Forecasting several North Dakota macroeconomic variables-a Bayesian vector autoregressions approach [1]

by

Feng Yao
Economics Department
West Virginia University
Morgantown, WV 26505, U.S.A.
email: feng.yao@mail.wvu.edu
Voice: + 1 304 293 7867
Fax: + 1 304 293 2233
February, 2011

Abstract: We discuss and implement the Bayesian vector autoregressions approach proposed by Litterman (1980) to estimate and forecast several macroeconomic variables in North Dakota, including Employment, Personal Income and Tax receipts. Using data from 1^{st} quarter of 1998 to 3^{rd} quarter of 2005 as hold out sample, we evaluate and compare its out of sample forecasting performance together with Vector Autoregressions. The results confirm our belief that properly incorporating the prior information into the Bayesian vector autoregressions deliver accurate and responsive forecasts.

Keywords and Phrases: Bayesian Vector Autoregressions, flat prior, forecast

JEL classifications. C22;C53

[1] We would like to thank Professor Lloyd Blackwell for comments and suggestions. The work is supported by Bureau of Business and Economic Research in University of North Dakota.

1 Introduction

Forecasting macroeconomic variables in North Dakota has been of interest to policy makers, private business owners and researchers. It is closely related to forecast macroeconomic variables in general, which has been the focus of intense research in the past two decades. However, facing the idiosyncrasy of business cycle and the unpredictable influence from the weather, model based forecasting has been a risky task, see Litterman (1986). Admitting the uncertainty and allowing data to reveal the underlying economy characteristics, Bayesian Econometrician has developed the Bayesian Vector Autoregressions model (BVAR)in forecasting macroeconomic variables.

When there are macroeconomic relationship derived from economic theory, it makes sense to incorporate it into the estimation process via structure models. However, as pointed out by Sims(1980), the identification claimed for existing large scale models is incredible. Following the criticism on structural equation, there has been a trend to model economic systems using joint time series behavior of the variables. Litterman (1980, 1986) appears to be the first attempt to use BVAR to forecast macroeconomic variables using flat priors, without considering the interdependence across equations. Doan, Litterman and Sims (1986) develop conditional projection using BVAR method. Kadiyala and Karlsson (1993) propose generalized BVAR by incorporating the dependence of equations into BVAR, and compare Normal-Wishart prior, diffuse prior with the flat prior in the specification of BVAR. Sim and Zha (1998) provides up to date BVAR methods using both structural and reduced form equation, where forecasts are provided with error bands. See Geweke and Whiteman (2006) for comprehensive review of BVAR and see also Stock (2006) for alternative modeling using dynamic factor models.

In this paper, we discuss the BVAR to estimate and forecast several macroeconomic variables in North Dakota, including employment, personal income and tax receipts. We provide detailed implementation procedures and compare its performance with Vector Autoregressions. The original BVAR proposed by Litterman (1986) is easy to use and popular in practice, so we focus mainly on this approach. In situation where the dependence across equation is prominent, we expect forecasting performance to be improved by adopting other developments mentioned above. Besides this introduction, we describe the Bayesian forecasting rule in Section 2, detail the BVAR model in section 3, provide an empirical illustration in Section 4, conclude in Section 5 and defer the tables and figures in Appendix in Section 6.

2 Bayesian Forecasting Rule

Following Geweke and Whiteman(2006), we assume model A is specified by finite dimensional parameters θ_A, a $k \times 1$ vector of unobservables, and $\theta_A \in \Theta_A \subseteq \Re^K$. y_t is a $p \times 1$ vector of observations, where $t = 1, \cdots, T$, and $Y_t = \{y_s\}_{s=1}^t$ is the history of the sequence y_t at time t. Let the conditional probability density function of y_t given Y_{t-1}, θ_A and model A be $p(y_t|Y_{t-1}, \theta_A, A)$.

The ultimate interest is represented by $\omega \in \Omega \subseteq \Re^q$ for $q > 0$, so correspondingly we would like to learn $p(\omega|Y_T, A)$, the conditional density of ω given observed data Y_T and model A. Specifically, for $p = 1$, if we are forecasting univariate macroeconomic variable y_t, $\omega' = (y_{T+1}, \cdots, y_{T+q})'$ and our interest is basically focused on the posterior predictive distribution $p(\omega|Y_T, A)$.

Since $p(\omega|Y_T, A) = \int_{\theta_A} p(\theta_A|Y_T^o, A)p(\omega|\theta_A, Y_T^o, A)d\theta_A$ and we could formulate,

$$p(\theta_A|Y_T^o, A) = \frac{p(\theta_A, Y_T^o|A)}{p(Y_T^o|A)} = \frac{p(\theta_A|A)p(Y_T^o|\theta_A, A)}{p(Y_T^o|A)} \propto p(\theta_A|A)p(Y_T^o|\theta_A, A)$$

$p(\theta_A|Y_T^o, A)$ is called the posterior density, which for given data Y_T^o, is determined by the product of $p(\theta_A|A)$, the prior distribution, and $p(Y_T^o|\theta_A, A)$, the likelihood distribution.

The Bayesian decision problem is to find action α, controlled by the decision maker, such that the Bayesian risk function is minimized, i.e.,

$$\hat{\alpha} = \arg_\alpha \min E[L(\alpha, \omega)|Y_T^o, A] = \arg_\alpha \min \int_\Omega L(\alpha, \omega)p(\omega|Y_T^o, A)d\omega$$

For quadratic loss function $L(\alpha, \omega) = (\alpha - \omega)'Q(\alpha - \omega)$ with some positive definite matrix Q, by Theorem 2.4.1 of Geweke (2005), we know that $\hat{\alpha}(Y_T^o, A) = E(\omega|Y_T^o, A)$. Though quadratic loss function is symmetric, this criteria is simple to implement and we adopt it for its popularity. It is worth mentioning that asymmetric loss function are available, which implies that the clients do need to examine whether his loss function is well approximated by the one we assume here.

Based on the Bayesian approach described above, we adopt the following updating rule:

1.Calculate the posterior density $p(\theta_A|Y_T^o, A) \propto p(\theta_A|A)p(Y_T^o|\theta_A, A)$. Often simulation would enable this step if the distribution is unconventional.

2. Provide $p(\omega|Y_T^o, A) = p(y_{T+1}, \cdots, y_{T+q}|Y_T^o, A) = \int_{\theta_A} p(\theta_A|Y_T^o, A)p(\omega|\theta_A, Y_T^o, A)d\theta_A$.

3. Obtain $\hat{\alpha}(Y_T^o, A) = \hat{E}(\omega|Y_T^o, A)$.

Repeat steps $1 - 3$ when Y_T^o is updated to, i.e., Y_{T+q}^o.

3 Bayesian Vector Autoregressions

The updating rule is implemented using the Bayesian Vector Autoregressions approach (BVAR) proposed by Litterman (1986). We specify the VAR model as

$$y_t = \theta_0 + \theta_1 y_{t-1} + \cdots + \theta_m y_{t-m} + \epsilon_t$$

where y_t is $P \times 1$, θ_0 is $P \times 1$, θ_l is $P \times P$ with $l = 1, \cdots, m$, ϵ_t is $P \times 1$ such that

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Pt} \end{bmatrix}, \theta_0 = \begin{bmatrix} \theta_{10} \\ \theta_{20} \\ \vdots \\ \theta_{P0} \end{bmatrix}, \theta_l = \begin{bmatrix} \theta_{l,11} & \theta_{l,12} & \cdots & \theta_{l,1P} \\ \theta_{l,21} & \theta_{l,22} & \cdots & \theta_{l,2P} \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{l,P1} & \theta_{l,P2} & \cdots & \theta_{l,PP} \end{bmatrix}, \epsilon_t = \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{Pt} \end{bmatrix}$$

To be more specific, the i^{th} equation for $i \in \{1, 2, \cdots, P\}$ is expressed as

$$\begin{aligned} y_{it} = \theta_{io} & + \theta_{1,i1}y_{1,t-1} + \theta_{1,i2}y_{2,t-1} + \cdots + \theta_{1,iP}y_{P,t-1} \\ & + \theta_{2,i1}y_{1,t-2} + \theta_{2,i2}y_{2,t-2} + \cdots + \theta_{2,iP}y_{P,t-2} \\ & \vdots \\ & + \theta_{m,i1}y_{1,t-m} + \theta_{m,i2}y_{2,t-m} + \cdots + \theta_{m,iP}y_{P,t-m} \end{aligned}$$

$$+ \quad \epsilon_{it}$$

Before we implement BVAR, let's state a result in restricted OLS regression. Suppose we have the restriction on the parameters

$$R_i \theta^i = r_i + v_i$$

where R_i is $(mP+1)\times(mP+1)$, and of rank $(mP+1)$ in general. $\theta^i = \{\theta_{i0}, \theta_{1,i1}, \cdots, \theta_{1,iP}, \cdots, \theta_{m,i1}, \cdots, \theta_{m,iP}\}'$, i.e., the i^{th} element or row of θ_l, $l = 0, 1, \cdots, m$. r_i is $(mP+1)\times 1$ vector as the prior mean of θ^i, and v_i is a random variable such that $v_i \sim N(0, \Omega_i)$.

With arbitrary restriction as above for equation i, we could run restricted OLS regression. Let $y_i = \{y_{it}\}_{t=1}^T$, $ones((T-m+1)\times 1)$ be a $(T-m+1)$ vector of ones, call

$$\begin{aligned} X = \{ones((T-m+1)\times 1), \quad & \{y_{1,t-1}\}_{t=m+1}^T, \quad \{y_{2,t-1}\}_{t=m+1}^T, \cdots \{y_{P,t-1}\}_{t=m+1}^T, \\ & \{y_{1,t-2}\}_{t=m+1}^T, \quad \{y_{2,t-2}\}_{t=m+1}^T, \cdots \{y_{P,t-2}\}_{t=m+1}^T, \\ & \vdots \\ & \{y_{1,t-m}\}_{t=m+1}^T, \quad \{y_{2,t-m}\}_{t=m+1}^T, \cdots \{y_{P,t-m}\}_{t=m+1}^T\} \end{aligned}$$

Let $\epsilon_i = \{\epsilon_{it}\}_{t=1}^T$. Upon imposing the restrictions, we have

$$\begin{aligned} y_i &= X\theta^i + \epsilon_i \\ r_i &= R_i\theta^i - v_i \end{aligned}$$

Or, in matrix notation

$$\begin{bmatrix} y_i \\ r_i \end{bmatrix} = \begin{bmatrix} X \\ R_i \end{bmatrix} \theta^i + \begin{bmatrix} \epsilon_i \\ -v_i \end{bmatrix} \tag{1}$$

Since $E(v_i v_i') = \Omega_i$, and we assume $E(\epsilon_i \epsilon_i') = \Sigma_i$, ϵ_i and v_i are uncorrelated, we obtain

$$E\left(\begin{pmatrix}\epsilon_i\\ v_i\end{pmatrix}(\epsilon_i'\,v_i')\right)=\begin{pmatrix}\Sigma_i & 0\\ 0 & \Omega_i\end{pmatrix} \tag{2}$$

For positive definite variance-covariance matrices Σ_i and Ω_i, we have

$$\left(E\left(\begin{pmatrix}\epsilon_i\\ v_i\end{pmatrix}(\epsilon_i'\,v_i')\right)\right)^{-1}=\begin{pmatrix}\Sigma_i^{-1} & 0\\ 0 & \Omega_i^{-1}\end{pmatrix} \tag{3}$$

So the Generalized Least Square or Mixed estimator by Theil and Goldberger(1961) is given by the following transformation

$$\begin{pmatrix}\Sigma_i^{-1/2} & 0\\ 0 & \Omega_i^{-1/2}\end{pmatrix}\begin{bmatrix}y_i\\ r_i\end{bmatrix}=\begin{pmatrix}\Sigma_i^{-1/2} & 0\\ 0 & \Omega_i^{-1/2}\end{pmatrix}\begin{bmatrix}X\\ R_i\end{bmatrix}\theta^i+\begin{pmatrix}\Sigma_i^{-1/2} & 0\\ 0 & \Omega_i^{-1/2}\end{pmatrix}\begin{bmatrix}\epsilon_i\\ -v_i\end{bmatrix} \tag{4}$$

$$\begin{aligned}\Rightarrow \hat{\theta}^i &= \left(\begin{bmatrix}\Sigma_i^{-1/2}X\\ \Omega_i^{-1/2}R_i\end{bmatrix}'\begin{bmatrix}\Sigma_i^{-1/2}X\\ \Omega_i^{-1/2}R_i\end{bmatrix}\right)^{-1}\begin{bmatrix}\Sigma_i^{-1/2}X\\ \Omega_i^{-1/2}R_i\end{bmatrix}'\begin{pmatrix}\Sigma_i^{-1/2}y_i\\ \Omega_i^{-1/2}r_i\end{pmatrix}\\ &= (X'\Sigma_i^{-1}X+R_i'\Omega_i^{-1}R_i)^{-1}(X'\Sigma_i^{-1}y_i+R_i'\Omega_i^{-1}r_i)\end{aligned}$$

We use above result to implement our BVAR method. Adopting Litterman (1986)'s prior, we assume a reasonable approximation of the behavior of an economic variable is a random walk around an unknown, deterministic component. For equation i, y_{it} is centering around

$$y_{it}=y_{i,t-1}+\theta_{i0}+\epsilon_{it}$$

The prior will reflect the following belief

a. The coefficients are having prior mean of zeroes except the first lag of the dependent variable, which has mean equal to one.

b. The parameters are uncorrelated. The more back into the past, the smaller the standard deviation of the parameters.

c. The prior standard deviation of the dependent variable should be larger, which implies the parameters for other variables in the equation is believed to center more tightly around zero.

We let r_i be a vector of zeros except for an one corresponding to the 1^{st} lag of variable i, i.e., in our order of the coefficient, for the restriction on $\theta_{1,ii}$ there will an one in the r_i vector, which is our restriction in a. R_i is a $(mP+1)\times(mP+1)$ diagonal matrix. The $(1,1)$ position is specified to be $\frac{\lambda}{\gamma_3 \hat{\sigma}_i}$. Obviously for $k>1$, $k=1+lj$ in equation i and $k \in \{1,2,\cdots,mP+1\}$, the k^{th} row of R_i imposes restriction on the k^{th} element of θ^i, i.e., $\theta_{l,ij}$ as the l^{th} lag of variable j in equation i. The k^{th} element in row k is set to be $\frac{\lambda}{\delta_{ij}^l}$, where

$$\delta_{ij}^l = \begin{cases} \frac{\lambda}{l^{\gamma_1}} & i=j \\ \frac{\lambda\gamma_2\hat{\sigma}_i}{l^{\gamma_1}\hat{\sigma}_j} & i\neq j \end{cases} \tag{5}$$

The hyperparameters in above BVAR are

- λ: the overall tightness
- γ_1: the decay rate
- γ_2: others' weight, $0<\gamma\leq 1$
- γ_3: the weight assigned to the constant
- $\hat{\sigma}_i$: estimated residual standard deviation in unrestricted OLS of equation i from univariate autoregression of y_{it} on its own m lags

λ is the prior standard deviation of the ii^{th} element in the parameter matrix θ_1, reflecting how closely the random walk approximation is to be imposed. As λ approaches 0, the diagonal element of θ_1 goes to 1 more tightly, and correspondingly the off-diagonal terms tend to 0.

$\gamma_1>0$: determines the extent to which coefficients on lags beyond the 1^{st} one are likely to be different from zero. With larger γ_1, coefficients on higher order lags are shrunk towards zero more tightly. When $\gamma_1=1$, the rate of decay in the weight is called harmonic.

γ_2: we expect most of the variations in each of the dependent variables in VAR is accounted for by its own lag. So we assign to, in each row of θ_l, coefficients of lags of other variables smaller variance. As γ_2 approaches zero, the off-diagonal elements of θ_l go to zero. For $\gamma_2 = 1$, there is no distinction between the coefficients of lags of dependent variables and those of lags other variables.

$\gamma_3 \hat{\sigma}_i$ will be the weight or standard deviation assigned to the constant in equation i. The prior mean for the constant is zero. As γ_3 approaches 0, the constant coefficient is shrunk toward the prior mean.

$\frac{\sigma_i}{\sigma_j}$: accounts for the differences in the units of measurement of different variables. If the variance of $y_{i,t}$ is far smaller than the variance of $y_{j,t}$, then the coefficients on $y_{i,t-1}, y_{i,t-2}, \cdots, y_{i,t-m}$ in the i^{th} equation is shrunk toward zero. As mentioned above, we use the $\hat{\sigma}_i$ to estimate this hyperparameter.

For $v_i \sim N(0, \lambda^2 I_{mP+1})$, $\Omega_i = \lambda^2 I_{mP+1}$. With the restriction or the prior we impose above, θ^i is believed to have mean r_i and variance as $R_i^{-1} \Omega_i R_i^{-1}$. Specifically, $\theta_{1,ii}$ is having mean$= 1$ and all other parameters have mean zero. The constant θ_{i0} have variance $(\gamma_3 \hat{\sigma}_i)^2$. The parameter in the i^{th} equation for the j^{th} variable with l-lag, $\theta_{l,ij}$, have variance $(\delta_{ij}^l)^2$.

Our specification is consistent with Litterman (1986) equation (7), with the $\hat{b}_i^{LT}$ on page 9 of Robertson and Tallman (1999), and with equation (58) of Geweke and Whiteman (2006). It is worth mentioning that the flat prior we impose here is also called improper prior (Litterman 1986, Zellner 1971) since the prior $p(\theta_A | A)$ does not have a proper probability distribution. Essentially, we are letting the coefficients to be random variables, and we only restrict the first and the second moments of the coefficients. Furthermore, the variance of the error term ϵ is approximated by univariate OLS autoregression, all of which are the areas that have been improved in the later literatures, see Geweke and Whiteman 2006. However, without observing repeated data, very little is known about the underlying deterministic components in the system. The use of noninformative prior just represents the ignorance, and enables the data to display more characteristics for the underlying model. As mentioned in Litterman (1986), the justification of this prior is that through its use, "we are able to express more realistically the true state of our knowledge and the uncertainty about the structure of the economy". It is true that we still face the task of choosing the hyperparameters in the prior, but as demonstrated in Litterman (1986) and Robertson and Tallman (1999), the performance of BVAR is not very sensitive to the choice of hyperparameters.

In Doan, Litterman and Sims (1984) and Sims (1992), several other types of inexact prior information have been introduced as restrictions on the linear combination of the coefficients in the VAR system. Specifically, following Doan et al. (1984) they allow for a VAR model that contains stochastic trends or unit root in the first differences of the data, which is represented by restriction in the VAR model that the sums of the coefficients on the lags of dependent variables equal one while the sums of the coefficients on other variables are zero. In Sims (1992), we could allow the number of stochastic trends in VAR model to be less than the number of equations in the system. Implementation of the two types of priors based on Litterman (1986) is straightforward by adding dummy variables and specifying two additional hyperparameters μ_5, μ_6 as described in Robertson and Tallman (1999). For this reason, we call this modified Litterman approach and include it into our data illustration section. For comparison purpose, we also include the VAR without any prior information as the benchmark approach.

4 Empirical Illustration

We include the following three variables in the VAR

- Employment: total number of monthly employment in thousands in North Dakota's nonfarming sectors. The data is ranging from January 1982 to September 2005 obtained from the website of Bureau of Labor Statistics and is seasonally adjusted. To transform employment into quarterly variable, we simply take average of the three months in a quarter.

- Income: quarterly personal income in millions of dollars received by all persons from all sources in North Dakota from 1^{st} quarter of 1982 to the 3^{rd} quarter of 2005. The data is collected from Bureau of Economic Analysis. By definition provided on the website, personal income is the sum of net earnings by place of residence, rental income of persons, personal dividend income, personal interest income, and personal current transfer receipts. Net earnings is earnings by place of work (the sum of wage and salary disbursements (payrolls), supplements to wages and salaries, and proprietors' income) less contributions for government social insurance, plus an adjustment to convert earnings by place of work to a placeofresidence basis. Personal income is measured before the deduction of personal income taxes and other personal taxes and is reported in current dollars (no adjustment is made for price changes). Quarterly estimates are expressed at seasonally adjusted annual rates.

- Tax: quarterly North Dakota state sales and use tax receipts, expressed in thousands of dollars and ranging from 1^{st} quarter of 1982 to the 3^{rd} quarter of 2005. It is not seasonally adjusted, however, it seems seasonality is not a prominent feature of this data series.

To give the audience a better understanding about the three macroeconomic variables of North Dakota, we provide several descriptive statistics in Table 1.

Table 1 Descriptive Statistics

	Mean	Standard Dev.	Skewness	Kurtosis	Minimum	Maximum
Employment	290.57	33.41	0.108	-1.611	248.03	345.93
Income	12355.80	3686.78	0.43	-1.02	6908.00	19952.00
Tax	68849.02	26536.98	0.31	-0.26	16739.33	127931.31

To implement the Bayesian VAR methods as in last section, we need to choose the hyperparameters, the quality of which would definitely influence the forecasting performance. As mentioned in Robertson and Tallman (1999), in practice, the values are determined based on examining historical forecast performance across a range of parameter settings. To make results comparable with the past literature, we pick the values as in Litterman (1986) and Robertson and Tallman (1999). Specifically, we set $\lambda = 0.1$, $\gamma_1 = 1$ for harmonic decay rate, $\gamma_2 = 0.2$, $\gamma_3 = 0.2$, $\mu_5 = 5$ and $\mu_6 = 5$. Due to relatively small sample size considered, we set the number of lags in VAR to be 4. We call the three approaches that we described in last section as Litterman, modified Litterman, VAR for the reasons thereforth. In total we have 95 quarterly observations on each variables from 1^{st} quarter of 1982 to 3^{rd} quarter of 2005. We use the first 64 observations up to the

4^{th} quarter of 1997 to start the forecasting one to four period ahead using Chain Rule of Forecasting. Then we increase the information set forward one period and carry through the same forecasting procedure. So to evaluate the out of sample forecasting performance based on 1^{st} quarter of 1998 to 3^{rd} quarter of 2005, we generate 31 one-step ahead forecasts, 30 two-step ahead forecasts, 29 three-step ahead forecasts, and 28 four-step ahead forecasts. The performance of each method in terms of Bias and Root Mean Squared Error (RMSE) is summarized in Table 2 for Employment, Table 3 for Income and Table 4 for Tax.

The message from the empirical illustration is clear. As we expect, for all three methods considered in forecasting the three North Dakota macroeconomic variables, the less step ahead forecasting we make, the better the performance in terms of both Bias and RMSE. This basically confirms the belief that the remote past is providing less valuable information for conditional forecasting than recent record. In general in forecasting all variables and all four step ahead forecasting considered, we observe Bayesian VAR with prior information performs better than VAR, while modified VAR is better than Litterman (1986), with exceptions in Bias of Employment and tax. This reflects that properly incorporated information in BVAR improves the performance relative to VAR, as has been discussed in Litterman (1986) and Robertson and Tallman (1999). It is also worth noting that further incorporation of restriction on allowing the stochastic trends in differenced data improves the performance even more. For example, in terms of the one step forecast, using two BVAR methods all produce forecasts that carry bias less than 5% of the magnitude of the data. The RMSE is reduced by at least 10% with the BVAR methods. All three methods considered, they are slightly overeforecasting employment and underestimate the income and tax variables.

To provide visual illustration of the results, we provide in Figure 1-3 the time series plot of true data in the out of sample forecast period together with the one-step ahead forecast produced by the three methods. All three variables exhibits upward trends in general in the time series plot. The two BVAR methods appear to perform quite similarly in forecasting three variables and the deviation of VAR from the true is in general larger than the two BVAR methods. As is clear in Figure 3, where tax receipts show more volatility, the two BVAR forecasts seem to follow the change much closely, while VAR method is relatively not responsive to the change.

5 Conclusion

In this paper, we discuss in detail and implement the Bayesian VAR methods proposed in Litterman (1986) to estimate and forecast several North Dakota macroeconomic variables, including Employment, Income and Tax receipts. Using 1^{st} quarter of 1998 to 3^{rd} quarter of 2005 as hold-out sample, we evaluate the out of sample performance of the BVAR methods and compare them with Vector Autoregression models. The forecasting performance confirms our belief that properly incorporating prior information in the BVAR methods improve the performance and thus we expect BVAR delivers relatively accurate and responsive forecasts for the macroeconomic variables in North Dakota.

6 Appendix

Table 2 Bias and Root Mean Squared Error for Employment				
	1-step ahead	2-step ahead	3-step ahead	4-step ahead
	Bias			
Litterman	0.370	0.781	1.181	1.608
Modified Litterman	0.401	0.847	1.293	1.773
VAR	0.297	0.528	0.925	1.473
	Root Mean Squared Error			
Litterman	1.347	1.821	2.343	3.060
Modified Litterman	1.391	1.969	2.616	3.450
VAR	1.703	2.482	3.255	4.352

Table 3 Bias and Root Mean Squared Error for Income				
	1-step ahead	2-step ahead	3-step ahead	4-step ahead
	Bias			
Litterman	-46.721	-81.435	-132.719	-170.247
Modified Litterman	-52.816	-94.990	-156.901	-206.323
VAR	-140.187	-277.977	-400.682	-498.529
	Root Mean Squared Error			
Litterman	286.050	419.779	509.469	563.633
Modified Litterman	293.333	439.263	543.053	613.100
VAR	327.068	485.274	583.926	675.761

Table 4 Bias and Root Mean Squared Error for Tax				
	1-step ahead	2-step ahead	3-step ahead	4-step ahead
	Bias			
Litterman	-830.671	-1858.038	-3232.1165	-3977.065
Modified Litterman	-2505.105	-4873.956	-7335.754	-9045.967
VAR	-3584.375	-4822.861	-4636.523	-6403.689
	Root Mean Squared Error			
Litterman	8868.614	9076.448	8333.182	9996.734
Modified Litterman	8749.592	9719.642	10239.011	12269.889
VAR	9602.945	10264.180	9708.982	11830.808

FIGURE 1 ONE STEP AHEAD EMPLOYMENT FORECASTS FROM 1998.1 TO 2005.3

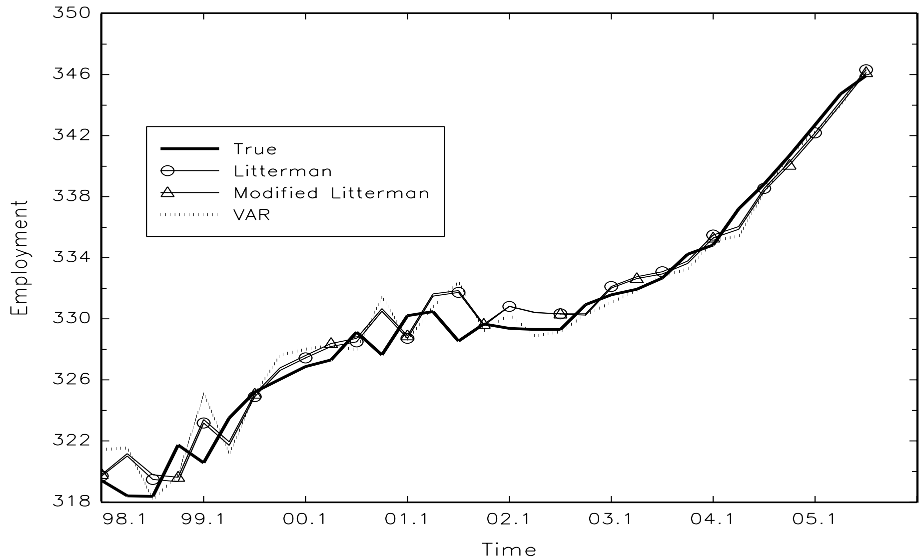


FIGURE 2 ONE STEP AHEAD INCOME FORECASTS FROM 1998.1 TO 2005.3

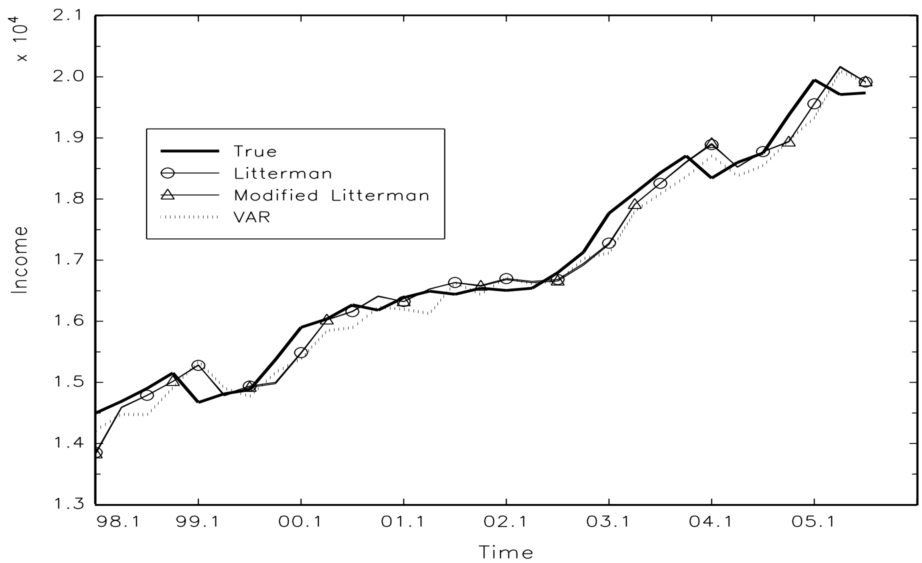
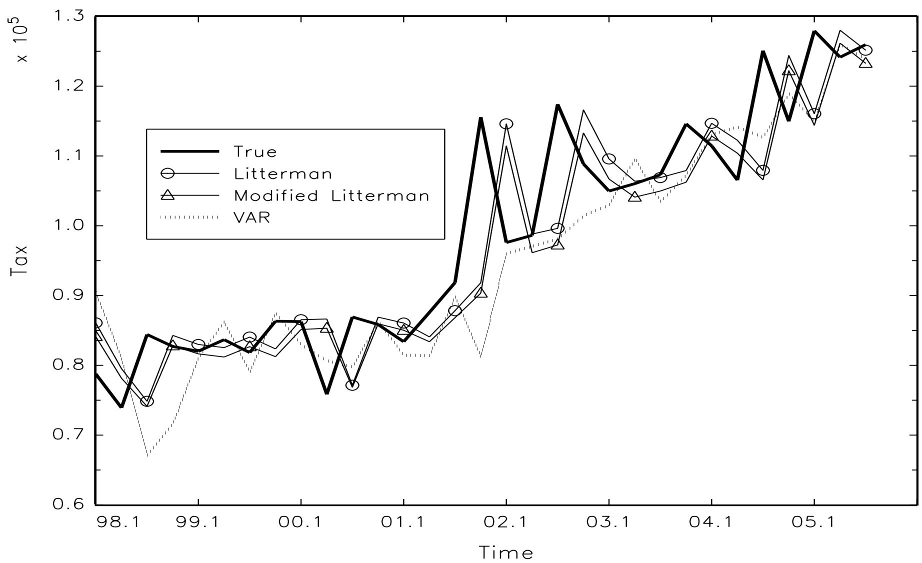


Figure 3 one step ahead tax forecasts from 1998.1 to 2005.3



References

[1] Doan, T., R. B. Litterman, and C. A. Sims 1984, Forecasting and Conditional Projection Using Realistic Prior Distributions, *Econmetric Reviews,* 3: 1-100

[2] Geweke, J. 2005, *Contemporary Bayesian Econometrics and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey*

[3] Geweke, J. and C. Whiteman 2006, Bayesian Forecasting. In G. Elliott, C. Granger, and A. Timmermann, eds., Handbook of Economic Forecasting, Amsterdam: North-Holland.

[4] Kadiyala, K. R. and S. Karlsson 1993, Forecasting with Generalized Bayesian Vector Autoregressions, *Journal of Forecasting,* 12, no. 4:365-78

[5] Litterman, R. B. 1980, A Bayesian Procedure for Forecasting with Vector Autoregression, Massachusetts Institute of Technology, Department of Economics Working Paper

[6] Litterman, R. B. 1986, Forecasting with Bayesian Vector Autoregressions-Five Years of Experience. *Journal of Business & Economic Statistics,* Vol.4, No.1.

[7] Robertson, J.C. and E. W. Tallman, 1999, Vector Autoregressions: Forecasting and Reality. *Federal Reserve Bank of Atlanta Economic Review,* 1st Quarter

[8] Sims, C.A. 1992, A Nie-Variable Probabilistic Macroeconomic Forecasting Model, in *Business Cycles, Indicators, and Forecasting,*edited by J. H. Stock and M. W. Watson, Chicago: University of Chicago Press

[9] Sims, C.A. and T. A. Zha 1998, Bayesian Methods for Dynamic Multivariate Models, *International Economic Review,*39, no.4:949-68

[10] Stock, J. 2006, Forecasting and Now-Casting with Disparate Predictors: Dynamic Factor Models and Beyond, Manuscript, Department of Economics, Harvard University

[11] Theil, H. and A. S. Goldberger 1961, On Pure and Mixed Statistical Estimation in Economics, *International Economic Review,* , pp. 65-78

[12] Zellenr, A. 1971, *An Introduction to Bayesian Inference in Econometrics,* New York, John Wiley