Tasmanian School of Business and Economics
University of Tasmania

Discussion Paper Series N 2020-12

Real-time Forecasting of the Australian Macroeconomy Using Flexible Bayesian VARs

Bo Zhang
Shanghai University, China

Bao H. Nguyen
University of Tasmania, Australia

ISBN 978-1-922352-75-0

Real-time Forecasting of the Australian Macroeconomy Using Flexible Bayesian VARs*

Bo Zhang† Bao H. Nguyen‡

October 23, 2020

Abstract

This paper evaluates the real-time forecast performance of alternative Bayesian Vector Autoregressive (VAR) models for the Australian macroeconomy. To this end, we construct an updated vintage database and estimate a set of model specifications with different covariance structures. The results suggest that a large VAR model with 20 variables tends to outperform a small VAR model when forecasting GDP growth, CPI inflation and unemployment rate. We find consistent evidence that the models with more flexible error covariance structures forecast GDP growth and inflation better than the standard VAR, while the standard VAR does better than its counterparts for unemployment rate. The results are robust under alternative priors and when the data includes the early stage of the COVID-19 crisis.

JEL-codes: C11, C32, C53, C55

Keywords: Australia, real-time forecast, Non-Gaussian, Stochastic Volatility.

*We thank Joshua Chan, Jamie Cross, Chenghan Hou and Trung-Duc Tran for their valuable discussions. We also benefited from comments of seminar participants at Tasmanian School of Business and Economics.

†School of Economics, Shanghai University, Shanghai, 200444 P. R. China. Email: bozhangyc@gmail.com

‡*Corresponding author*: Tasmanian School of Business and Economics, University of Tasmania and Centre for Applied Macroeconomic Analysis (CAMA), Australian National University, Australia. Email: b.nguyen@utas.edu.au

1 Introduction

Forecasting key macroeconomic variables, such as output growth, inflation and unemployment is an important but difficult task for policy makers. To obtain accurate and timely forecasts of such indicators, forecasters have to deal with uncertainty around forecasting models and data. The accurate forecasts would only be obtained under a model that is able to capture the salient feature of macroeconomic data. Producing timely forecasts, on the other hand, requires a forecasting approach that can handle with data that have been just released and subject to revisions in the future.

In this paper, we take the aforementioned issues into account. We contribute to the literature by, for the first time, investigating real-time data for Australia. Much of literature has looked at forecasting using the latest available vintage of data at the time of investigation. However, as highlighted in Clements and Galvão (2013) and Clements (2017), among others, using truncated series from such a single vintage would lead to an inaccurate assessment. One of the main reasons is that most macroeconomic variables are subject to data revisions and these revisions are often not small and random. Using a single vintage of data implies that the data used in model estimation have been revised many times, while the forecast is conditioned on data that have been just released. As a result, data revisions have a major impact on forecasts (Croushore, 2011a,b). To minimize these potential forecasting problems, we employ a real-time database that includes all possible data vintages for the Australian macroeconomy. To that end, we collect data vintages from various sources and construct an updated and comprehensive real-time dataset of key macroeconomic variables for Australia.[1]

In addition to constructing and utilizing real-time data, we also exploit the usefulness of non-standard Vector autoregressive (VARs) in the context of the Australian macroeconomy. VARs have been a successful tool in the forecasting literature since the mid-1980s. Starting with the early work by Doan et al. (1984) and Litterman (1986) which exploited Bayesian methods and focused on VARs with a small number of dependent variables. Because VARs tend to have a great number of parameters, Bayesian approach offers a formal way to shrinking parameters and improves forecast performance. The family of priors that they used is commonly called Minnesota prior and one of the most popular priors in the Bayesian VAR (BVAR) literature. This approach was followed by the seminal work of Bańbura et al. (2010), who considered larger BVAR models with more than 20 variables. With a slight modification of the Minnesota prior, Bańbura et al. (2010) found that large BVAR models even forecast better than small BVARs and factor models. Similar conclusions are also found in Carriero et al. (2009), Koop (2013) and Carriero

[1]This dataset is available at https://sites.google.com/site/nguyenhoaibao.

et al. (2015). Recently, a variety of extensions of the standard VAR with conventional assumptions of error disturbances (e.g., homoscedastic, Gaussian and serially independent) has been proposed, including alternative model specifications that can feature flexible covariance structures (Cross et al., 2020; Chan, 2020a,b). For example, in the context of large BVARs and using US macroeconomic data, Chan (2020b) shows that one can further improve the forecast performance of VARs by replacing the standard covariance structure with a more flexible structure, such as non-Gaussian, heteroscedastic, and serially dependent innovations. These extensions are crucial because they can take into account salient features of macroeconomic time series and thus enhance the forecasting power of BVARs (Carriero et al., 2015; Clark and Ravazzolo, 2015).

In light of this emerging literature, we evaluate the forecast performance of a set of small and large BVAR models for the Australian economy. To this end, we first consider a small VAR with three core macroeconomic variables, including GDP growth, CPI inflation and unemployment rate as a benchmark model. We then compare the forecast performance of this benchmark model with those associated with a larger VAR model. As in Chan (2020a), we consider a set of VAR models combining three error covariance structures: common stochastic volatility, serial dependence moving error and t innovations. While these features are found to be important in forecasting for many economies (Chan, 2020b; Zhang et al., 2020), evidence for the Australian economy has been limited. Recent work by Cross and Poon (2016) considers a wide range of univariate and small multivariate models and points out that models with heavy-tailed error distributions, such as the t distribution, provide the most accurate forecasts for Australian GDP. With this idea in mind, along with the t distribution, we also consider heteroscedastic and serially dependent errors. Our approach therefore takes into account all possible combinations of non-standard error assumptions. In addition, other than a small VAR considered in Cross and Poon (2016), this paper evaluates the forecast performance of a relative larger model with 20 variables. The number of variables considered in this paper is motivated by recent evidence found by Panagiotelis et al. (2019), who show that a VAR model that is not beyond 20 variables tends to generate more accurate macroeconomic forecasts for Australia.

Our out-of-sample forecasting experiment delivers the following results. We find that a large VAR model with 20 variables tends to outperform a small VAR model when forecasting GDP growth, CPI inflation and unemployment rate. Specifically, we find consistent evidence that the models with more flexible error covariance structures forecast GDP growth and inflation better than the standard VAR, whereas the standard VAR using non-standard priors does better than its counterparts for unemployment rate. These findings are found to remain unchanged under alternative priors and when the data includes the early stage of the Covid-19 crisis.

The rest of the paper is organized as follows. In Section 2 we describe how we collect the real-time dataset and their sources. We then introduce in Section 3 alternative covariance structures that incorporate into BVARs. Forecast results and discussion are presented in Section 4. We report the sensitivity analysis in Section 5 and Section 6 concludes the paper.

2 Data

In this paper, we use a dataset that includes a variety of standard macroeconomic and financial variables, such as GDP and its components, prices, unemployment and money supply. These variables are similar to variables that commonly include in a large-size VARs in the macroeconomics forecasting literature (Bańbura et al., 2010; Koop, 2013; Chan, 2020a). To capture the fact that the Australian economy is a small open economy and relies on commodity resources, we also include a real exchange rate measure, terms of trade and commodity prices. While Eickmeier and Ng (2011) find that adding international predictors can improve forecast for New Zealand GDP, Panagiotelis et al. (2019) and Bjørnland et al. (2017) highlight that such predictors do not add much value to predicting GDP growth for Australia. With this idea in mind, we only consider a medium number of predictors that consists of 20 variables and runs from 1982Q3 to 2020Q1. Series which are originally observed at a monthly frequency are transformed to quarterly by averaging over the 3 months in a quarter. Table A1 in Appendix A provides a brief description of each variable, along with the methods of transformation.

Except for financial variables that are not subject to revision, we use real-time data for the remaining variables. Data vintages before 2017Q1 are taken from the *Australian Real-Time Macroeconomics Database* maintained by the University of Melbourne.[2] These data vintages are collated from various sources, which are originally published by the Australian Bureau of Statistics (ABS) and the Reserve Bank of Australia (RBA). The construction of this database is described in Lee et al. (2012) and the reader is referred to that paper for further details about the data. To update the dataset, we collected data vintages from 2017Q2 to the most recent release from ABS and RBA website. The real-time data used in this paper therefore consists of vintages for 1995Q1 through 2020Q2, each covering data extending back to 1982Q3. The starting date of 1995Q1 for the first vintage was chosen because data for some of the variables of interest are only available

[2]The database is publicly available for download at https://fbe.unimelb.edu.au/economics/macrocentre/artmdatabase#databases-and-documentation

until 1982.[3] Starting with vintage 1995Q1 makes our first sample of evaluation period long enough to allow reasonable estimation inference. The models are first estimated with data from 1982Q3 to 1995Q1, and then recursively estimated with expending sample windows starting in 1982Q4 and ending in 1995Q2, 1995Q3, ..., 2019Q1. Due to reporting lags, the real-time data vintage released at time t contains observations only up to time $t-1$. We report results for horizons of current quarter nowcasts as well as one-quarter-ahead, two-quarter-ahead and one-year-ahead forecasts.

Aside from the evaluation period, another issue is what vintage to be taken as actual data in calculating forecast errors. In real-time forecasting literature, either the first release following the forecast date or the most recent vintage can be used. As discussed in Lee et al. (2012), the Australian real-time data has been revised multiple times for various reasons, reflecting "definitional changes" and "revisions". In this case, data released in the latest vintage is presumably closer to the underlying "true" value of the time series. Thus, we decided to take the latest vintage as actuals in evaluating forecast accuracy, as in Garratt et al. (2009), Schorfheide and Song (2015), Carriero et al. (2015) and Chan (2020a).

3 Flexible Bayesian VARs

BVARs with flexible covariance matrix assumptions are considered as the main specifications of the competing models in our forecast exercise. In the following sections, we first introduce BVARs with conventional error assumptions and then common stochastic volatility (CSV), heavy tailedness (e.g., Student's t distribution), and serial dependence moving average error (MA).

3.1 Standard VARs with Conventional Error Assumptions

We start from an expression of the standard VAR model, which can be written in a reduced form of VAR with order p as below:

$$\mathbf{y}_t = \mathbf{b} + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t^y, \quad \boldsymbol{\varepsilon}_t^y \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \tag{1}$$

where $\mathbf{y}_t = (y_{1t}, \cdots, y_{nt})'$ denote an $n \times 1$ vector of endogenous variables in a BVAR, $\mathbf{b}$ is an $n \times 1$ vector of intercepts, and $\mathbf{B}_1, \cdots, \mathbf{B}_p$ are $n \times n$ coefficient matrices, and $\boldsymbol{\Sigma}$ is an $n \times n$ cross-sectional covariance matrix of VAR. In a standard VAR, $\boldsymbol{\varepsilon}_t^y$ can be assumed

[3] For example, from the *Australian Real-Time Macroeconomics database*, the first vintage for real GDP is 1971Q3 and the sample collected begins in 1959Q3, while commodity prices for Australia is only available from 1982Q3.

to be independent and identically Gaussian distributed (iid). In practice, Equation 1 can be rewritten as below for parameter estimation:

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t^y, \tag{2}$$

where $\mathbf{X}_t = \mathbf{I}_n \otimes [1, y'_{t-1}, \cdots, y'_{t-p}]$ in which notation $\otimes$ denotes the Kronecker product, and $\boldsymbol{\beta}$ is stacked by rows of $[\mathbf{b}, \mathbf{B}_1, ..., \mathbf{B}_p]'$ with the size of $(1 + np)n \times 1$.

Let $x'_t = (1, y'_{t-1}, \cdots, y'_{t-p})$ be a $1 \times (1 + np)$ vector, when stacking the observations over time T, we get $\mathbf{X}$ which is a $T \times (1 + np)$ matrix. Then we have:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}, \tag{3}$$

where $\mathbf{Y}$ is $\mathbf{y}_t$ stacked over time T, $\mathbf{B} = (\mathbf{b}, \mathbf{B}_1, ..., \mathbf{B}_p)'$ with a size of $(1 + np) \times n$, $\mathbf{E} = (\boldsymbol{\varepsilon}_1^y, \cdots, \boldsymbol{\varepsilon}_T^y)'$, so that

$$\text{vec}(\mathbf{E}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}), \tag{4}$$

where $\boldsymbol{\Omega}$ is the serial covariance matrix of VAR model.

As mentioned, in improving model fitness and forecastability, the standard BVAR model with iid Gaussian innovations can be extended in different ways in order to capture important features of macroeconomic time series. In what follows, we introduce these extensions in details and consider those proposed models as flexible BVARs.

3.2 VARs with a Common Stochastic Volatility

One of the most useful extensions of VARs is the adoption of a common stochastic volatility (CSV) factor. There has been recognized that the volatilities of a wide ranges of macroeconomic variables are time-varying and tend to move together (Carriero et al., 2016; Mumtaz and Theodoridis, 2018; Poon, 2018). However, standard VARs with homoscedastic error, would not be able to capture this feature. The inclusion of CVS error specification allows VARs to capture any common structural shifts in the macroeconomic time series. In the modeling framework of VARs with CSV, we firstly consider time-varying volatility. Suppose $\boldsymbol{\varepsilon}_t^y \sim \mathcal{N}(\mathbf{0}, e^{h_t}\boldsymbol{\Sigma})$, where h is the stochastic volatility parameter and e^{h_t} is the common stochastic volatility (Carriero et al., 2016). More specifically, h follows an AR(1) process:

$$h_t = \phi_h h_{t-1} + \boldsymbol{\varepsilon}_t^h, \quad \boldsymbol{\varepsilon}_t^h \sim \mathcal{N}(\mathbf{0}, \sigma_h^2), \tag{5}$$

where $|\phi_h| < 1$. In this assumption, the variances of all the variables share the same stochastic volatility parameter which is a restrictive assumption. There is empirical evidence that the volatilities of macroeconomic time series have a comovement (Carriero et al., 2016), thus it is also a parsimonious assumption for parameter estimation.

3.3 VARs with a CSV and t Errors

Recent empirical studies also show that the forecast performance of macroeconomic variables can be improved when normal distribution is replaced by heavy-tailed distribution, e.g. Student's t distribution, in covariance matrix of VARs. The importance of this extension is that when the model accounts for t-disturbances, this specification of heavy-tailed innovations turns out to present good features, such as reducing the variation of estimates, dealing well with outliers, such as the Great Recession, and thus providing good model fitness (e.g., Clark and Ravazzolo, 2015; Cross and Poon, 2016; Chiu et al., 2017). In modelling VARs that can capture such fat tail events, the distribution of error terms $\boldsymbol{\varepsilon}_t^y$ has one more hyperparamter λ_t for t innovations:

$$\boldsymbol{\varepsilon}_t^y \sim \mathcal{N}(\mathbf{0}, \lambda_t e^{h_t} \boldsymbol{\Sigma}), \tag{6}$$

where $\lambda_t \sim \mathcal{IG}(\boldsymbol{\nu}_\lambda/2, \boldsymbol{\nu}_\lambda/2)$ following an inverse-gamma distribution with degree of freedom parameter $\boldsymbol{\nu}_\lambda$, and $\lambda_1, ..., \lambda_T$ are independent from each other.

3.4 VARs with a CSV and MA(1) t Errors

Another property of macroeconomic variables that has been recognized is serially dependent (Chan, 2013). To handle this property, the conventional assumption of serially independent innovations can be replaced by a moving average of error terms. Following Chan (2020b), for the serial dependence of covariance matrix over time, suppose the error term $\boldsymbol{\varepsilon}_t^y$ follows a heteroscedastic moving average innovation process. More precisely, we assume $\boldsymbol{\varepsilon}_t^y$ has an MA(1) stochastic volatility process:

$$\boldsymbol{\varepsilon}_t^y = u_t + \boldsymbol{\psi}_{\boldsymbol{\varepsilon}} u_{t-1}, \quad u_t \sim \mathcal{N}(\mathbf{0}, \lambda_t e^{h_t} \boldsymbol{\Sigma}). \tag{7}$$

Here, the covariance matrix $\boldsymbol{\Omega}$ in Equation 4 has $((1+\boldsymbol{\psi}_{\boldsymbol{\varepsilon}}^2)\lambda_1 e^{h_1}, ..., (1+\boldsymbol{\psi}_{\boldsymbol{\varepsilon}}^2)\lambda_T e^{h_T})$ along its main diagonal, $(\boldsymbol{\psi}_{\boldsymbol{\varepsilon}}\lambda_1 e^{h_1}, ..., \boldsymbol{\psi}_{\boldsymbol{\varepsilon}}\lambda_{T-1} e^{h_{T-1}})$ above and below the main diagonal, and 0 elsewhere.

Table 1 summarizes specifications considered in our main analysis. For our forecasting exercise, we start with a small BVAR model with conventional error assumptions and consider this model as a benchmark. We then include a larger BVAR model and augment the aforementioned features of the covariance structure into the standard BVAR.

Table 1: A list of competing models.

Model	Description
Small BVAR	3-variable VAR with standard error assumptions
BVAR	20-variable VAR with standard error assumptions
BVAR-CSV	20-variable VAR with a common stochastic volatility
BVAR-CSV-t	20-variable VAR with a CSV and t errors
BVAR-CSV-t-MA	20-variable VAR with a CSV and MA(1) t errors

All models are estimated using Markov chain Monte Carlo method (MCMC), see Appendix B for details on simulation. The estimation results in our empirical studies are all based on 5000 posterior samples obtained after a burn-in period of 1000. With regard to priors, for the comparison purposes, whenever possible we choose exactly the same priors for the common parameters across models. In particular, the Minnesota prior and the natural conjugate prior is used for the standard VARs and flexible VARs respectively. Details of values of the hyperparameter of these prirors are reported in Appendix C.

4 Forecast Results

In this section, we perform a recursive out-of-sample forecasting exercise to evaluate the performance of the proposed VARs in terms of both point and density forecast. For expository purposes, in the analysis below we focus on the performance of the models listed in Table 1. Additional results under other possible combinations of CSV, t innovations and MA, such as BVAR-t, BVAR-t-MA, ..., can be found in Appendix D.1.

4.1 Forecast Evaluation Metrics

To evaluate the forecast performance of each of the Baysian VAR models listed in Table 1, we perform a recursive out-of-sample forecasting exercise to obtain both point and density forecast. The recursive exercise will involve using data available up to time $t-1$ released in vintage t to forecast at time $t+k$ for $k = 0, 1, 2$ and 4. Thus, the forecast horizons are nowcasts, one-quarter-ahead, two-quarter-ahead and one-year-ahead. We focus on three target variables: real GDP growth, CPI inflation and unemployment rate. Following standard practice, we set the lag length to p = 4.

The accuracy of the point forecast is assessed by root mean square forecast error (RMSFE). RMSFE is a commonly used scale dependent measure for each time series with the same unit. For RMSFE, a smaller value comes from a smaller forecast error and stands for a better forecast performance. The value of RMSFE for the target variable i ($i = 1, 2, 3$) at forecast horizon k ($k = 0, 1, 2, 4$) is calculated as:

$$\text{RMSFE}_{i,k} = \frac{1}{T - k - T_0} \sum_{t=T_0-1}^{T-k-1} \sqrt{(y^{\text{o}}_{i,t+k} - \mathbb{E}(y_{i,t+k} \,|\, \mathbf{y}^t_{1:t-1}))^2},$$

where T_0 is the start of the evaluation period, $y^{\text{o}}_{i,t+k}$ is the observed value of the interested variable in the latest vintage, and $\mathbb{E}(y_{i,t+k} \,|\, \mathbf{y}^t_{1:t-1})$ is the sample mean of forecasts given information of the variable up to time $t - 1$ in vintage t.

As point forecast ignores the predictive distribution of forecast results, we also evaluate the forecast performance from predictive distribution of density forecast by the average of log predictive likelihood (ALPL). For the estimation $y_{i,t+k}$ in vintage t, the predictive likelihood is obtained by the predictive density evaluated at the observation $y^{\text{o}}_{i,t+k}$. More specifically, the ALPL is defined as:

$$\text{ALPL}_{i,k} = \frac{1}{T - k - T_0} \sum_{t=T_0-1}^{T-k-1} \log p(y_{i,t+k} = y^{\text{o}}_{i,t+k} \,|\, \mathbf{y}^t_{1:t-1}),$$

where $p(y_{i,t+k} = y^{\text{o}}_{i,t+k} \,|\, \mathbf{y}^t_{1:t-1})$ is the predictive likelihood with information of the interested variable up to time $t - 1$ in vintage t. Given the predictive distribution, a larger value of predictive likelihood means that the observation $y^{\text{o}}_{i,t+k}$ is more likely under the predicted density forecast. In other words, a larger value of ALPL indicates better forecast performance.

4.2 Forecasting Results

In this section we discuss the forecast performance of the proposed BVAR models for GDP growth, CPI inflation and unemployment rate. We report the point and density forecast results of these models for each variable in Table 2, 3 and 4, respectively. For easy comparison, we report the ratios of RMSFEs of a given model to those of the benchmark BVAR using the three core variables. Hence, values smaller than unity indicate better forecast performance than the small BVAR. For ALPLs, we report differences from that of the small BVAR. In this case, positive values indicate better forecast performance than the benchmark.

Overall, the results suggest the covariance structure that the forecaster chooses to embed to the BVAR model, along with the model size, effectively impacts the forecast performance. In particular, for the case of the Australian macroeconomy, we find three consistent patterns. First, the large BVAR models tend to outperform the small BVAR for all three core variables, especially for horizons after nowcasts. These findings are consistent with the results in Koop (2013), Panagiotelis et al. (2019) and Chan (2020b). For example, using real-time dataset for the US, Chan (2020b) finds that a BVAR model with 20 variables tend to forecast real variables better than a small BVAR. In the Australian context, our results further reveal that the large VAR models also do well for norminal variables. A similar conclusion is also found in Panagiotelis et al. (2019). Using truncated series from a single vintage, this study finds that a large model that is not beyond 20 variables tends to provide better forecasts for Australia.

Second, the results also show that the models with more flexible error covariance structures can improve the forecast accuracy of GDP growth and CPI inflation. For GDP growth, BVAR-CSV and BVAR-CSV-t forecast relatively better than the benchmark for both point and density forecasts. For example, BVAR-CSV does better in all horizons in terms of point forecast and BVAR-CSV-t is found to be the best model in terms of density forecasts. More interesting, we find BVAR-CSV-t-MA, the most flexible covariance structure among our proposed models, forecasts CPI inflation substantially better than the benchmark model for both point and density forecasts. This model reduces the RMSFE of BVAR about 7% for all horizons. Our results for Australia further confirm those in recent forecasting literature, such as Clark (2011), D'Agostino et al. (2013), Clark and Ravazzolo (2015), and Chan (2020a). Investigating real-time dataset for the US, these studies consistently find that BVAR models with stochastic volatility tend to outperform their counterparts with constant variance. Other than that, as highlighted in Cross and Poon (2016) and Chan (2020a), forecasting results for Australia also suggest, in many instances, the forecasting accuracy can be further improved by adding more features, such as t error distribution and MA component, to the covariance of the VAR model.

Table 2: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR	0.982	0.965	0.991	0.988	-0.015	-0.028	-0.051	-0.057
BVAR-CSV	**0.978**	**0.988**	**0.998**	**0.985**	0.059	0.034	0.027	-0.004
BVAR-CSV-t	0.983	0.990	1.000	0.989	**0.070**	**0.055**	**0.048**	**0.082**
BVAR-CSV-t-MA	0.988	0.992	1.000	0.988	0.061	0.049	0.044	0.078

Note: Values in bold indicate the best relative RMSFE and ALPL. Gray cells indicate the significant difference of the predictive accuracy between an alternative models and the benchmark small BVAR, at 1% level of significance using the related asymptotic test introduced by Diebold and Mariano (1995).

Table 3: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR	0.936	0.945	0.960	0.971	0.019	0.018	0.006	0.007
BVAR-CSV	0.932	0.922	0.923	0.918	0.116	0.152	0.153	0.192
BVAR-CSV-t	0.925	0.925	0.925	0.921	0.182	0.195	0.197	0.206
BVAR-CSV-t-MA	**0.920**	**0.917**	**0.919**	**0.912**	**0.188**	**0.203**	**0.198**	**0.209**

Note: see Table 2.

Table 4: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR	**0.892**	**0.928**	**0.884**	**0.842**	**0.215**	**0.301**	**0.335**	**0.348**
BVAR-CSV	1.014	0.969	0.934	0.888	0.167	0.229	0.258	0.582
BVAR-CSV-t	0.998	0.963	0.925	0.875	0.184	0.249	0.283	0.310
BVAR-CSV-t-MA	1.001	0.969	0.934	0.885	0.182	0.218	0.252	0.290

Note: see Table 2.

Third, while the flexible BVAR models with more general error distribution produce better forecasts for GDP growth and CPI inflation than other standard models, we find that the BVAR models with standard error assumption provide the most accurate forecasts for unemployment rate. As reported in Table 4, both RMSFE and ALPL indicate that the first four proposed models perform substantially better than the others for all forecast horizons. As described in Section 3, these models fall within a class of standard BVAR model embodying the conventional assumption of homoscedastic, Gaussian and serially independent errors. Our findings for Australia also reflect results observed for the US. Indeed, considering a range of large BVAR models, Chan (2020a) finds that no models can consistently outperform the standard VAR model when forecasting the US unemployment rate. This finding is important because it reflects the natural property of macroeconomic time series that inflation tends to be much more volatile than output growth and unemployment rate. Therefore, BVAR-CSV-t-MA, the most flexible model, is likely the best model to forecast inflation, while the standard model retains enough flexibility to forecast unemployment rate.

5 Sensitivity Analysis

To examine whether our findings are sensitive to the choice of prior and the unprecedented shock caused by the Covid-19 crisis, we ran the models using a set of alternative priors and with the pre-crisis data. The main conclusions of the paper are robust to all of these sensitivity checks. Below we provide a brief summary.

5.1 Prior Sensitivity

Recently, there have been a number of studies, such as Jochmann et al. (2010), Chan (2020b) and Cross et al. (2020), highlighting that the forecast performance might be sensitive to alternative prior choices. Motivated by these empirical observations, aside from the Minnesota prior, we also consider three other priors for the standard VARs: the natural conjugate prior, the independent normal and inverse-Wishart prior and the stochastic search variable selection (SSVS) prior as a sensitivity analysis. Details about these priors are described in Appendix C. A list of competing models with different priors for this exercise are described in Table 5 and the corresponding results are presented in Table 6-8. For easy comparison, we also report the results performed by the standard VAR with the Minnesota prior. Similar to the comparison method used in the main analysis, we use RMSFEs and ALPLs of the small BVAR as a benchmark.

Table 5: A list of competing large VAR models with alternative priors.

Model	Description
BVAR-Minn	20-variable VAR with the Minnesota prior
BVAR-NCP	20-variable VAR with the natural conjugate prior
BVAR-IP	20-variable VAR with the independent prior
BVAR-SSVS	20-variable VAR with the SSVS prior

Overall, our main conclusions are robust to these sensitivity checks. In line with our main findings, for GDP growth and inflation, none of the standard VARs with these proposed priors outperforms the flexible VARs considered in the main analysis. For unemployment, the results suggest the standard VAR remains the best model. Interestingly, we find that the forecast performance of the standard VAR model can be slightly improved under a particular prior class. This is, for the point forecast, it is likely that the BVAR model with independent prior can slightly enhance forecast accuracy for unemployment rate. For density forecasts, the model with the natural conjugate prior relatively does better than its counterparts. As reported in Appendix D.2, we also observe a similar results for data up to vintage 2020Q1, excluding the period of the Covid-19 crisis. We further discuss this event and its potential impacts on our forecast performance in the next section.

Table 6: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-Minn	0.982	0.965	0.991	0.988	-0.015	-0.028	-0.051	-0.057
BVAR-NCP	1.026	0.989	1.015	1.005	0.004	0.014	-0.011	-0.003
BVAR-IP	0.987	0.950	0.982	1.002	0.022	0.035	0.014	0.006
BVAR-SSVS	0.961	0.929	1.007	1.071	0.035	0.035	-0.016	-0.050

Note: see Table 2.

Table 7: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-Minn	0.936	0.945	0.960	0.971	0.019	0.018	0.006	0.007
BVAR-NCP	0.998	1.008	1.001	1.017	0.021	0.017	0.014	0.015
BVAR-IP	0.945	0.965	0.978	0.974	0.061	0.045	0.051	0.054
BVAR-SSVS	1.016	1.010	1.030	0.975	-0.011	0.005	-0.007	0.034

Note: see Table 2.

Table 8: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-Minn	0.892	0.928	0.884	0.842	0.215	0.301	0.335	0.348
BVAR-NCP	0.869	0.914	0.869	0.851	**0.279**	**0.403**	**0.404**	0.342
BVAR-IP	**0.851**	**0.914**	**0.862**	**0.808**	-0.040	0.289	0.374	**0.411**
BVAR-SSVS	0.854	0.913	0.862	0.821	0.035	0.345	0.396	0.391

Note: see Table 2.

5.2 The COVID-19 Crisis

By the time of writing this paper, the COVID-19 outbreak has disrupted the world and produced extremely large variation in many key marcoeconomic variables of its economies and the Australian economy is not an exception. This unprecedented shock thus creates a tremendous challenge for macroeconomic forecasting as it demands unusual assumptions (Schorfheide et al., 2020; Primiceri and Tambalotti, 2020; Lenza and Primiceri, 2020). With that in mind, we conduct the sensitivity analysis of our forecasts by re-estimating the models using pre-crisis data up until the end of 2019. The point and density forecast results for GDP, CIP inflation and unemployment rate are reported in Table 6, 7 and 8, respectively. We find that the main results remains unchanged. This is, a large BVAR model remains a better choice for forecasting the Australian macroeconomy. In particular, models with flexible covariance structures are still competitive models when forecasting GDP and CPI inflation, while standard large BVAR models are useful when forecasting unemployment rate.

Table 9: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR	0.982	0.966	0.988	0.985	-0.016	-0.029	-0.050	-0.056
BVAR-CSV	**0.979**	**0.988**	**0.998**	**0.983**	0.053	0.036	0.028	-0.009
BVAR-CSV-t	0.982	0.991	0.998	0.986	**0.066**	**0.056**	**0.052**	**0.086**
BVAR-CSV-t-MA	0.988	0.990	0.997	0.985	0.058	0.053	0.048	0.081

Note: see Table 2.

Table 10: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
L-BVAR	0.940	0.944	0.959	0.970	0.017	0.018	0.011	0.001
BVAR-CSV	0.933	0.923	0.923	0.918	0.108	0.145	0.157	0.197
BVAR-CSV-t	0.927	0.926	0.927	0.924	0.186	0.192	0.194	0.194
BVAR-CSV-t-MA	**0.923**	**0.917**	**0.919**	**0.912**	**0.186**	**0.197**	**0.198**	**0.202**

Note: see Table 2.

Table 11: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR	**0.893**	**0.931**	**0.888**	**0.847**	**0.216**	**0.300**	**0.333**	**0.345**
BVAR-CSV	1.012	0.969	0.934	0.889	0.170	0.230	0.262	0.588
BVAR-CSV-t	0.997	0.964	0.925	0.877	0.187	0.248	0.283	0.308
BVAR-CSV-t-MA	1.005	0.971	0.937	0.888	0.180	0.217	0.250	0.285

Note: see Table 2.

6 Conclusion

In this paper, we have studied the forecast performance of a set of BVARs for the Australian macroeconomy. In light of the recent development in BVAR models, we considered a wide range of BVAR modifications that is equipped with alternative priors and allow for various flexible error covariance structures. In addition, we also constructed and for the first time we utilized the real-time data in forecasting core indicators for Australia. We focused on three core variables, including GDP growth, CPI inflation and unemployment rate and found that a large BVAR model with 20 variables tends to outperform a small BVAR model. Specifically, we find consistent evidence that the models with more flexible error covariance structures forecast GDP growth and CPI inflation better than the standard VAR, whereas the standard VAR using conventional covariance assumptions does better than its counterparts when forecasting unemployment rate. These findings are found to remain unchanged under alternative priors and when we examine the early stage of the Covid-19 crisis.

References

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector autoregressions. *Journal of Applied Econometrics*, 25(1):71–92.

Bjørnland, H. C., Ravazzolo, F., and Thorsrud, L. A. (2017). Forecasting GDP with global components: This time is different. *International Journal of Forecasting*, 33(1):153–173.

Carriero, A., Clark, T., and Marcellino, M. (2016). Common drifting volatility in large Bayesian VARs. *Journal of Business & Economic Statistics*, 34(3):375–390.

Carriero, A., Clements, M. P., and Galvão, A. B. (2015). Forecasting with Bayesian multivariate vintage-based VARs. *International Journal of Forecasting*, 31(3):757–768.

Carriero, A., Kapetanios, G., and Marcellino, M. (2009). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting*, 25:400–417.

Chan, J. C. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172.

Chan, J. C. (2020a). Large Bayesian VARs: A flexible Kronecker error covariance structure. *Journal of Business & Economic Statistics*, 38(1):68–79.

Chan, J. C. (2020b). *Large Bayesian Vector Autoregressions*. Macroeconomic forecasting in the era of big data. Springer.

Chan, J. C. and Hsiao, C. (2014). *Estimation of stochastic volatility models with heavy tails and serial dependence*. Bayesian Inference in the Social Sciences. John Wiley & Sons, Hoboken.

Chiu, C.-W. J., Mumtaz, H., and Pinter, G. (2017). Forecasting with VAR models: Fat tails and stochastic volatility. *International Journal of Forecasting*, 33(4):1124–1143.

Clark, T. E. (2011). Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, 29(3):327–341.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Clements, M. P. (2017). Assessing macro uncertainty in real-time when data are subject to revision. *Journal of Business & Economic Statistics*, 35(3):420–433.

Clements, M. P. and Galvão, A. B. (2013). Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. *Journal of Applied Econometrics*, 28(3):458–477.

Cross, J. L., Hou, C., and Poon, A. (2020). Macroeconomic forecasting with large Bayesian VARs: Global-local priors and the illusion of sparsity. *International Journal of Forecasting*, https://doi.org/10.1016/j.ijforecast.2019.10.002.

Cross, J. L. and Poon, A. (2016). Forecasting structural change and fat-tailed events in Australian macroeconomic variables. *Economic Modelling*, 58:34–51.

Croushore, D. (2011a). Forecasting with real-time data vintages. *The Oxford handbook of economic forecasting*, pages 247–267.

Croushore, D. (2011b). Frontiers of real-time data analysis. *Journal of economic literature*, 49(1):72–100.

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Econometric Statistics*, 13(3):134–144.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distribution. *Econometric Reviews*, 3(1):1–100.

Eickmeier, S. and Ng, T. (2011). Forecasting national activity using lots of international predictors: An application to New Zealand. *International Journal of Forecasting*, 27(2):496–511.

Garratt, A., Koop, G., Mise, E., and Vahey, S. P. (2009). Real-time prediction with U.K. monetary aggregates in the presence of model uncertainty. *Journal of Business & Economic Statistics*, 27(4):480–491.

George, E. I., Sun, D., and Ni, S. (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, 142(1):553–580.

Giannone, D., Lenza, M., and Primiceri, G. E. (2015). Prior selection for vector autoregressions. *The Review of Economics and Statistics*, 97(2):436–451.

Jochmann, M., Koop, G., and Strachan, R. W. (2010). Bayesian forecasting using stochastic search variable selection in a VAR subject to breaks. *International Journal of Forecasting*, 26(2):326–347.

Koop, G. and Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4):267–358.

Koop, G. M. (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, 28(2):177–203.

Lee, K., Olekalns, N., Shields, K., and Wang, Z. (2012). Australian real-time database: An overview and an illustration of its use in business cycle analysis. *Economic Record*, 88(283):495–516.

Lenza, M. and Primiceri, G. E. (2020). How to Estimate a VAR after March 2020. Technical report, National Bureau of Economic Research.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.

Mumtaz, H. and Theodoridis, K. (2018). The changing transmission of uncertainty shocks in the us. *Journal of Business & Economic Statistics*, 36(2):239–252.

Panagiotelis, A., Athanasopoulos, G., Hyndman, R. J., Jiang, B., and Vahid, F. (2019). Macroeconomic forecasting for Australia using a large number of predictors. *International Journal of Forecasting*, 35(2):616–633.

Poon, A. (2018). Assessing the synchronicity and nature of australian state business cycles. *Economic Record*, 94(307):372–390.

Primiceri, G. E. and Tambalotti, A. (2020). Macroeconomic forecasting in the time of COVID-19. *Manuscript, Northwestern University.*

Schorfheide, F. and Song, D. (2015). Real-time forecasting with a mixed-frequency VAR. *Journal of Business & Economic Statistics*, 33(3):366–380.

Schorfheide, F., Song, D., et al. (2020). Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic. Technical report.

Zhang, B., Chan, J. C., and Cross, J. L. (2020). Stochastic volatility models with ARMA innovations: An application to G7 inflation forecasts. *International Journal of Forecasting*, 36:1318–1328.

Appendix A Data

Table A1: Description of variables used in the recursive forecasting exercise.

Variable	Data revision	Frequency	Transformation
Three main variables using in the small VAR and large VAR			
Real GDP	Y	Q	$400\Delta\log$
CPI	Y	Q	$400\Delta\log$
Unemployment Rate, seasonally adjusted	Y	M	no
Remaining variables using in the large VAR			
Real Household Final Consumption	Y	Q	$400\Delta\log$
Real Gross Fixed Capital Formation	Y	Q	$400\Delta\log$
Real General Government Final Expenditure	Y	Q	$400\Delta\log$
Real Exports of Goods and Services	Y	Q	$400\Delta\log$
Real Imports of Goods and Services	Y	Q	$400\Delta\log$
Manufacturing Production Index	Y	Q	$400\Delta\log$
Industrial Production Index	Y	Q	$400\Delta\log$
Employed Persons, seasonally adjusted	Y	M	$400\Delta\log$
M3, seasonally adjusted	N	M	$400\Delta\log$
Broad Money, seasonally adjusted	N	M	$400\Delta\log$
90 Days Bank Accepted Bills	N	M	no
Interbank Overnight Cash Rate	N	M	no
Real Exchange Rate Measure	N	Q	$400\Delta\log$
10 Year Australia Government Security	N	M	Δ
Commodity price index	N	M	$400\Delta\log$
SP ASX AllOrds	N	M	$400\Delta\log$
Terms of Trade	N	M	$400\Delta\log$

Notes: As mentioned in Section 2, the real-time data taken the Australian Real-time Macroeconomics Database (ARMD) are only available up to 2017Q1. We extended the database from 2017Q2 to the latest vintages by collecting data from ABS and RBA.

Appendix B Estimation

The posterior estimation for parameters of the BVAR models can be obtained by sampling sequentially by Markov chain Monte Carlo (MCMC) methods. Here, we take the estimation of parameters in BVAR-CSV-MA-t as an example. There are seven steps in one loop of posterior draws for each parameter. Specifically, the posterior draws are obtained for the coefficients of VAR $\mathbf{B}$, the cross-sectional covariance matrix $\mathbf{\Sigma}$, the hyperparameter λ_t and ν of t distribution, the stochastic volatility parameter h and the related truncated normal parameter ρ and variance σ_h^2, and the moving average coefficient ψ. The simulation can be implemented as below:

1. $p(\mathbf{B}, \mathbf{\Sigma} \,|\, \mathbf{Y}, \lambda_t, \mathbf{h}, \sigma_h^2, \rho_h, \psi_{\varepsilon}, \nu_{\lambda})$;
2. $p(\lambda_t \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \mathbf{h}, \sigma_h^2, \rho_h, \psi_{\varepsilon}, \nu_{\lambda})$;
3. $p(\nu_{\lambda} \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \lambda_t, \mathbf{h}, \sigma_h^2, \rho_h, \psi_{\varepsilon})$;
4. $p(\mathbf{h} \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \lambda_t, \sigma_h^2, \rho_h, \psi_{\varepsilon}, \nu_{\lambda})$;
5. $p(\sigma_h^2 \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \lambda_t, \mathbf{h}, \rho_h, \psi_{\varepsilon}, \nu_{\lambda})$;
6. $p(\rho_h \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \lambda_t, \mathbf{h}, \sigma_h^2, \psi_{\varepsilon}, \nu_{\lambda})$;
7. $p(\psi_{\varepsilon} \,|\, \mathbf{Y}, \mathbf{B}, \mathbf{\Sigma}, \lambda_t, \mathbf{h}, \sigma_h^2, \rho_h, \nu_{\lambda})$;

In the first step, given that the coefficients and covariance matrix are the natural conjugate prior, the joint posterior distribution of $(\mathbf{B}, \mathbf{\Sigma})$ is a normal-inverse-Wishart distribution, so the posterior draws can be obtained from their posterior distribution directly.

The second and third steps draw the parameter λ_t and ν_{λ} for t distribution which can be written as a scale mixture of Gaussian distribution. This multivariate t distribution has a mean vector $\mathbf{0}$, scale matrix $\mathbf{\Sigma}$ and degree of freedom ν, and $(\lambda_t \,|\, \nu_{\lambda})$ follows an inverse-gamma distribution. Then we have $\mathbf{\Omega} = \text{diag}(\lambda_1, ..., \lambda_T)$. The hyperparameter ν_{λ} in the inverse-gamma distribution of λ_t can be sampled by an independence-chain Metropolis-Hastings step described in Chan and Hsiao (2014).

The following three steps are related to the common stochastic volatility parameter $\mathbf{h}$ and its hyperparameter σ_h^2 and ρ_h. The simulation of common stochastic volatility can follow Carriero et al. (2016) and the models are assumed to have a stationary AR(1) stochastic volatility. We assume that σ_h^2 has an inverse-gamma prior and ρ_h has an independent truncated normal distribution. Then the posterior distribution of parameter $\mathbf{h}$ can be obtained by implementing Newton-Raphson algorithm and the acceptance-rejection Metropolis-Hastings step.

Lastly, the posterior distribution of moving average parameter $\psi_{\boldsymbol{\varepsilon}}$ can be sampled by an independence-chain Metropolis-Hastings step, while the related estimation method and efficient algorithm are discussed in Chan (2013).

Appendix C Priors

The selection of priors is a crucial step in BVAR estimation, as the number of coefficients which needs to be estimated can be a great amount. This overparameterization problem can be eliminated by using informative priors or regularization. In the setup of coefficient prior, the Minnesota Prior is considered in the standard VARs, and the natural conjugate prior is used in the VARs with various flexible covariance structures. We also present the forecast results of models with other prior settings for sensitivity analysis (e.g., the independent normal and inverse-Wishart prior, and the stochastic search variable selection prior. The aim of these priors are the same, which is try to shrink the BVAR to a more parsimonious structure so that the estimation is applicable.

C.1 Minnesota Prior

The Minnesota prior is firstly introduced with small VARs by Doan et al. (1984). It uses an approximation $\widehat{\sigma^2}$ for error covariances in each VAR equation by OLS estimation, so it is not limited by the size of VAR and can be applied to a large BVAR. In the prior distribution of the coefficients, the means and the variances imposed distributions associated with the lag length l of variable's own lag and the lag of another variable. Specifically, a modified version is used which is discussed in Koop and Korobilis (2010):

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\beta}_{Minn}, \mathbf{V}_{Minn}), \tag{8}$$

$$\mathbf{V}_{Minn} = \begin{cases} b_1 & \text{for intercept,} \\ b_2/l^2 & \text{for own lags,} \\ b_3\widehat{\sigma_i^2}/(l^2\widehat{\sigma_j^2}) & \text{otherwise,} \end{cases} \tag{9}$$

where $\boldsymbol{\beta}_{Minn} = 0$ indicates that growth rate data are used and they are stationary time series. $\mathbf{V}$ is the variance operator, b_1, b_2 and b_3 are hyperparameters of $\mathbf{V}_{Minn}$.

The shrinkage degree of $\mathbf{V}_{Minn}$ is consistent with the variable's own lag with l^2 for parameters with either own or cross lag. In other words, more reliable information is provided by more recent lags which should be given more weight in the estimation. In practise, the value of $\mathbf{V}_{Minn}$ is smaller when the lag length l turns larger. In addition, the value of $\mathbf{V}_{Minn}$ is also controlled by the ratio of prior variance from two variables.

For the cross lags, it is supposed that the lags of other variables can not explain more variation of one variable than its own lags, so the $\mathbf{V}_{Minn}$ of cross lags should be smaller than that of own lags.

In the application part, for the small standard BVAR with the Minnesota prior, the hyperparameters of the variance operator are set to be $b_1 = 10^2, b_2 = 0.2^2$ and $b_3 = 0.1^2$, where b_2 is bigger than b_3 indicating that variables' own lags are more important than their cross lags. With the Minnesota prior, the BVARs are models with constant variances, then a two-step Gibbs sampler can be used to estimate the models. The VAR coefficients $\boldsymbol{\beta}$ are drawn from a conditional posterior distribution that is multivariate normal in the first step, and the covariance matrix $\boldsymbol{\Sigma}$ is simulated from an inverse Gamma distribution in the second step. The additional detail on algorithms and priors can be found in Koop and Korobilis (2010). For the 20-variable BVAR with the Minnesota prior, the prior settings and the estimation of the model are the same as those of the standard BVAR.

The setting of Minnesota prior provides a way of shrinkage for the standard VARs with considerable amount of coefficient, but the parameters of Minnesota prior are restricted to be fixed and the covariance matrix is a diagonal matrix. To cover these concerns, alternative priors in the sensitivity analysis section introduce hyperparameters or other flexible specifications on the covariance matrix to the VAR models.

C.2 The Natural Conjugate Prior

The natural conjugate prior (NCP)is used as the prior of VARs with flexible covariance structures, which assumes that the error covariance matrix of VARs is an unknown symmetric matrix. It can be considered as the Minnesota prior with a normal-inverted-Wishart assumption on the error covariance matrix $\boldsymbol{\Sigma}$ instead of a fixed diagonal matrix. This prior takes into account of the uncertainty of the error covariance matrix. Moreover, it is computational tractable and has a closed form of the marginal likelihood comparing with the Minnesota prior. The normal-inverted-Wishart prior takes the following form:

$$\mathbf{B}|\boldsymbol{\Sigma} \sim \mathcal{N}(\mathbf{B}_0, \boldsymbol{\Sigma} \otimes \mathbf{V}_B), \quad \boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{\nu}_0, \mathbf{S}_0), \tag{10}$$

where $\mathbf{B}_0, \mathbf{V}_B, \boldsymbol{\nu}_0$ and $\mathbf{S}_0$ are prior hyperparameters of Normal distribution and inverted Wishart distribution, parameters with the subscript 0 stand for those of the prior distributions. Equation (10) can be written as:

$$(\mathbf{B}, \boldsymbol{\Sigma}) \sim \mathcal{NIW}(\mathbf{B}_0, \mathbf{V}_B, \boldsymbol{\nu}_0, \mathbf{S}_0). \tag{11}$$

In NCP, parameters of larger lag lengths are conducted higher degree of shrinkage. However, there is no difference between the prior variances of variables' own lags and other lags comparing with the feature of Minnesota prior. NCP gives the same degree of shrinkage of variables' own lags and other lags, thus, they share the same tightness hyperparameter on variables' lags.

In the application part, We set $\boldsymbol{\beta}_0 = \mathbf{0}$, the hyperparameters for the covariance matrix $\mathbf{V}_0$ are $b_1 = 10^2, b_2 = 0.2^2$ (so that the parameters with larger lag lengths are conducted higher degree of shrinkage, which is consist with the setting of the Minnesota prior), $\boldsymbol{\nu}_0 = n + 3$ and $\mathbf{S}_0 = diag(s_1^2, ..., s_n^2)$ (where $s_1^2, ..., s_n^2$ are obtained from the standard OLS estimates of the error variance for each equation). To estimate the models with NCP, the Kronecker structure of the posterior covariance matrix can be considered for fast simulation. This approach is based on the algorithm of drawing posteriors from the matrix normal distribution. As the posterior distributions of the VAR coefficients $\boldsymbol{\beta}$ and the covariance matrix $\boldsymbol{\Sigma}$ have the same distributions of the priors, $\boldsymbol{\Sigma}$ can be drawn marginally from an inverse gamma distribution, then $\boldsymbol{\beta}$ can be simulated from a normal distribution.

The detailed algorithm for BVAR with NCP is described in Giannone et al. (2015) and Carriero et al. (2009). When conducting forecast, one-step-ahead forecast can be obtained from the analytical form of the predictive density, but there is no analytical formula for forecasting over one period ahead. In other word, direct forecast method needs to be used when forecast is more than one-step-ahead.

C.3 Independent Normal and Inverse-Wishart Prior

The independent normal and inverse-Wishart prior does not have any restriction on the prior covariance matrix of coefficients $\mathbf{V}$. Thus, it is more flexible on the assumption of prior parameters. By assumption, the priors of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are independent, and they have normal and inverse-Wishart distributions, respectively:

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\beta}_0, \mathbf{V}_{\boldsymbol{\beta}}), \quad \boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{\nu}_0, \mathbf{S}_0), \tag{12}$$

The posteriors of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ with independent normal and inverse-Wishart prior do not have an analytical expression. They need to be simulated from their conditional distributions $p(\boldsymbol{\beta}|\mathbf{y}, \boldsymbol{\Sigma})$ and $p(\boldsymbol{\Sigma}|\mathbf{y}, \boldsymbol{\beta})$. In the simulation of posterior distributions, forward-backward substitution and precision-based algorithm should be considered for rapid computation (see Chan (2020b)). For the independent normal and inverse-Wishart prior, the values of $\boldsymbol{\beta}_0$ and $\mathbf{V}_0$ are set the same as those in the Minnesota prior, and the values of $\boldsymbol{\nu}_0$ and $\mathbf{S}_0$ are set as those in the natural conjugate prior. The posterior distributions of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are drawn from a two-step Gibbs sampler. In the simulation, the forward-backward substitution and the precision-based algorithm can be considered for rapid computation (see Chan (2020b)).

C.4 Stochastic Search Variable Selection Prior

The stochastic search variable selection prior (SSVS) is also a shrinkage prior for VAR coefficients which is introduced firstly by George et al. (2008). The hierarchical structure of parameter's prior is consisted with the independent normal and inverse-Wishart prior. The difference between SSVS and priors in the previous sections is that the restriction on coefficients of a variable's lags to be zero needs to follow a selection procedure. Comparing with the Minnesota prior, SSVS is also a data-based prior, but parameters of variables' cross lags are not restricted to be zero or close to zero when using SSVS.

In its prior selection procedure, the parameters of VAR conduct "stochastic search" and do "variable selection" from two groups with a probability q_s and $1 - q_s$ in an independent Bernoulli distribution during simulation. In one of these two groups, the priors of coefficients are strongly shrunk to zero with small variances σ^2_{s1}, while in the other group, the priors are relative non-informative. More specifically, the SSVS can be writen in a mixture distribution as below:

$$\boldsymbol{\beta}_{s,j}|q_s \sim \begin{cases} \mathcal{N}(0, \sigma^2_{s1,j}), & \text{with prob. } q_s, \\ \mathcal{N}(0, \sigma^2_{s2}), & \text{with prob. } 1 - q_s, \end{cases} \tag{13}$$

where $\boldsymbol{\beta}_{s,j}$ stands for coefficients estimated by SSVS. In the application part, $\sigma^2_{s1,j}$ are estimated by the Minnesota prior and are the elements of the diagonal of $\mathbf{V}_{Minn}$. For the other group, the non-informative prior of σ^2_{s2} is set to be 10 in the estimation. The weight of mixture distribution q_s reflect the information from the prior that whether the parameter is different from zero, so it can be set to any value between 0 and 1 or estimated by hyperparameters. Here, it is set to be 0.5 for equal chances for simplicity.

For the covariance matrix $\boldsymbol{\Sigma}$, we assume it has the inverse-Wishart prior:

$$\boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{\nu}_0, \mathbf{S}_0), \tag{14}$$

where $\boldsymbol{\nu}_0$ and $\mathbf{S}_0$ are prior hyperparameters of inverted Wishart distribution.

The estimation of models with SSVS prior can use a three-step Gibbs sampler for parameter simulation. The first two steps of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are the same of those in the independent normal and inverse-Wishart prior. In the step of $\boldsymbol{\beta}$ posterior distribution, the hyperparameter $\sigma^2_{s1,j}$ are estimated by the Minnesota prior and are the elements of the diagonal of $\mathbf{V}_{Minn}$, while the non-informative prior of σ^2_{s2} is set to be 10 in the estimation for the other group. The third step is to simulate the success rate of probability q_s in the mixture Gaussian distribution. More details can be seen in George et al. (2008).

Appendix D Additional Results

D.1 Results under other Covariance Structure Combinations

Table D1.2-D1.7 below present additional results obtained under a large VAR model when we consider other covariance structure combinations. A long with the specifications examined in the main analysis, this extension provides a full possible combination of non-standard error assumptions. As described in Table D1.1, these specifications include BVAR-t, BVAR-MA, BVAR-t-MA and BVAR-CSV-MA. The point and density forecast results for GDP, CPI inflation and unemployment rate for data up to vintage 2020Q2 are reported in Table D1.2, D1.3 and D1.4 respectively. The corresponding results for data up to vintage 2020Q1 are presented in Table D1.5, D1.6 and D1.7. Overall, these results support our main conclusion that models with flexible covariance structure tend to forecast better than a small VAR with standard error covariance assumptions. Although these specifications tend to forecast well, none of them are found to perform better than the selected specifications discussed in our main analysis.

Table D1.1: A list of other competing models.

Model	Description
BVAR-t	20-variable VAR with t errors
BVAR-MA	20-variable VAR with MA(1) errors
BVAR-t-MA	20-variable VAR with MA(1) t errors
BVAR-CSV-MA	20-variable VAR with a common SV and MA(1) errors

Table D1.2: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-t	1.007	1.009	1.020	1.004	0.062	0.032	0.013	-0.017
BVAR-MA	1.033	0.999	1.016	1.015	-0.004	0.007	-0.014	-0.009
BVAR-t-MA	1.013	1.016	1.027	1.005	0.053	0.011	0.000	-0.029
BVAR-CSV-MA	0.985	0.994	1.002	0.983	0.048	0.031	0.024	-0.008

Note: see Table 2.

Table D1.3: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-t	0.934	0.949	0.949	0.952	0.236	0.277	0.279	0.267
BVAR-MA	0.992	1.014	1.014	1.029	0.034	-0.006	0.003	-0.001
BVAR-t-*MA*	0.932	0.947	0.963	0.952	0.238	0.293	0.296	0.283
BVAR-CSV-MA	0.926	0.919	0.932	0.917	0.124	0.156	0.145	0.199

Note: see Table 2.

Table D1.4: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q2.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-t	0.943	0.942	0.903	0.865	0.212	0.439	0.501	0.583
BVAR-MA	0.869	0.916	0.869	0.845	0.279	0.398	0.413	0.365
BVAR-t-*MA*	0.944	0.943	0.903	0.857	0.208	0.445	0.516	0.603
BVAR-CSV-MA	1.019	0.972	0.936	0.882	0.163	0.225	0.262	0.617

Note: see Table 2.

Table D1.5: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k=1$	$k=2$	$k=4$	*nowcast*	$k=1$	$k=2$	$k=4$
BVAR-t	1.006	1.010	1.019	1.004	0.063	0.033	0.010	-0.015
BVAR-MA	1.033	1.002	1.019	1.006	-0.004	0.005	-0.014	-0.009
BVAR-t-*MA*	1.011	1.016	1.024	1.002	0.052	0.008	-0.008	-0.033
BVAR-CSV-MA	0.983	0.994	1.001	0.981	0.047	0.034	0.027	-0.012

Note: see Table 2.

Table D1.6: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR-t	0.937	0.952	0.953	0.953	0.239	0.277	0.278	0.264
BVAR-MA	0.996	1.013	1.012	1.022	0.021	0.016	0.007	0.007
BVAR-t-*MA*	0.936	0.948	0.963	0.955	0.239	0.290	0.292	0.279
BVAR-CSV-MA	0.928	0.919	0.933	0.916	0.114	0.148	0.142	0.210

Note: see Table 2.

Table D1.7: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR-t	0.943	0.945	0.906	0.866	0.215	0.432	0.504	0.586
BVAR-MA	0.870	0.919	0.873	0.849	0.278	0.401	0.414	0.359
BVAR-t-*MA*	0.946	0.946	0.906	0.861	0.211	0.445	0.523	0.612
BVAR-CSV-MA	1.018	0.972	0.937	0.885	0.165	0.231	0.265	0.620

Note: see Table 2.

D.2 Results under alternative priors with pre-crisis data

Table C2.1: Relative root MSFE and average log likelihood for nowcasts, one-, two-, and four-step-ahead GDP forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR-Minn	0.982	0.966	0.988	0.985	-0.016	-0.029	-0.050	-0.056
BVAR-NCP	1.023	0.992	1.016	1.003	0.004	0.010	-0.010	-0.002
BVAR-IP	0.987	0.949	0.981	1.000	0.021	0.035	0.014	0.007
BVAR-SSVS	0.949	0.944	1.001	1.062	0.042	0.026	-0.015	-0.045

Note: see Table 2.

Table C2.2: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead CPI inflation forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR-Minn	0.940	0.944	0.959	0.970	0.017	0.018	0.011	0.001
BVAR-NCP	0.999	1.010	1.002	1.017	0.019	0.015	0.005	-0.004
BVAR-IP	0.947	0.965	0.978	0.973	0.059	0.045	0.051	0.053
BVAR-SSVS	1.018	1.009	1.027	0.978	-0.022	0.006	0.009	0.031

Note: see Table 2.

Table C2.3: Relative root MSFE and average log likelihood for nowcast, one-, two-, and four-step-ahead unemployment rate forecasts for data up to vintage 2020Q1.

	relative RMSFE				relative ALPL			
	nowcast	$k = 1$	$k = 2$	$k = 4$	*nowcast*	$k = 1$	$k = 2$	$k = 4$
BVAR-Minn	0.893	0.931	0.888	0.847	0.216	0.300	0.333	0.345
BVAR-NCP	0.868	0.918	0.873	0.853	**0.281**	**0.396**	**0.404**	0.342
BVAR-IP	**0.852**	**0.917**	**0.866**	**0.812**	-0.046	0.286	0.371	**0.406**
BVAR-SSVS	0.855	0.918	0.866	0.825	0.027	0.339	0.391	0.388

Note: see Table 2.