Macroeconomic Forecasting in Large Bayesian VARs: The Case of Taiwan

Kuo-Hsuan Chin and Po-Chuan Liao

Department of Economics
Feng Chia University

July 9, 2025

Introduction

- **Macroeconomic forecasting** is **important** for

    - **Making the appropriate policy**, particularly for the **monetary** policymakers .

    - **Guiding the financial investment decision**, particularly for the public.

    - Something else......

1. **Model-based** approach: the forecasts are generated from **reduced-form**, **structural** and **semi-structural** models.

   1. **Reduced-form model**: **time-series regression**, state-space representation, and most of them are **statistical-type** models. **(Stock and Watson, 2007)**

   2. **Structural model**: dynamic stochastic general equilibrium (DSGE) model, and it particularly has **economic meaning**. **(Herbst, 2024)**

   3. **Semi-structural model**: the combination of reduced-form and structural models. **(Chin and Li, 2019)**

2. **Survey-based** approach: the forecasts are collected and aggregated on the basis of the **individual forecasts from a group of experts**.

   - Example: **Survey of Professional Forecasts (SPF)** is constructed by the Federal Reserve Bank of Philadelphia.

Introduction
Types of Macroeconomic Forecasting

1. **"Point" forecast**: it is the **"best"** guess of the **single value** for the variable of interest.

   - **Example**: the point forecast of the inflation rate in the next quarter is 2.3%.

   - **"Best"**: hopefully it is minimized the forecast error.

2. **"Density" forecast**: it shows the **probability distribution** of the forecast.

   - **Example**: we are confident that inflation rate has 95% probability between 2.1% and 2.5%.

   - The density forecast for the variable of interest is usually generated via **repeated sampling**, such as the Bayesian approach or bootstrapping.

1. We adopt the **model-based approach**. In particular,

   - It is a **reduced-form**, **multivariate time series** model: **vector autoregression, VAR**. **(Sims, 1980)**

   - **VAR: small-scale, medium-scale and large scale**. **(Bańbura, Giannone, and Reichlin, 2010)**

2. We consider both the **point and density forecasts**. **(Wolters, 2015)**

3. Research objective: **Taiwan**

   - We find very limit studies in the past take into account the **Bayesian approach in a large-scale VAR model**. **[Contribution]**

- Using **Bayesian approach**, **treating the (parameter, model) uncertainty explicitly**, is helpful

  - To incorporate the **useful prior information (belief)** into the parameter estimation and statistical inference.

    1. The prior information: it could be coming from **expert's judgment, past data or economy theory**.

    2. The useful prior information is formally and directly incorporated into the **prior density** of something interest, say the parameter, the so-called **information priors**.

- Using **Bayesian approach**, **treating the (parameter, model) uncertainty explicitly**, is helpful

  - To solve the **"over-parameterization" problem**

    1. The problem is usually occurred in a **large-scale** VARs.

    2. The **shrinkage prior**, the so-called **Minnesota prior [it imposes the hypothesis that individual variables all follow random walk processes]**, helps to solve the problem.

- We find the forecasting performance of the **small-scale (3-variable) VARs is dominant** in most rounds of forecast competition, which is also confirmed by the testing results.

  1. Forecast competition: VARs with **3, 7, 12, 20, and 32 variables**.

  2. The above finding is **applied to both point and density forecasts**.

  3. The above finding is further confirmed by the **appropriate tests of relatively forecasting ability**.

- Accordingly, **we do not find any benefit of using a rich set of data information in predicting Taiwan's macroeconomic variables**.

Vector Autoregressions (VARs)

Vector Autoregressions (VARs)
Model Specification (I)

- The pth lag of a VAR model with the intercept, $VAR(p)$, can be written as

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t, \quad t = 1, 2, \ldots, T$$

  1. $y_t = (y_{1,t}, y_{2,t}, \ldots, y_{n,t})$ is a $n \times 1$ vector of **endogenous variable**.
  2. A_0 is a $n \times 1$ vector of intercept.
  3. $A_1, A_2, \ldots, A_p$ are $n \times n$ matrices of sloping parameter.
  4. ε_t is a $n \times 1$ vector of error terms following a multivariate normal density with zero mean and variance-covariance matrix Σ.

- It is a **dense parameterization**, especially in a **large-scale** model. **[The curse of dimensionality or the so-called "over-parameterization" problem]**

- The equation above could be be further **transposed and stacked** as

$$\underbrace{\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} 1 & y_0' & \cdots & y_{1-p}' \\ 1 & y_1' & \cdots & y_{2-p}' \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_{T-1}' & \cdots & y_{T-p}' \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} A_0' \\ A_1' \\ \vdots \\ A_p' \end{bmatrix}}_{B} + \underbrace{\begin{bmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{bmatrix}}_{\varepsilon},$$

$$Y = XB + \varepsilon,$$

- One could further **vectorize** the above equation as follows,

$$y = (I_n \otimes X)\beta + \xi,$$

where $y = vec(Y)$, $\beta = vec(B)$, $\xi = vec(\varepsilon)$, and $\xi \sim N(0, \Sigma \otimes I_T)$. We use the expression above in the application of the Bayesian approach.

Bayesian Approach

Bayesian Approach
Basic Principle

- Bayes rule:

$$\underbrace{p\left(\theta \mid y\right)}_{\text{posterior density}} \propto \underbrace{p\left(y \mid \theta\right)}_{\text{likelihood function}} \underbrace{p\left(\theta\right)}_{\text{prior density}}$$

  - θ refers to the parameter of interest. That is, $\theta = \{\beta, \Sigma\}$; y is data.

  - Regarding the parameters, so-called **hyper-parameters**, that characterize the **prior density**, it is denoted as γ.

  - **Regarding the choice of the prior density (priors) for the parameter and its hyper-parameters, we follow Giannone, Lenza and Primiceri (2015).**

Bayesian Approach
In the Large-Scale VARs

- The **informative conjugate prior density** is proven to be **effectively to solve the "curse of dimensionality" (Bańbura, Giannone, and Reichlin, 2010)**:

  1. **"Informative" priors**:
     - Example: the **random-walk behavior** for some of macroeconomic time series data.
     - It moves some of the parameters toward zero, becoming a **parsimonious model**, **largely reducing the estimation error and improving the out-of-sample forecasting accuracy**.

  2. **"Conjugate" priors**: **time-saving** for the sampling process of the posterior density.

Bayesian Approach
In the Large-Scale VARs

- We follow **Giannone, Lenza and Primiceri (2015)** to specify the priors and its hyper-parameters:

  - It is a **"Conjugate" priors**: it falls in the family of the **Normal-Inverse-Wishart (NIW)** density.

    1. Technically, it is expressed as

       $$\beta \mid \Sigma \sim N(b, \Sigma \otimes \Omega)$$

       $$\Sigma \sim IW(\Psi, d)$$

       where b, Ω, Ψ, and d are hyper-parameters, denoted by γ.

    2. In this paper, the conditional prior density of β is treated as a **Minnesota-type prior**. Regarding the specific moments specified for the hyper-parameters of the priors, please refer to **Kuschnig and Vashold (2021)**.

- We follow **Giannone, Lenza and Primiceri (2015)** to specify the priors and its hyper-parameters:

  - It sets **prior hyper-parameters** in a **data-based fashion**,

    1. Technically, it is expressed as

    $$p\left(\gamma, \theta \mid y\right) \propto p\left(y \mid \theta, \gamma\right) p\left(\theta \mid \gamma\right) p\left(\gamma\right)$$

    2. The **marginal likelihood (ML)** could be obtained on the basis of the previous equation as

    $$p\left(y \mid \gamma\right) = \int p\left(y \mid \theta, \gamma\right) p\left(\theta \mid \gamma\right) d\theta$$

    where it could be used to choose the optimal hyper-parameters.

Empirical Implementation & Results

- **Data source**: National Statistics, R.O.C (Taiwan); statistical database at central bank of the Republic of China.

- Total number of data series: **32 data series**.

- Sample range of the raw data: **1987Q1-2024Q4**. (Quarterly data)

- Pre-estimation data analysis: **transforming each series into percentage form**.

- Sample range for the estimation: **1997Q2-2024Q4. (due to missing data)**

Empirical Implementation & Results
VAR Model Setting

- Regarding the specification of the Bayesian VAR model,

  - The length of the **lag is four**.

  - The number of variables considered in the VARs includes **3, 7, 12, 20, and 32**.

  - The specification of the prior density for the parameter and its hyper-parameters are based on **Giannone, Lenza and Primiceri (2015)**.

  - Regarding the empirical implementation, we use **BVAR package operated in RStudio**. **(Kuschnig and Vashold, 2021)**

Empirical Implementation & Results

Out-of-Sample Forecasting Setting & Forecast Evaluation

- Forecasting scheme: **rolling-sample forecasting**.

  1. In-sample observations: 80 observations; **the first rolling-sample is 1997Q2-2017Q1**.

  2. We generate **one-step-ahead, four-step-ahead, and eight-step-ahead point & density forecasts**.

  3. For the **point forecast** exercise, we compare the relatively forecasting accuracy based on **root mean squared forecast error (RMSE)**, which is further confirmed with modified **Diebold and Mariano (DM) test**. **(Harvey, Leybourne and Newbold, 1997)**

  4. Regarding the **density forecast** exercise, we compare the relatively forecasting accuracy based on **log predictive likelihood**, which is further confirmed with **Amisano and Giacomini (AG) test**. **(Amisano and Giacomini, 2007)**

Empirical Implementation & Results

Empirical Results - Point Forecast (I)

Table 1. Root Mean Squared Forecast Error (RMSE)

VARs	GDP Growth Rate			Inflation Rate		
	$h=1$	$h=4$	$h=8$	$h=1$	$h=4$	$h=8$
3V	**1.58252**	**2.21935**	**2.36091**	**0.47813**	**0.41898**	**0.42526**
7V	1.8861	2.54423	2.69505	0.625	0.60237	0.54296
12V	2.05495	3.08991	2.48975	0.55549	0.5103	0.47491
20V	1.87175	2.72179	2.45624	0.54277	0.58306	0.57043
32V	1.89612	2.62628	2.40223	0.64828	0.60126	0.60815

VARs	Rediscount Rate			
	$h=1$	$h=4$	$h=8$	
3V	**0.02925**	**0.07**	**0.09181**	3V: 3-variable VARs
7V	0.05434	0.11823	0.1278	7V: 7-variable VARs
12V	0.06936	0.12704	0.11596	12V: 12-variable VARs
20V	0.07718	0.11965	0.10903	20V: 20-variable VARs
32V	0.08031	0.12216	0.11245	32V: 32-variable VARs

Table 2. Diebold and Mariano Test

VARs	GDP Growth Rate ($h = 1$)					GDP Growth Rate ($h = 4$)				
	3V	7V	12V	20V	32V	3V	7V	12V	20V	32V
3V	-	0.23	0.13	0.34	0.14	-	0.35	0.12	0.27	0.27
7V		-	0.31	0.74	0.87		-	**0.09***	0.99	0.79
12V			-	0.30	0.58			-	**0.06***	0.19
20V				-	0.41				-	0.61
32V					-					-

VARs	Inflation Rate ($h = 1$)					Inflation Rate ($h = 4$)				
	3V	7V	12V	20V	32V	3V	7V	12V	20V	32V
3V	-	0.21	0.20	**0.06***	**0.01****	-	0.19	0.10	0.11	**0.09***
7V		-	0.67	0.53	0.69		-	0.27	0.76	0.90
12V			-	0.78	0.25			-	0.31	0.18
20V				-	**0.03****				-	0.18
32V					-					-

Table 2. Diebold and Mariano Test (Continued)

VARs	Rediscount Rate ($h = 1$)				
	3V	7V	12V	20V	32V
3V	-	**0.03****	**0.00*****	**0.00*****	**0.00*****
7V		-	**0.00*****	**0.00*****	**0.00*****
12V			-	0.15	0.13
20V				-	0.47
32V					-

VARs	Rediscount Rate ($h = 4$)				
	3V	7V	12V	20V	32V
3V	-	0.28	**0.02****	0.20	0.21
7V		-	0.39	0.92	0.84
12V			-	0.24	0.28
20V				-	0.69
32V					-

Empirical Implementation & Results

Empirical Results - Density Forecast (I)

Table 3. Log Predictive Likelihood

VARs	GDP Growth Rate			Inflation Rate		
	$h = 1$	$h = 4$	$h = 8$	$h = 1$	$h = 4$	$h = 8$
3V	-139.2249	**-75.5243**	**-0.9109**	**-126.9068**	**-66.9956**	**-11.5767**
7V	**-137.7596**	-82.3822	-10.7741	-127.1735	-73.0306	-23.1664
12V	-141.6779	-82.9880	-16.4381	-130.3263	-73.5064	-24.549
20V	-150.4485	-92.3936	-26.4885	-137.5988	-82.4224	-29.8589
32V	-146.3405	-92.1894	-28.0283	-134.9876	-82.8581	-31.3159

VARs	Rediscount Rate		
	$h = 1$	$h = 4$	$h = 8$
3V	-108.7573	**-57.1779**	**-15.7577**
7V	**-108.0958**	-62.3598	-25.7387
12V	-110.9181	-61.8344	-25.0936
20V	-117.5620	-69.4780	-28.9670
32V	-115.5476	-69.9695	-30.6383

Table 4. Amisano and Giacomini Test for Relatively Forecasting Accuracy

	GDP Growth Rate ($h = 1$)					GDP Growth Rate ($h = 4$)				
VARs	3V	7V	12V	20V	32V	3V	7V	12V	20V	32V
3V	-	0.05*	0.00***	0.00***	0.00***	-	0.68	0.00***	0.00***	0.00***
7V		-	0.00***	0.00***	0.00***		-	0.00***	0.00***	0.00***
12V			-	0.00***	0.00***			-	0.00***	0.00***
20V				-	0.00***				-	0.00***
32V					-					-
	Inflation Rate ($h = 1$)					Inflation Rate ($h = 4$)				
VARs	3V	7V	12V	20V	32V	3V	7V	12V	20V	32V
3V	-	0.00***	0.00***	0.00***	0.00***	-	0.00***	0.00***	0.00***	0.00***
7V		-	0.56	0.00***	0.00***		-	0.6218	0.00***	0.00***
12V			-	0.00***	0.00***			-	0.00***	0.00***
20V				-	0.80				-	0.595
32V					-					-

Table 4. Amisano and Giacomini Test for Relatively Forecasting Accuracy (Continued)

VARs	Rediscount Rate ($h = 1$)					Rediscount Rate ($h = 4$)				
	3V	7V	12V	20V	32V	3V	7V	12V	20V	32V
3V	-	0.00***	0.00***	0.00***	0.00***	-	0.00***	0.00***	0.00***	0.00***
7V		-	0.00***	0.00***	0.00***		-	0.35	0.00***	0.00***
12V			-	0.00***	0.00***			-	0.00***	0.00***
20V				-	0.33				-	0.22
32V					-					-

Empirical Implementation & Results
Summary of Empirical Results

- In terms of **point forecast competition**,

  1. Table 1 shows that the **small-scale (3-variable) VAR model has the smallest RMSE**, compared to other VARs, in forecasting three variables of interest.

  2. The results are further **confirmed by the Diebold and Mariano test**.

- Regarding the **density forecast competition**,

  1. **The small-scale VAR model is dominant in most of the competing rounds**.

  2. The results are further **confirmed by the Amisano and Giacomini test**.

- In sum, we **do not find any benefit of using a rich set of information** in predicting output growth, inflation and rediscount rate for Taiwan's macroeconomy.

Conclusion

Conclusion

- We estimate the **VAR models with 3, 7, 12, 20 and 32 variables** respectively.

- To solve the "over-parameterization" problem, we apply the **Bayesian approach**. In particular, the specification of prior density for the parameter and its hyper-parameters is implemented according to the approach of **Giannone, Lenza and Primiceri (2015)**.

- We conduct the **rolling-sample estimation and generate out-of-sample point and density forecasts**.

- **We do not find any benefit of using a rich set of data information in predicting Taiwan's macroeconomic variables**.

References

References

- Amisano, Gianni, and Raffaella Giacomini. (2007) "Comparing Density Forecasts via Weighted Likelihood Ratio Tests." ***Journal of Business & Economic Statistics***, 25, 2, 177-190.

- Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. (2010) "Large Bayesian Vector Autoregressions." ***Journal of Applied Econometrics***, 25, 1, 71-92.

- Chin, Kuo-Hsuan, and Xue Li. (2019) "Bayesian Forecast Combination in VAR-DSGE Models." ***Journal of Macroeconomics***, 59, 278-298.

- Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri. (2015) "Prior Selection for Vector Autoregressions." ***Review of Economics and Statistics***, 97, 2, 436-451.

- Harvey, David, Stephen Leybourne, and Paul Newbold. (1997) "Testing the Equality of Prediction Mean Squared Errors." ***International Journal of Forecasting***, 13, 2, 281-291.

References

- Herbst, Edward. (2024) "Forecasting with DSGE Models." In **_Handbook of Research Methods and Applications in Macroeconomic Forecasting_**, 126-155. Edward Elgar Publishing.

- Kuschnig, Nikolas, and Lukas Vashold. (2021) "BVAR: Bayesian Vector Autoregressions with Hierarchical Prior Selection in R." **_Journal of Statistical Software_**, 100, 1-27.

- Sims, Christopher A. (1980) "Macroeconomics and Reality." **_Econometrica_**, 1-48.

- Stock, James H., and Mark W. Watson. (2007) "Why has US Inflation Become Harder to Forecast?" **_Journal of Money, Credit and Banking_**, 39, 3-33.

- Wolters, Maik H. (2015) "Evaluating Point and Density Forecasts of DSGE Models." **_Journal of Applied Econometrics_**, 30, 1, 74-96.

Thank you