Bayesian forecasting of Rupee/Dollar exchange rate: Does Minnesota prior matter?

Prabheesh K.P[*,†] and Rakshita Nagalla[*,‡]

†*Assistant Professor, Department of Liberal Arts, IIT Hyderabad*

‡*Undergraduate STudent, Department of Electrical Engineering, IIT Hyderabad*

E-mail: prabheesh@iith.ac.in; rakshita.nagalla@gmail.com

Abstract

In this paper, we extend the Bayesian literature to VAR models to forecast the Indian rupee-dollar exchange rate. We use information about money supply, interest rate and output for this purpose. By evaluating various priors, we conclude that Bayesian VAR using Minnesota prior induce shrinkage and give best performance by overcoming the issue of overparameterisation, especially in case of short term forecasts.

Keywords

Exchange rate, Forecasting, Bayesian VAR

Introduction

In the present globalized world, exchange rate plays a major role in a country's macroeconomic performance. Any fluctuations in the exchange rate affect the competitiveness of the country's exports in the international market and hence affect its external trade balance. Therefore, fore-casting the exchange rate become a major task to formulate appropriate policies to mitigate the impact of exchange rate fluctuations. Theories of the exchange rate of exchange rate such as purchasing power parity (PPP), interest rate parity, monetary approach, the port-folio approach can provide reasonable understanding on the factors that affect the exchange rates, but the predictability power of the macroeconomic factors are still an unsolved debate. For instance, Meese and Rogoff (1983)[1] argued that the exchange rate forecast based upon macroeconomic factors cannot outperform over a random walk model. Given the lack of predictability of macroeconomic factors, the exchange rate is largely disconnected from the macroeconomic fundamentals of the economy (Backmann and Schussler, 2016)[2] and this phenomenon establishes the "exchange rate disconnect puzzle" (Engel, 2008).[3] However, the survey study by Rossi (2013)[4] suggests that the success of forecast of the exchange rate is primarily determined by choice of regressors, the forecast horizon, the sample period, the types of forecasting methods and forecast evaluation method. This further instigates the researchers to search for alternative models rather than exploring new macroeconomic determinants of exchange rates. Among the alternative forecasting models, the Bayesian vector autoregression (BVAR) model is widely used and has been found to outperform the random walk forecast.

BVAR models are estimated based on Bayesian statistical theory and which combines priors with sample information in model estimation. The main advantage of BVAR as compared to standard unrestricted VAR is it reduces the risk of overparameterization by imposing restrictions on the parameters or by shrinking them towards zero. Bayesian methods offer a consistent and logical way of imposing these restrictions, based on prior probability distribution functions, and can thus significantly improve the forecasting accuracy. This model is relevant in the case of emerging economy's exchange rate forecast, where the researcher faces the lack of high-frequency data. Hence, in this paper we forecast exchange rate of Indian rupee against the US dollar, using BVAR approach. Many studies deal with exchange rate forecasting using BVAR.[5–7] However, to our knowledge, there are only a few studies that forecast Indian Rupee exchange rate using BVAR. Dua and Ranjan (2011)[8] found that BVAR outperforms over other models in predicting exchange rate during 1996-2008. Our study is different from the previous study by testing various prior information in the BVAR estimation process as well as extending the forecasting capacity of BVAR to recent data till 2016. Our strategy in this paper is as follows: First, we estimate exchange rate based on random walk model and multivariate model using both standard unrestricted VAR and BVAR and then carry out a recursively out-of-sample forecast focusing in twelve-month forecasting horizons (h=1 4 8 12). Second, we compare the performance of the models using Root Mean Squared Forecast Error (RMSFE) from the forecast. Third, we use various prior information to forecast the exchange rate such as Normal-Wishart prior, Minnesota prior and Natural conjugate prior and then compare the forecasting power of the model. Our overall empirical findings suggest that the forecast based on BVAR approach outperform over the random walk and unrestricted VAR model, especially in the short-run. We also found that forecasting performance improves using the Minnesota prior, due to the shrinkage on the parameters of the model.

Empirical Model and Data

To forecast exchange rate, we adopt two models, one is based on random walk model and second is based on multivariate model. According to the Random walk model, h-step ahead forecast of the exchange rate is given by:

$$y_{i,t+h} = y_{i,t} \tag{1}$$

Whereas in the case of multivariate model, we follow Meese and Rogoff (1983),[1] which express the variation in exchange rate as a function of relative change in money supply, output and interest rates in the two economies.

$$e_t = f(m_t - m_t^*, y_t - y_t^*, i_t - i_t^*) \tag{2}$$

where

e_t denotes log of exchange rate of India (Rs./$)

$i_t - i_t^*$ denotes the difference between Indian (domestic) and US (foreign) Treasury bill rate,

$y_t - y_t^*$ denotes the difference between log of Indian and US index of industrial production, and $m_t - m_t^*$ denotes the difference between log of Indian and US money supply.

Monthly data for the following variables from July 1996 to July 2016 was used for the analysis. This data has been collected from the websites of Reserve Bank of India and U.S. Bureau of Economic Analysis .

Methods

In this study, we employ multivariate forecasting models in the vector autoregressive (VAR) and Bayesian VAR framework. A VAR model with lag p can be written as:

$$y_t = a_0 + \sum_{j=1}^{p} A_j y_{t-j} + e_t \tag{3}$$

where y_t for $t = 1....T$ is an $MX1$ vector containing observations on M time series variables, e_t is an $MX1$ vector of errors, a_0 is an $MX1$ vector of intercepts and A_j is an MXM matrix of coefficients. We assume e_t to be i.i.d. $N(0, \sum)$.

If we define $x_t = (1, y'_{t-1}, ..., y'_{t-p})$ and

$$X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_T \end{bmatrix} \tag{4}$$

Note that, if we let $K = 1 + Mp$ be the number of coefficients in each equation of the VAR, then X is a TXK matrix.

Finally, if we let $A = (a_0 A_1 ... A_p)$, then $\alpha = vec(A)$ is a $KMX1$ vector which stacks all the VAR coefficients (and the intercepts) into a vector. With all these definitions, we can write the VAR either as:[9]

$$Y = XA + E \tag{5}$$

or

$$y = (I_M \otimes X)\alpha + e \tag{6}$$

where $e \sim N(0, \sum \otimes I_T)$

This VAR model is estimated using both the traditional OLS approach and using Bayesian approach. The OLS estimates of A and $\sum$ are given by:

$$\hat{A} = (X'X)^{-1}X'Y \tag{7}$$

and

$$\hat{\sum} = \hat{S}/(T-K) \tag{8}$$

where $\hat{S} = (Y - X\hat{A})'(Y - X\hat{A})$

Bayesian estimation is performed using the three priors, namely, Minnesota prior, natural conjugate prior, and non-informative prior which is a special case of natural conjugate prior. These three prior are discussed in detail below:

Minnesota prior

It must be observed that α in Equation 6 contains KM parameters that need to be estimated. Thus, VAR models are nor parsimonious. However, precise estimates can be obtained by using Bayesian VARs with shrinkage priors, commonly known as Minnesota priors. This prior was developed at the Federal Reserve Bank of Minneapolis and the University of Minnesota.[10] Minnesota prior for α is given by:

$$\alpha \sim N(\alpha_{Mn}, V_{Mn}) \tag{9}$$

Since all the four variables are non-stationary, the Minnesota prior uses a prior mean expressing a belief that the individual variables exhibit random walk behavior. Thus, $\alpha_{Mn} = 0_{KM}$ except for the elements corresponding to the first own lag of the dependent variable in each equation which are set to one. The prior covariance matrix V_{Mn} is assumed to be diagonal and is given by:

$$V_{i,jj} = \begin{cases} \frac{a_1}{r^2} & \text{for coefficients of own lag r for r = 1...p} \\ \frac{a_2\sigma_{ii}}{r^2\sigma_{jj}} & \text{for j} \neq \text{i and for r = 1...p} \\ a_3\sigma_{ii} & \text{for coefficients on exogenous variables} \end{cases} \tag{10}$$

where V_i denotes the block of V_{Mn} associated with the K coefficients in equation i and $V_{i,jj}$ denotes its diagonal elements.

Specifying the whole prior covariance matrix V_{Mn} is thus reduced to choosing three scalars, a_1, a_2 and a_3. It must be observed that as lag length increases, $V_{i,jj}$ decrease. This incorporates the information that the confidence that the coefficient is zero is greater for coefficient of longer lags. By setting a_1 greater than a_2, we include information that own lags are likely to be more important predictors than lags of other variables. Further, $\sigma_{ii} = s_i^2$ where s_i^2 is the standard OLS estimate of error variance in i^{th} equation. This ensures scaling to adjust for difference in units of different variables.

In this method, $\sum$ is replaced with the estimate $\hat{\sum}$ obtained using ordinary least squares (OLS) approach. While replacing an unknown matrix of parameters by an estimate rather than integrating it out in a Bayesian fashion is not a convenient assumption, it greatly simplifies computation since it allows posteriors to be analytically computed. We choose optimum hyperparamters based on extensive grid check. The optimum values of a_1, a_2 and a_3 were found to be 0.2, 0.14 and 10^{-4} respectively.

Natural conjugate prior

Natural conjugate priors are those for which the prior, likelihood and posterior come from the same family of distributions. They allow posteriors to be analytically computed. The natural conjugate prior form for the VAR model discussed above is given by:

$$\alpha|\sum \sim N(\underline{\alpha},\sum \otimes \underline{V}) \tag{11}$$

and

$$\sum^{-1} \sim W(\underline{S}^{-1},\underline{\nu}) \tag{12}$$

where $\underline{\alpha},\underline{V},\underline{\nu}$ and $\underline{S}$ are prior hyperparameters. c

Non-informative prior

The non-informative prior or the flat prior does not incorporate any shrinkage in the VAR model. It can be obtained by setting $\underline{\nu} = \underline{S} = \underline{V}^{-1} = cI$ and letting $c \to 0$ in the natural conjugate prior. This choice of these hyperparameters leads to posterior and predictive results which are based on familiar OLS quantities. The drawback of the non-informative prior is that it does not allow for any shrinkage which we have found to be important for VAR modelling. However, since it is a special case of natural conjugate prior, Bayesian estimation and prediction can be calculated analytically.

Evaluation

The forecasting exercise is performed using a rolling estimation window of 10 years (120 months), and projecting the models forward up to 12 steps ahead. The initial estimation window is 1996:7 2006:6, and the initial forecast window is 2006:7-2007:7. The last estimation window is 2005:7 to 2015:6, while the last forecast window is 2015:7 to 2016:7. The driftless random walk given in equation 1 is used as the benchmark model.

We will evaluate our results in terms of the Root Mean Squared Forecast Error (RMSFE) generated by model M when forecasting the exchange rate (vis-a-vis the US Dollar) of currency i at horizon h. Defining $y^M_{i,t+h|t}$ as the h-step ahead forecast of $y_{i,t+h}$ given the information available at time t, the h-step ahead forecast forecast error at time t is:

$$FE^{M}_{h,t} = y^{M}_{i,t+h|t} - y_{i,t+h} \tag{13}$$

and the h-step ahead MSFE is defined as:

$$RMSFE^{M}_{i,h} = \sqrt{1/T_0 \sum_{t=1}^{T_0} (FE^{M}_{h,t})^2} \tag{14}$$

where T_0 is the total number of computed forecasts.

The RMSFE values were calculated after taking the antilog of forecasted log exchange rate values.

Results and Conclusion

Figure below shows the plot of RMSFE values versus the number of steps ahead forecast for different methods discussed above. The values are mentioned in Table 1 in Appendix.

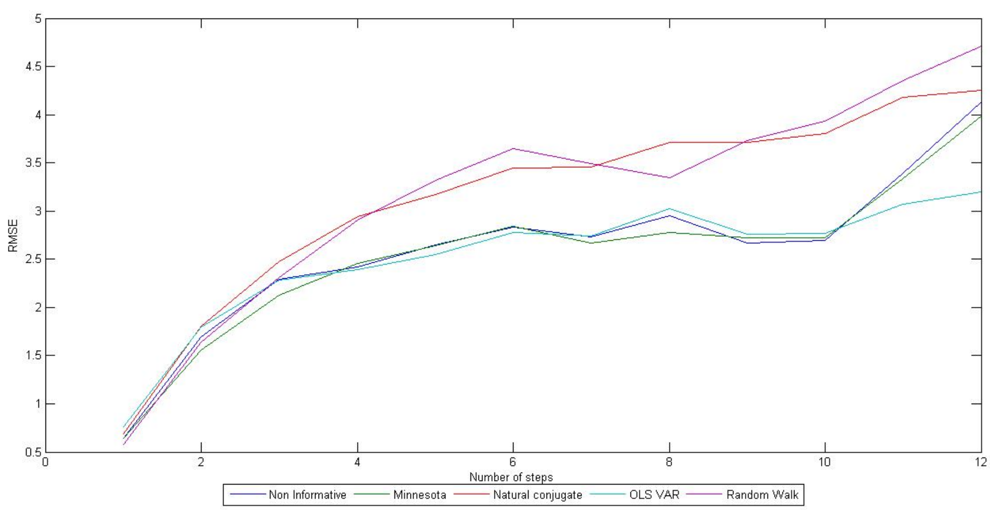

Figure 1: RMSE of different models for different forecat horizons

We observe that BVAR using Minnesota consistently outperforms the benchmark random walk model for forecasts ranging from 1-step to 12-steps ahead. However, VAR estimated using the standard OLS method fails to perform better than the naive random walk model for 1-step and 2-step ahead forecasts. Minnesota BVAR gives better forecast than VAR especially short term forecasts (i.e 1-step to 3-step). Thus, we can conclude that inducing shrinkage through Minnesota resulted in significantly better forecasting performance especially in the short-run. We observe that the performance of BVAR using flat prior (or non-informative prior) is almost similar to the VAR estimated using OLS method. This makes intuitive sense as no additional information or shrinkage is being incorporated in case of a non-informative prior.

However, we observe that BVAR using Normal Wishart prior performs much worser than the almost all the other methods. This could be because the prior co-variance of α in case of natural conjugate (or normal wishart prior) doesnot provide as much flexibility as the Minnesota prior. To explain this possibly undesirable property of this prior, we introduce notation where individual elements of $\sum$ are denoted by σ_{ij}. The fact that the prior covariance matrix in case of this prior has the form $\sum \otimes V$ (which is necessary to ensure natural conjugacy of the prior), implies that the prior covariance of the coefficients in equation i is $\sigma_{ii}V$. This means that the prior covariance of the coefficients in any two equations must be proportional to one another. As, a result it is not possible to incorporate prior information regarding different priors.

Thus, we can conclude that forecasting performance of Rupee-Dollar exchange rate using VAR models can be improved, by inducing shrinkage on the parameters of the model appropriately using the Minnesota prior and Bayesian framework atleast in the short run.

Table 1: RMSE values for different forecast horizons

	Forecast horizon											
Method	1	2	3	4	5	6	7	8	9	10	11	12
Non-Informative	0.638	1.695	2.295	2.421	2.653	2.837	2.732	2.949	2.665	2.695	3.395	4.137
Minnesota	0.640	1.556	2.130	2.457	2.638	2.843	2.666	2.776	2.724	2.725	3.339	3.983
Normal-Wishart	0.689	1.805	2.475	2.938	3.171	3.447	3.457	3.717	3.715	3.800	4.176	4.250
VAR-OLS	0.762	1.798	2.279	2.396	2.549	2.775	2.742	3.023	2.755	2.772	3.067	3.201
Random Walk	0.582	1.641	2.314	2.905	3.314	3.648	3.493	3.348	3.735	3.931	4.350	4.7146

Appendix

References

1. Meese, R. A.; Rogoff, K. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics* **1983**, *14*, 3–24.

2. Beckmann, J.; Schüssler, R. Forecasting exchange rates under parameter and model uncertainty. *Journal of International Money and Finance* **2016**, *60*, 267–288.

3. Engel, C.; Mark, N. C.; West, K. D.; Rogoff, K.; Rossi, B. Exchange Rate Models Are Not as Bad as You Think. *NBER Macroeconomics annual* **2007**, *22*, 381–473.

4. Rossi, P. H.; Wright, J. D.; Anderson, A. B. *Handbook of survey research*; Academic Press, 2013.

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

6. Liu, T.-R.; Gerlow, M. E.; Irwin, S. H. The performance of alternative VAR models in forecasting exchange rates. *International Journal of Forecasting* **1994**, *10*, 419–433.

7. Sarantis, N.; Stewart, C. Structural, VAR and BVAR models of exchange rate determination: a comparison of their forecasting performance. *Journal of Forecasting* **1995**, *14*, 201–215.

8. Dua, P.; Ranjan, R. Modelling and forecasting the Indian RE/US dollar exchange rate. *Centre for Development Economics, Delhi School of Economics Working Paper* **2011**, *197*.

9. Koop, G.; Korobilis, D. Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends® in Econometrics* **2010**, *3*, 267–358.

10. Doan, T.; Litterman, R.; Sims, C. Forecasting and conditional projection using realistic prior distributions. *Econometric reviews* **1984**, *3*, 1–100.