about half in terms of correlation from the 1990’s to 2000’s. Therefore, there is evidence of heterogeneity present within each State’s business cycle.

5.2 Future research

A question that is left unanswered in Chapter 2 is Malaysia’s monetary policy rule in regards to unexpected shocks to real GDP, inflation and the Nominal Effective Exchange Rate. To investigate this issue further, one must fully identify the impact/contemporaneous matrix. One potential avenue for this research agenda is to follow Ellis, Mumtaz and Zabczyk (2014) and utilise a Dynamic Stochastic General Equilibrium (DSGE) model, simulate the impulse responses, and use these responses as motivating restrictions for the impact/contemporaneous matrix. In order for this agenda to begin, further research first needs to be undertaken in regards to the deep parameters of the Malaysian economy. In regards to Chapter 3, we note that we have only provided an out of sample study of the proposed modeling features. For future research it would be useful to analyze the in-sample fit by incorporating structural instabilities and fat-tails into general equilibrium models of the Australian economy. For instance, the New Keynesian model of Australia developed by Jääskelä and Nimark (2011) could be extended by allowing for time varying Student’s-t distributed disturbances within both aggregate demand and supply shocks. Finally, for Chapter 4 there are several questions that our study has left unanswered. The pairwise correlations show that synchronicity varies among each State. It would be intriguing to discover an explanation towards why some States share a higher correlation together whilst others have a lower correlation together. Furthermore, it will be interesting to examine the
relationship between Australia’s major trading partner business cycles and each States business cycle, and whether there is synchronicity and commonality between the States and their trading partners economic performances.
6 Bibliography

References


