Hacettepe Journal of Mathematics and Statistics
Volume 38 (1) (2009), 85 – 93

USAGE OF DIFFERENT PRIOR DISTRIBUTIONS IN BAYESIAN VECTOR AUTOREGRESSIVE MODELS

Volkan Sevinç* and Gül Ergün†

Received 05 : 06 : 2008 : Accepted 05 : 01 : 2009

Abstract

In Bayesian vector autoregressive models, the Litterman or Minnesota Prior is widely used. However, in some cases, the Minnesota prior is not the best prior distribution that can be used. Thus, other prior distributions can also be applied. In this paper, as well as the Minnesota prior, four other prior distributions have been studied. Based on these prior distributions, five different Bayesian vector autoregressive models have been built to forecast the Turkish unemployment rate and the industrial production index for the two periods of the year 2008. Finally, the five priors have been compared with each other according to the forecasting performances of the models that they are used in.

Keywords: Bayesian vector autoregressive models, Vector autoregressive models, Prior Distributions, Bayes' Theorem, Bayesian approach.

2000 AMS Classification: 37 M 10, 62 F 15, 62 C 10.

1. Introduction

Multivariate time series models such as vector autoregressive (VAR) and Bayesian vector autoregressive (BVAR) models have been widely used in many areas of economics.

BVAR models were first proposed by Litterman [11] as an alternative to the VAR models due to their some advantages such as solving the overparameterization of VAR models and giving better forecasts. Litterman [11] in his study used the Minnesota prior which has been traditionally used in BVAR analysis. Many computer programs that build BVAR models use the Minnesota prior as the default prior distribution. However, although the Minnesota prior is an important prior distribution, it is not the only prior

*Department of Statistics, Muğla University, 48000, Kötekli, Muğla, Turkey. E-mail: volkansevinc@yahoo.com

†Department of Statistics, Hacettepe University, 06800 Beytepe, Ankara, Turkey.
E-mail: gul@hacettepe.edu.tr

distribution that can be used in BVAR analyses. In some cases some other prior distributions could be applied in BVAR models with a better performance. In this paper, as well as the classical Minnesota prior distribution, the four different prior distributions proposed by Kadiyala and Karlsson [9] also have been given and based on these five different prior distributions, five different BVAR model have been developed to forecast the Turkish unemployment rate and the industrial production index. Finally, the comparison of these models in terms of their forecasting performances has been given.

2. Bayesian vector autoregressive models

When fitting macroeconomic models, structural models proposed by Cowles Commission were used until 1970's. However, due to some changes in economic environments and new relations appearing, those models used have become inadequate or invalid. In this case, new models considering the dynamic relations were needed. Sims [13] has proposed vector autoregressive models (VAR) as alternative to complicated structural models and since then, VAR models have been a very important tool in macroeconomic analyses. VAR models remove the constraints arising from the economics theory and use the advantage of multivariate analysis. But it is seen that in large models with many parameters VAR models have a disadvantage of overparameterization problem.

There are two different solutions proposed against the over-parameterization problem which is seen in vector autoregressive models. The first solution is using a model called structural VAR which has theoretical constraints and the second one is using a Bayesian vector autoregressive model (BVAR) which has been introduced by Litterman [11] and become a base for the recent studies.

The BVAR approach starts with the assumption that available data don't involve information at every dimension. This means, in a VAR model which involves too many parameters, some parameters might be different from zero by coincidence. In this case the contribution of the variables matching those parameters would be erroneous. Therefore, the forecasts given by the model built this way would be quite inconsistent. In BVAR approach, this shortcoming is solved by defining proper prior distributions for parameters. The role of the prior distribution can be thought as a barrier preventing the parameters from appearing as nonzero very easily by adding information to them. This barrier set for the parameters by the prior distribution can be broken only when the sample set really provides information.

While developing the BVAR model, Litterman has imposed some assumptions on the unrestricted VAR model given by the following equation.

$$y_t = \mu + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \cdots + \Pi_p y_{t-p} + \varepsilon_t \tag{2.1}$$

The threshold point in Litterman's study was that the series used to estimate the VAR models are unpredictable. This idea can be expressed as each series can be defined as a random walk around an unknown deterministic component. Therefore, the prior distribution for the variable t is focused on the definition of a random walk.

$$y_t - y_{t-1} = c + \varepsilon_t \tag{2.2}$$

The i th equation in the VAR model can be written as below:

$$\begin{aligned} y_{it} = c_i &+ \phi_{i1}^{(1)} y_{1,t} + \phi_{i2}^{(1)} y_{2,t-1} + \cdots + \phi_{in}^{(1)} y_{n,t-1} + \phi_{i1}^{(2)} y_{1,t-2} + \phi_{i2}^{(2)} y_{2,t-2} \\ &+ \cdots + \phi_{in}^{(2)} y_{n,t-2} + \cdots \phi_{i1}^{(p)} y_{1,t-p} + \phi_{i2}^{(p)} y_{2,t-p} + \cdots + \phi_{in}^{(p)} y_{n,t-p} \end{aligned} \tag{2.3}$$

where, $\phi_{ij}^{(s)}$ is the coefficient relating y_{it} to $y_{j,t-s}$

The restriction (2.2) requires that $\phi_{ii}^{1} = 1$ and that all other $\phi_{ij}^{(s)}$ be zero. These 0 or 1 values represent the mean value of the prior distribution for the coefficients. Litterman has chosen the γ value as the standard deviation of the prior distribution. The variance-covariance matrix of the prior distribution was taken as the diagonal matrix.

$$\phi_{ii}^{(1)} \sim N(1, \gamma^2). \tag{2.4}$$

Although each equation $i = 1, 2, \ldots, n$ of the VAR is estimated separately, the same value γ is used for each i. Smaller values of γ mean greater confidence in the prior information. For example, $\gamma = 0.2$ means that, without seeing the data, the researcher is 95% confident that ϕ_{ii}^{1} is not smaller than 0.60 and no greater than 1.40. The coefficients relating y_{it} to further lags are estimated to be zero and Litterman states that confidence in this estimation is greater when the lag is greater. Therefore, he suggested taking

$$\begin{aligned}
&\phi_{ii}^{(2)} \sim N(0, (\gamma/2)^2)\\
&\phi_{ii}^{(3)} \sim N(0, (\gamma/3)^2)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\phi_{ii}^{(p)} \sim N(0, (\gamma/p)^2)
\end{aligned}$$

That is, tightening the prior distribution with a harmonic series for the standard deviation as the lag increases. After the means have been determined, the remaining process is to obtain an estimation of the dispersion around the prior mean. Litterman [12] defines a standard error function for the coefficient of the lth lag of the jth variable in the ith equation as

$$S(i, j, l) = \frac{[\gamma\, g(l)\, f(i, j)]\, s_i}{s_j}. \tag{2.5}$$

3. Prior and posterior distributions for Bayesian vector autoregressive models

Bayesian analysis requires explicit specification of the prior distribution to be used in the analysis. Since, both the prior distribution and sample data are required to obtain the posterior distribution, the choice of the prior distribution depends on the knowledge and experience of the researcher. Some theoretical assumptions are also effective in the choice of the priors. There are a lot of different views about including the prior information in the analysis (Lindley [10], Bernardo [2], Efron [6], Berger and Bernardo [1], Canova [3]).

The main idea in BVAR models is that the model parameters are random variables. The mechanism of this idea is representing the prior information for all the unknown quantities through a prior distribution and combining them with the objective information coming from observations to obtain the posterior distributions. Posterior distributions are obtained by the application of Bayes' Theorem. In general, the choice of the prior distribution depends on the structure of the available information. In this paper, prior distributions such as Minnesota, Diffuse, Normal-Wishart, Normal-Diffuse and Extended Natural Conjugate distributions that are suitable for BVAR models are considered, and their structure and appropriate posterior distributions are given separately in the following subsections. The most widely used prior distribution in BVAR models is the Minnesota prior distribution proposed by Litterman which is based on the Normal distribution.

3.1. The Minnesota Prior. The variance-covariance matrix of the prior Ψ is defined as fixed and diagonal. Therefore, the prior for the ith equation is as given below:

$$\gamma_i \sim N(\widetilde{\gamma}_i, \overline{\Sigma}_i). \tag{3.1}$$

Then using Bayes' Theorem the posterior distribution is given by

$$\gamma_i / y \sim N(\overline{\gamma}_i, \overline{\Sigma})_i, \tag{3.2}$$

where

$$\overline{\Sigma}_i = (\widetilde{\Sigma}_i^{-1} + \psi_{ii}^{-1} \mathrm{Z}'\mathrm{Z})^{-1}$$

and

$$\overline{\gamma}_i = \overline{\Sigma}_i (\widetilde{\Sigma}_i^{-1} \widetilde{\gamma}_i + \psi_{ii}^{-1} \mathrm{Z}' y_i).$$

The diagonal elements ψ_{ii} of Ψ and are obtained from the data.

3.2. The diffuse and Normal–Wishart priors. These priors are the ones that were first proposed by Geisser [7], Tiao and Zellner [14], and the diffuse prior is as follows.

$$p(\gamma, \Psi) \propto |\Psi|^{-(m+1)/2}. \tag{3.3}$$

Using that prior in the BVAR model, the posterior distribution is obtained as

$$\begin{aligned} &\gamma | \Psi, y \sim N(\widehat{\gamma}, \Psi \otimes (\mathrm{Z}'\mathrm{Z}^{-1}) \\ &\quad \Psi | y \sim \operatorname{inv} \mathrm{Wishart}\left((Y - Z\widehat{\Gamma})'(Y - Z\widehat{\Gamma}), T - k\right) \end{aligned} \tag{3.4}$$

The marginal posterior distribution of Γ in the joint posterior distribution is

$$\Gamma | y \sim MT(\mathrm{Z}'\mathrm{Z}, (\Upsilon - \mathrm{Z}\widehat{\Gamma})'(\Upsilon - \mathrm{Z}\widehat{\Gamma}), \widehat{\Gamma}, T - k). \tag{3.5}$$

In the case where variance-covariance matrix for normally distributed data is not known, the unknown parameters would be γ and Ψ. The specification of the joint prior distribution of these two parameters is as follows:

$$f(\gamma, \Psi) = f(\gamma | \Psi) f(\Psi). \tag{3.6}$$

If the assumption of a fixed and diagonal variance-covariance matrix is loosed, the natural joint prior for normal data is the Normal-Wishart distribution

$$\gamma | \Psi \sim N(\widetilde{\gamma}, \Psi \otimes \widetilde{\Omega}), \Psi \sim iW(\widetilde{\Psi}, \alpha). \tag{3.7}$$

The γ parameters need to be estimated, so Ψ is a parameter that can be neglected (a nuisance parameter). The main target is to find the posterior moments of the posterior distribution of γ. The posterior distribution is obtained as follows:

$$\gamma | \Psi, y \sim N(\overline{\gamma}, \Psi \otimes \overline{\Omega}), \Psi | y \sim iW(\overline{\Psi}, T + \alpha), \tag{3.8}$$

where

$$\begin{aligned} \overline{\Omega} &= (\widetilde{\Omega}^{-1} + \mathrm{Z}'\mathrm{Z})^{-1}, \\ \overline{\Gamma} &= \overline{\Omega}(\widetilde{\Omega}^{-1}\widetilde{\Gamma} + \mathrm{Z}'\mathrm{Z}\widehat{\Gamma}), \\ \overline{\Psi} &= \widehat{\Gamma}'\mathrm{Z}'\mathrm{Z}\widehat{\Gamma} + \widetilde{\Gamma}'\widetilde{\Omega}^{-1}\widetilde{\Gamma} + \widetilde{\Psi} + (\Upsilon - \mathrm{Z}\widehat{\Gamma})'(Y - Z\widehat{\Gamma}) - \overline{\Gamma}'(\widetilde{\Omega}^{-1} + Z'Z)\overline{\Gamma}. \end{aligned}$$

The marginal posterior distribution of Γ is again a multivariate t-distribution.

$$\Gamma | y \sim MT(\overline{\Omega}^{-1}, \overline{\Psi}, \overline{\Gamma}, T + \alpha). \tag{3.9}$$

3.3. The Normal-Diffuse prior. This prior distribution was first proposed by Zellner [15]. This prior avoids the constraints that are imposed by the Normal-Wishart prior on the variance-covariance matrix of γ, and allows a non-diagonal variance-covariance matrix. The multivariate normal prior belonging to the regression parameters of the Minnesota prior is combined with a diffuse prior on the residual variance-covariance matrix. That is there is a prior independence between γ and Ψ given by

$$\gamma \sim N(\widetilde{\gamma}, \widetilde{\Sigma}), p(\Psi) \propto |\Psi|^{-(m+1)/2}. \tag{3.10}$$

By an application of Bayes' Theorem in the BVAR model, the marginal posterior distribution of γ is obtained as

$$\begin{aligned} p(\gamma|y) \propto \exp\big\{ -\tfrac{1}{2}(\gamma-\widetilde{\gamma})'\widetilde{\Sigma}^{-1}(\gamma-\widetilde{\gamma})\big\}\big|(\Upsilon - \mathrm{Z}\widehat{\Gamma})'(\Upsilon - \mathrm{Z}\widehat{\Gamma}) \\ + (\Gamma - \widehat{\Gamma})'\mathrm{Z}'\mathrm{Z}(\Gamma - \widehat{\Gamma})\big|^{-T/2} \end{aligned} \tag{3.11}$$

When the Minnesota, Normal-Wishart and Diffuse priors are used, the posterior distributions can be obtained in a closed form. In other words, the multivariate distribution obtained by the multiplication of the prior distribution and the likelihood is in a distribution form which is known. By the help of this distribution, the necessary estimations can be analytically obtained. However, the posterior distribution for the Normal-Diffuse and ENC priors cannot be obtained in a form which allows an analytical process. To overcome this problem, Gibbs sampling, which is one of the MCMC methods, or other numerical methods need to be used.

After some mathematical arrangement the following conditional posterior distribution is obtained.

$$\gamma|\Psi, y \sim N(\overline{\gamma}, (\Sigma^{-1} + \Psi^{-1} \otimes \mathrm{Z}'\mathrm{Z})^{-1}), \tag{3.12a}$$

$$\Psi^{-1}|\gamma, y \sim W\big(\big[(\Upsilon - \mathrm{Z}\widehat{\Gamma})'(\Upsilon - \mathrm{Z}\widehat{\Gamma}) + (\Gamma - \widehat{\Gamma})'\mathrm{Z}'\mathrm{Z}(\Gamma - \widehat{\Gamma})\big]^{-1}, T\big), \tag{3.12b}$$

where,

$$\overline{\gamma} = (\Sigma^{-1} + \Psi^{-1} \otimes \mathrm{Z}'\mathrm{Z})^{-1}\left[\Sigma^{-1}\widetilde{\gamma} + (\Psi^{-1} \otimes \mathrm{Z}'\mathrm{Z})\widehat{\gamma}\right].$$

Using equations (3.12a), (3.12b) and the algorithm given by Geweke [8], which is necessary to make draws from (3.12b), Gibbs sampling can easily be done. The problem is that the equation (3.12a) requires the factorization of a $mk \times mk$ matrix, and taking the inverse of the factor matrix at every step. Therefore, the speed of the algorithm will sharply decrease as the number of parameters in the VAR model increases. Gibbs sampling starts with the generation of Ψ as the least squares estimation of γ from equation (3.12a). A "burning" period of 200 draws is determined (that is the values obtained in this period are not used). Some experiments have shown that Gibbs sampling is insensitive to the choice of initial values for Ψ.

3.4. The Extended Natural Conjugate prior. The extended natural conjugate prior brings about a solution to the constraints on $\mathrm{Var}\,(\gamma)$ of the Normal-Wishart prior. This solution is obtained by the re-parametrization of the VAR equation given below.

$$y_t = \sum_{i=1}^{p} y_{t-i}A_i + x_t C + \varepsilon_t. \tag{3.13}$$

Let Δ be a $mk \times m$ matrix. Let the γ_i 's be the diagonal elements and the other elements be zero,

$$\Delta = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_m \end{bmatrix}.$$

Let also $\Xi = \iota' \otimes \mathrm{Z}$ and let ι be a $m \times 1$ vector of 1's. The equation (3.13) can be written as $\Upsilon = \Xi\Delta + \mathrm{E}$. For the prior distribution below

$$\begin{aligned} &p(\Delta) \propto \left|\widetilde{\Psi} + (\Delta - \widetilde{\Delta})'\widetilde{M}(\Delta - \widetilde{\Delta})\right|^{-\alpha/2}, \\ &\Psi|\Delta \sim iW(\widetilde{\Psi} + (\Delta - \widetilde{\Delta})'\widetilde{\mathrm{M}}(\Delta - \widetilde{\Delta}), \alpha) \end{aligned} \tag{3.14}$$

and normal data, the posterior distribution is given by Dréze and Morales [4] as follows:

$$\begin{aligned} &p(\Delta|y) \propto \left|\overline{\Psi} + (\Delta - \overline{\Delta})'\overline{M}(\Delta - \overline{\Delta})\right|^{-(T+\alpha)/2}, \\ &\Psi|\Delta, y \sim iW(\widetilde{\Psi} + (\Delta - \overline{\Delta})'\mathrm{M}(\Delta - \overline{\Delta}), T + \alpha), \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} &\overline{\mathrm{M}} = \widetilde{\mathrm{M}} + \Xi'\Xi, \\ &\overline{\Psi} = \widetilde{\Psi} + \widetilde{\Delta}'\widetilde{\mathrm{M}}\widetilde{\Delta} + \Upsilon'\Upsilon - \overline{\Delta}'\overline{\mathrm{M}\Delta}, \end{aligned}$$

and

$$\overline{\Delta}, \overline{\mathrm{M}\Delta} = \widetilde{\mathrm{M}}\widetilde{\Delta} + \Xi'\Upsilon.$$

If $\widetilde{M}$ is of full rank, then $\overline{M}$ is also of full rank and $\overline{\Delta}$ is singular. The marginal distribution of Δ has the form of a multivariate t-distribution. However, due to the restricted structure of Δ, it does not have a multivariate t-distribution (Kadiyala and Karlsson [9]).

The Gibbs sampling algorithm for deriving a posterior distribution for the Extended Natural Conjugate prior distribution is based on the lemma of Dréze and Richard [5]. According to this rule the parameters of the i th equation have a multivariate-t distribution conditionally on the parameters of the remaining equations.

$$\gamma_i|\gamma_1, \ldots, \gamma_{i-1}, \ldots, \gamma_m \sim t\left(d_i \frac{q_{ii}}{T + \alpha - k} P_{ii}^{-1}, T + \alpha - k\right). \tag{3.16}$$

Thus, Gibbs sampling is implemented by cycling through equation (3.16) for $i = 1, \ldots, m$. In each cycle, q_{ii} values, the vectors d_i and the matrices P_{ii} need to be calculated. These calculations may be time consuming in large models. The prior means of γ_i are used as the initial values of the Gibbs sampling. A "burning" period for the first 200 draws is used. After the burning process the initial values do not have a real effect on the draws, and Gibbs sampling converges to the real posterior distribution rapidly.

4. A Bayesian vector autoregressive analysis of the Turkish unemployment rate and industrial production index

In this section, using the prior distributions emphasized above, five different BVAR models have been built and forecasts for the Turkish unemployment rate and industrial production index are obtained. The most important reason for choosing the unemployment rate and the industrial production index is that the unemployment rate within a country is a very important economic indicator and therefore that forecasting the Turkish unemployment rate would be a useful task.

The computer program used in this study is an open source program that allows the use of prior distributions other than the Minnesota prior. It is written by Kadiyala and Karlsson [9], and provided by the Stockholm School of Economics for academic use. The program has been run according to the aims of our study and then the forecasts have been obtained. The data used in the study is for the years 1990-2007.

The variables *unemployment rate* and *industrial production index* are used after taking their logarithm in each two-dimensional BVAR model. The data was collected from the records of TÜİK (Turkish Institute of Statistics). The industrial production index is recorded seasonally by the institute. However, although the unemployment rate has been recorded seasonally between the years 2000-2007, it has been recorded biannually before the year 2000. Therefore, the data set has been rearranged as biannual data after the necessary calculation of averages. The unemployment rate and industrial production index for the years 1990:1 – 2007:2 are shown in Figure 1 and Figure 2, respectively.

Figure 1. Turkish unemployment rate for the years 1990:1 - 2008:2

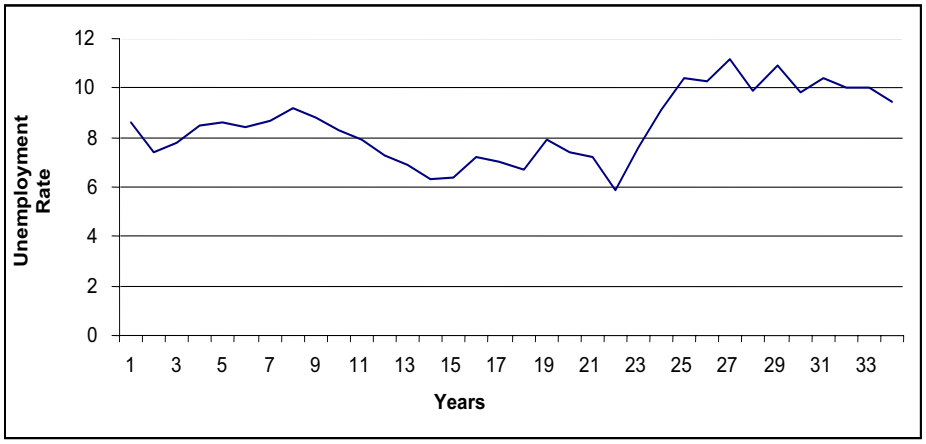
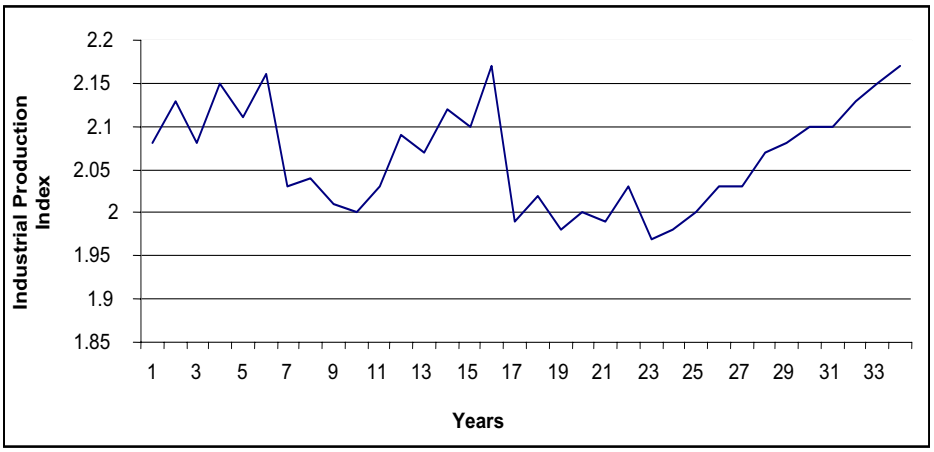


Figure 2. Turkish industrial production index for the years 1990:1 - 2008:2



In the following sections, with the help of the prior distributions mentioned in section 3, two steps ahead forecasts for the Turkish unemployment rate and the industrial production index are given based on four lags. The two steps ahead forecast values belong to the periods 2008:1 and 2008:2. The calculation of the RMSE values is as follows:

$$\text{RMSE} = \sqrt{\sum_{j=0}^{N_k-1} \frac{[F_{t+j+k} + A_{t+j+k}]^2}{N_k}}, \tag{4.1}$$

where

$k = 1, 2, \ldots, 12$: Forecast step,

A_t : The realized value of the exchange rate,

F_t : The forecast value of the exchange rate,

N_t : Total number of k-step ahead forecasts in the projection period for which the realized value of the exchange rate A_t is known.

5. The Bayesian vector autoregressive forecasts based on the five prior distributions

The forecast values for the unemployment rate and Industrial production index obtained from the model parameters updated by the application of Bayes' Theorem using the priors and RMSE values for the periods 2008:1 and 2008:2 are briefly given in Table 1 and Table 2, respectively.

Table 1. BVAR forecasts of the five different prior distributions and RMSE values for the Turkish unemployment rate

	2008:1		2008:2	
	UnemploymentRate	RMSE	UnemploymentRate	RMSE
Minnesota	10.14	1.07	9.74	1.65
Diffuse	9.24	1.66	7.94	2.14
Normal-Wishart	9.49	1.11	8.57	1.67
Normal-Diffuse	10.15	1.13	9.78	1.77
ENC	10.09	1.08	9.64	1.65

Table 2. BVAR forecasts of the five different prior distributions and RMSE values for the Turkish industrial production index

	2008:1		2008:2	
	ProductionIndex	RMSE	ProductionIndex	RMSE
Minnesota	2.07	0.58	2.12	0.82
Diffuse	2.09	0.88	2.13	4.65
Normal-Wishart	2.06	0.49	2.10	0.62
Normal-Diffuse	2.07	0.64	2.12	0.85
ENC	2.07	0.59	2.12	0.83

It can be seen from the tables that the results for the five different prior distributions are quite close to one another. In addition, the RMSE values of the four prior distributions other than the Diffuse prior have similar values. However, the Diffuse prior, which is a noninformative prior, gives the highest RMSE values. That means that it has the worst performance for both the unemployment rate and the industrial production index, as expected. It is remarkable that, although the Normal-Wishart prior has given the same forecast values for the industrial production index as the Minnesota prior, its RMSE value is less than one of the Minnesota prior's. This means that it displayed a better performance.

6. Conclusion

In this study it is pointed out that the BVAR models, which are proposed as an alternative to the VAR models, can be built not only based on the Minnesota prior but also on other prior distributions. The modeling application in the study shows that each prior distribution proposed can be an alternative to the Minnesota prior. The choice of the prior distribution that is to be used in the BVAR analysis involves certain criteria, such as choosing an informative prior distribution and a prior distribution which leads to a posterior distribution that can be analytically obtained.

References

[1] Berger, J. O. and Bernardo, J. M. *On the development of reference priors*, In: Bernardo, J. M. *et al.* (Bayesian Analysis IV, Oxford University Press, Oxford, 1992).

[2] Bernardo, J. M., *Reference posterior distributions for Bayesian*, Inference, Journal of Royal Statistical Society Series **B41**, 113–147, 1979.

[3] Canova, F. *Methods for Applied Research 9* (Introduction to Bayesian Methods, Universitat Pompeu Fabra, 2004).

[4] Dréze, J. H. and Morales, J. A. *Bayesian full information analysis of simultaneous equations*, Journal of the American Statistical Association **71**, 919–23. Reprinted in A. Zellner, ed. *Bayesian Analysis in Econometrics and Statistics* (North-Holland, Amsterdam, 1980).

[5] Dréze, J, H. and Richard, J. F. *Bayesian analysis of simultaneous equation systems* in Z. Griliches and M. D. Intrilligator eds. (Handbook of Econometrics, Vol. I, North-Holland, Amsterdam, 1980).

[6] Efron, B. *Why isn't everyone a Bayesian?*, American Statistician **40**, 1–11, 1986.

[7] Geisser, S. *Bayesian estimation in multivariate analysis*, Annals of Mathematical Statistics **36**, 150–159, 1965.

[8] Geweke, J. *Antithetic acceleration of Monte Carlo integration in Bayesian inference*, Journal of Econometrics **38**, 73–89, 1988.

[9] Kadiyala, K. R. and Karlsson, S. *Numerical methods for estimation and inference in Bayesian VAR models*, Journal of Applied Econometrics **12**, 99–132, 1997.

[10] Lindley, D. V. *The Bayesian approach*, Scandinavian Journal of Statistics **5**, 1–26, 1978.

[11] Litterman, R. B. A Bayesian Procedure for Forecasting with Vector Autoregressions (Mimeo, Massachussets Institute of Technology, 1980).

[12] Litterman, R. B. *Forecasting with Bayesian vector autoregressions - Five years of experience*, Journal of Business and Economic Statistics **4**, 25–38, 1986.

[13] Sims, C. A. *Macroeconomics and reality*, Econometrica **48** (1), 1–48, 1980.

[14] Tiao, G. C. and Zellner, A. *On Bayesian estimation of multivariate regression*, Journal of the Royal Statistical Society **B26**, 389–99, 1964.

[15] Zellner, A. *An Introduction to Bayesian Inference in Econometrics* (John Wiley, New York, 1974).