Title	Forecasting with Vector Autoregressions by Bayesian Model Averaging
Author(s)	Sugita, Katsuhiro
Citation	琉球大学経済学ワーキングペーパーシリーズ = Ryukyu Economics Working Paper Series(REWP#03): 1-13
Issue Date	2019-06-25
URL	http://hdl.handle.net/20.500.12000/44627
Rights	

Ryukyu Economics Working Paper Series, University of the Ryukyus
琉球大学 経済学ワーキングペーパー

Forecasting with Vector Autoregressions by Bayesian Model Averaging

Katsuhiro Sugita

Working Paper No. REWP#03

June 2019

Forecasting with Vector Autoregressions by Bayesian Model Averaging

Katsuhiro Sugita

Faculty of Global and Regional Studies, University of the Ryukyus, Nishihara,

Senbaru, Okinawa, 903-0213, Japan

E-Mail: ksugita@ll.u-ryukyu.ac.jp

Abstract

This paper examines how vector autoregression model by Bayesian model averaging method can improve forecasting performance. Bayesian model averaging selects significant variables in vector autoregression model that contains many insignificant variables, and thus alleviates over-parameterization problem. For empirical application, macroeconomic data for three countries - US, UK and Japan - are examined. I find that the Bayesian model averaging method can improve forecasting performance.

1 Introduction

There have been many empirical research that employ Bayesian model averaging (henceforth, BMA) to conduct statistical analysis in the presence of model uncertainty, for example Fernández et al. (2001), Koop and Potter (2004), Strachan et al. (2007), Durlauf et al. (2008), Kapetanios et al. (2008), Wright (2008), Wright (2009), Magnus et al. (2010), Clark and Mccracken (2010), Faust et al. (2013), Hasan et al. (2018) and so on. BMA is a kind of variable selection and thus used for clarify which variables are important or not in the model. Therefore, concept of BMA is similar to the stochastic search variable selection porposed by George and McCulloch (1993) and George and Mcculloch (1997), though approach to inference is quite different. While most of these studies deal with single equation regression, I focus on forecasting with vector autoregression model (VAR) by BMA in this paper.

BMA is more useful for multivariate VAR models than for single univariate models since VAR models typically contain more variables and thus have over-parameterization problem, which leads to imprecise inference and thus deterioarates the forecast peformance. To deal with this dimensionality problem, this paper examines applying BMA to VAR models to see whether BMA improves forecast performances. I consider BMA that uses both the marginal likelihood with the method by Fernández et al. (2001) Eklund and Karlsson (2007). For empirical application, the performance of the BMA method is evaluated by the out-of-sample forecasting of three major macroeconomic variables of the US, Japan and UK., using either direct and iterated forecasting methods.

2 Vector Autoregression by Bayesian Model Averaging

I briefly survey BMA, see Raftery et al. (1997) or Hoeting et al. (1999) for complete review. Let consider a set of m models $M_{1,...,}M_m$, of which each r^{th} model has a parameter vector θ_r. Inference based on the single "best" model as for the selected model were true ignores model uncertainty. BMA does not choose the single "best" model, instead it considers all possible models as a weight to deal with the model uncertainty. Let w be some quantity of interest, then the predictive quantity can be obtained as:

$$p(w|y) = \sum_{r=1}^{m} p(w|y, M_r)\, p(M_r|y) \tag{1}$$

where $p(M_r|y)$ is the posterior model probability. Let $p(M_r)$ is the prior model probability, then the posterior model probability $p(M_r|y)$ is defined as

$$p(M_r|y) = \frac{p(y|M_r)\, p(M_r)}{\sum_{j=1}^{m} p(y|M_j)\, p(M_j)} \tag{2}$$

where $p(y|M_r)$, the marginal likelihood of the rth model M_r, and the likelihood integrated with the prior on the parameters of M_r, denoted by $p(\theta_r|M_r)$, so that

$$p(y|M_r) = \int p(y|\theta_r, M_r)\, p(\theta_r|M_r)\, \mathrm{d}\theta_r \tag{3}$$

In many cases the number of models m under consideration is too big to compute the posterior model probability in (2) for every possible model. For example, in the case of three-variable VAR model with 10 lag length, where there are 93 parameters in the model, there are 2^{93} possible models. To deal with this problem, BMA uses a Markov chain Monte Carlo Model Composition (MC^3), which draws model itself from the model space.

Let y_t be a $n \times 1$ vector of $t = 1, ..., T$, then VAR model with p lag is given as:

$$y_t = \mu + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \varepsilon_t \tag{4}$$

where μ is a $n \times 1$ vector of an intercept term, Φ_i, $i = 1, \ldots, p$, are $n \times n$ matrices of lag coefficients and the errors ε_t are assumed to be $N(0, \Omega)$, where Ω is an $n \times n$ covariance matrix. Let $n \times (1 + np)$ matrix $\Theta = (\mu, \Phi_1, ..., \Phi_p)$ and $1 \times (1 + np)$ matrix $x_t = (1, y_t', ..., y_{t-p+1}')$, then equation (4) can be rewritten as a system of seemingly unrelated regressions (SUR):

$$y_t = \Theta x_{t-1}' + \varepsilon_t \tag{5}$$

$$= (I_n \otimes x_{t-1})\phi + \varepsilon_t \tag{6}$$

where $\phi = \mathrm{vec}(\Theta')$ is an $n(1 + np) \times 1$ matrix. To apply BMA to VAR model, the SUR model in equation (6) should be rewritten in a matrix form as:

$$y = z\phi + \varepsilon \tag{7}$$

where $y = (y_{p+1}', y_{p+2}', ..., y_T')'$ is a $Tn \times 1$ vector, $z = ((I_n \otimes x_p)', (I_n \otimes x_{p+1})', ..., (I_n \otimes x_{T-1})')'$ is a $Tn \times n(1 + np)$ matrix, and $\varepsilon \sim N(0_{Tn \times Tn}, \Sigma)$. Each column of z in equation (7) corresponds to variable of each equation.

As for prior specification, Moral-Benito (2015) review various priors for BMA. In this paper, I follow objective prior for BMA, developed by Fernández et al. (2001), which is a natural conjugate framework and has been used by many researchers for BMA including Koop and Potter (2004), Ley and Steel (2009), Wright (2008, 2009). The natural conjugate Normal-Gamma prior with g-prior implies

$$\phi_r|\Omega \sim N\left(0, \left[g_r z_r' z_r\right]^{-1} \Omega^{-1}\right) \tag{8}$$

where g_r is the scalar prior hyperparameter. $g_r = 1/k^2$ if $Tn \leqq k^2$ and $1/(Tn)$ otherwise where k is the number of regressors in ϕ, though Eklund and Karlsson (2007) choose $g_r = 1/k^3$, which yields a slightly more informative prior in their application. A prior for the variance-covariance matrix Ω is given as a noninformative prior as

$$p(\Omega) \propto \Omega^{-1} \tag{9}$$

The prior model probability $p(M_r)$ is given by

$$p(M_r) \propto \rho^{k_r}(1-\rho)^{k-k_r} \tag{10}$$

where ρ is prior expected model size divided by k, the number of regressors in ϕ. Note that ρ=0.5 leads to equal prior model probability.

With above priors, the posterior for ϕ_r is derived as a multivariate t distribution with mean

$$E(\phi_r|y,M_r) = \left[(1+g_r)z_r'z_r\right]^{-1} z_r'y \tag{11}$$

covariance matrix

$$var(\phi_r|y,M_r) = \frac{\bar{\nu}\bar{s}^2}{\bar{\nu}-2}\left[(1+g_r)z_r'z_r\right]^{-1} \tag{12}$$

where $\bar{\nu} = N$ degrees of freedom, and

$$\bar{s}^2 = \bar{\nu}^{-1}\left[\frac{1}{1+g_r}y'P_{z_r}y + \frac{g_r}{1+g_r}(y-\bar{y}\iota_T)'(y-\bar{y}\iota_T)\right] \tag{13}$$

where $P_{X_r} = I_{Tm} - z_r(z_r'z_r)^{-1}z_r'$.

The marginal likelihood for model M_r is derived as

$$p(y|M_r) \propto \left(\frac{g_r}{1+g_r}\right)^{\frac{k_r}{2}}\left[\frac{1}{1+g_r}y'P_{z_r}y + \frac{g_r}{1+g_r}(y-\bar{y}\iota_T)'(y-\bar{y}\iota_T)\right]^{-\frac{Tn-1}{2}}. \tag{14}$$

The Markov chain Monte Carlo model composition (MC^3) uses the random walk chain Metropolis-Hastings algorithm with the marginal likelihood (14) to draw model from model space.

Let $n(1+np)\times 1$ vector $M_r^{(k)}$ be kth cycle model indicator vector where its element is 0 if the element ϕ_j is restricted to be zero, otherwise 1 if there is no restriction. Let $z_t^{(k)}$ is kth cycle explanatory matrix in (7) which contains column of z if corresponding variable ϕ_j is contained, otherwise deleted the column of z. Given initial values $M_r^{(0)}$, $\phi_r^{(0)}$, $z_t^{(0)}\Omega^{(0)}$, the kth cycle is obtained from $M_r^{(k-1)}$, $\phi_r^{(k-1)}$, $z_t^{(k-1)}$, $\Omega^{(k-1)}$ by sequentially simulating the steps:

1. draw random integer l^* from discrete uniform distribution $l = \mathrm{U}(0, n(1+np))$, where $n(1+np)$ is the maximum number of the parameters in the model.

2. check whether l^*th row of $M_r^{(k-1)}$ is 1 or 0. If its value is 1, then delete l^*th row of ϕ_j, set M_r^* and make new z^* in (7) in which l^*th column of z is deleted, otherwise add l^*th row of $M_r^{(k-1)}$ and set M_r^* and z^* in which l^*th column of z is added.

3. compute new marginal likelihood for model M_r^*, $p(y|M_r^*)$ in (3).

4. by the random walk chain Metropolis-Hastings algorithm, compute the acceptance probability $\alpha\left(M_r^{(k-1)}, M_r^*\right) = \min\left[\frac{p(y|M_r^*)}{p\left(y|M_r^{(k-1)}\right)}, 1\right]$.

5. set $M_r^{(k)} = M_r^*$ with probability $\alpha\left(M_r^{(k-1)}, M_r^*\right)$ and set $M_r^{(k)} = M_r^{(k-1)}$ with probability $1-\alpha\left(M_r^{(k-1)}, M_r^*\right)$.

6. compute all parameters of model $M_r^{(k)}$, then set $k = k+1$.

7. repeat step 1 - 6.

Eklund and Karlsson (2007) point out that the use of the marginal likelihood entails the problem of in-sample overfitting of the data, and propose that the use of the predictive likelihood rather than the marginal likelihood for BMA procedure produces better forecast performance.

3 Multi-period Forecasts for VAR Models

For multi-step ahead forecasting, there are two method - direct forecasting method and iterated forecasting method. The iterated forecast method is based on one-step ahead forecast, that is to used for further step ahead forecast. The VAR model in (4) can be written as

$$\mathbf{y}_\tau = \mu + \Phi \mathbf{y}_{\tau-1} + \varepsilon_\tau \tag{15}$$

where $\mathbf{y}_\tau = (y'_\tau, ..., y'_{\tau-p+1})'$, $\mu = (\mu', 0, ..., 0)'$, $\varepsilon_\tau = (\varepsilon'_\tau, 0, .., 0)'$ and

$$\Phi = \begin{bmatrix} \Theta_1, ..., \Theta_{p-1} & \Theta_p \\ I_{n(p-1)} & 0_{n(p-1)\times n(p-1)} \end{bmatrix}$$

The one-step ahead forecast $\hat{y}'_{\tau+1|\tau}$ of the VAR model is obtained by estimating the parameters in (4) as $\hat{y}_{\tau+1|\tau} = \hat{\mu} + \sum_{i=1}^{p} \hat{\Phi}_i y_{\tau+1-i|\tau}$. Iterated forecasts for the h-period forecasts are obtained recursively as

$$\hat{y}_{\tau+h} = \hat{\mu} + \sum_{i=1}^{p} \hat{\Phi}_i \hat{y}_{\tau+h-i} \tag{16}$$

where $\hat{y}_{i|\tau} = y_j$ for $j \leq \tau$, thus the iterated forecasts can be computed as

$$\hat{\mathbf{y}}_{\tau+h} = \sum_{i=0}^{h-1} \hat{\Phi}^i_{(I)} \hat{\mu}_{(I)} + \hat{\Phi}^h_{(I)} \mathbf{y}_{\tau-1}. \tag{17}$$

Direct forecasting method predicts h-ahead forecast $y_{\tau+h}$ estimating the regression

$$y_\tau = \Theta x'_{\tau-h-1} + \varepsilon_\tau \tag{18}$$

then using the estimated coefficient $\hat{\Theta}$ directly to forecast $y_{\tau+h}$ with data through period τ as

$$\hat{\mathbf{y}}_{\tau+h} = \hat{\mu}_{(D)} + \hat{\Theta}_{(D)} \mathbf{y}_{\tau-1} \tag{19}$$

or

$$\hat{y}_{\tau+h} = \hat{\Theta} x'_{\tau-1} + \varepsilon_\tau \tag{20}$$

Thus, the relative forecast accuracy depends on how accurate $\hat{\Theta}_{(I)}$ and $\hat{\Theta}_{(D)}$ are estimated. If $\hat{\Theta}_{(I)}$ is badly estimated with large errors, then its powered values diverge increasingly from $\Theta_{(I)}$. Since the iterated method depends on one-period ahead coefficients $\hat{\Theta}_{(I)}$, the direct method is preferable when the one-period ahead model is not specified correctly. Chevillon and Hendry (2005) evaluate the asymptotic and finite-sample properties of direct forecasting method, and show that, compared with iterated method, the direct method is more efficient asymptotically, more precise in finite samples and more robust against model misspecification. The theoretical advantages of the direct forecasting method over the iterated method are shown by Bhansali (1996, 1997), Clements and Hendry (1996), and Kang (2003) among others. However, Marcellino et al. (2006) evaluates a large-scale empirical comparison of iterated and direct forecasts using U.S. macroeconomic time series data, and find that iterated forecasts tend to have smaller MSFEs than direct forecasts, contrary to the theoretical preference of direct forecasts.

To evaluate the forecasting performances among several different models in a recursive out-of-sample prediction exercise, the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE) are widely used. Let $y'_{\tau+h}$ is a vector of observations at time $\tau+h$ for $\tau=\tau_0,...,T-h-1$, and $h=2$, 4, 8 and 12- step ahead forecasts. Then, $\hat{\Phi}=\left(\hat{\mu}',\hat{\Theta}'_1,\dots,\hat{\Theta}'_p\right)'$ is estimated for both the direct and iterated method, using information up to $\tau-1$ to forecast values $\hat{y}_{\tau+h}$ starting from $\tau=\tau_0$ up to $\tau=T-h-1$, and calculate the MSFE defined as:

$$\text{MSFE}=\frac{1}{T-h-\tau_0+1}\sum_{\tau=\tau_0}^{T-h}\left[y_{\tau+h}-\hat{y}_{\tau+h}|\hat{\Phi},Y_{\tau-1}\right]^2. \tag{21}$$

and

$$\text{MAFE}=\frac{1}{T-h-\tau_0+1}\sum_{\tau=\tau_0}^{T-h}|y_{\tau+h}-\hat{y}_{\tau+h}|\hat{\Phi},Y_{\tau-1}| \tag{22}$$

where $Y_{\tau-1}=(X_{\tau-1},X_{\tau-2},\dots,X_1)$. For evaluating the forecast performance, I also use the bias squared as Jochmann et al. (2010), which is defined as:

$$\text{Bias Sq.}=\left[\frac{1}{T-h-\tau_0+1}\sum_{\tau=\tau_0}^{T-h}\left\{y_{\tau+h}-\hat{y}_{\tau+h}|\hat{\Phi},Y_{\tau-1}\right\}\right]^2 \tag{23}$$

4 Macroeconomic Forecasting Using VAR with BMA

In this section, multi-step forecasting performances of BMA are evaluated in a recursive pseudo out-of-sample using three macroeconomic time series of three countries - US, UK, and Japan. All data are quarterly and obtained from the Federal Reserve Bank of St. Louis[1]. The three variables are unemployment rate, inflation rate and interest rate. Inflation rates are obtained by the 400 times the difference of the log of CPI. Interest rates are 3-month treasury bill rates for US, government securities for Japan and UK. US data are from 1953:Q1 to 2018:Q4 with 264 observations. UK data are from 1960:Q2 to 2016:Q4 with 227 observations. Japan data are from 1970:Q1 to 2017:Q3 with 191 observations.

To estimate models by BMA, the prior parameters ρ in 10 is set to be 0.5 so that equal prior model probability is allocated to each model. MC^3 is run at each date τ with 400,000 draws after 100,000 burn-in. With regard to lag-length determination, both the AIC and BIC were considered. The AIC and BIC were recomputed at each date τ, so the order of the selected forecasting model changes from one period to the next.

Table 1 - 6 present the results for all forecasting exercises. Table 1 provides the summary of the forecast performance with direct method for US with lag length selected by AIC and BIC. Table 2 provides the results by iterated forecasts for US using AIC and BIC for lag selection.

1. BMA produces lower MSFE, MAFE and Bias squared than MLE for most cases.

2. Compared the results by AIC (long lag length) with the results by BIC (short lag length), the forecast performances by BMA tends to be insensitive to the choice of the lag length, while the MLE estimator is considerably affected by the selection of the lag length.

[1] https://fred.stlouisfed.org

3. The iterated forecasts tend to outperform direct forecast.

These three facts can be found for the UK (Table 3 for direct method Table 4), and for Japan, (Table 5 and Table 6 for iterated method).

5 Conclusion

In this paper, BMA method is applied to vector autoregressive model to select only significant variables in the model that contains many insignificant variables, so that the method can alleviate the over-parametrization problem and improves the forecasting performance. For empirical studies, VAR models with macroeconomic variables - unemployment rate, inflation rate and interest rate - for US, UK and Japan are examined. I find that the BMA can improve forecasting performance compared with the MLE method with any forecast horizon, by either direct or iterated forecasting method.

References

Bhansali, R. (1997). "Direct autoregressive predictors for multistep prediction: Order selection and performance relative to the plug in predictors." *Statistica Sinica*, 7: 425–449.

Bhansali, R. J. (1996). "Asymptotically efficient autoregressive model selection for multistep prediction." *Annals of the Institute of Statistical Mathematics*, 48(3): 577–602.

Chevillon, G. and Hendry, D. F. (2005). "Non-parametric direct multi-step estimation for forecasting economic processes." *International Journal of Forecasting*, 21(2): 201–218.

Clark, T. E. and Mccracken, M. W. (2010). "Averaging forecasts from VARs with uncertain instabilities." *Journal of Applied Econometrics*, 25: 5–29.

Clements, M. P. and Hendry, D. F. (1996). "Multi-Step Estimation for Forecasting." *Oxford Bulletin of Economics and Statistics*, 58(4): 657–684.

Durlauf, S. N., Kourtellos, A., and Tan, C. M. (2008). "Are Any Growth Theories Robust?" *The Economic Journal*, 118(527): 329–346.

Eklund, J. and Karlsson, S. (2007). "Forecast combination and model averaging using predictive measures." *Econometric Reviews*, 26(2-4): 329–363.

Faust, J., Wright, J. H., and Gilchrist, S. (2013). "Credit Spreads as Predictors of Real-time Economic Activity." *The Review of Economics and Statistics*, 95(5): 1501–1519.

Fernández, C., Ley, E., and Steel, M. F. (2001). "Model uncertainty in cross-country growth regressions." *Journal of Applied Econometrics*, 16(5): 563–576.

George, E. I. and McCulloch, R. E. (1993). "Variable selection via Gibbs sampling." *Journal of the American Statistical Association*, 88(423): 881–889.

George, E. I. and Mcculloch, R. E. (1997). "Approaches for bayesian variable selection." *Statistica sinica*, 7: 339–373.

Hasan, I., Horvath, R., and Mares, J. (2018). "What type of finance matters for growth? Bayesian model averaging evidence." *World Bank Economic Review*, 32(2): 383–409.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). “Bayesian model averaging: A tutorial.” *Statistical science*, 14(4): 382–401.

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.

Kang, I. B. (2003). “Multi-period forecasting using different models for different horizons: An application to U.S. economic time series data.” *International Journal of Forecasting*, 19(3): 387–400.

Kapetanios, G., Labhard, V., and Price, S. (2008). “Forecasting using Bayesian and information-theoretic model averaging: An application to U.K. inflation.” *Journal of Business and Economic Statistics*, 26(1): 33–41.

Koop, G. and Potter, S. (2004). “Forecasting in dynamic factor models using Bayesian model averaging.” *The Econometrics Journal*, 7(2): 550–565.

Ley, E. and Steel, M. F. J. (2009). “On the effect of prior assumptions in Bayesian model averaging with applications to growth regression.” *Journal of Applied Econometrics*, 24(4): 651–674.

Magnus, J. R., Powell, O., and Prüfer, P. (2010). “A comparison of two model averaging techniques with an application to growth empirics.” *Journal of Econometrics*, 154(2): 139–153.

Marcellino, M., Stock, J. H., and Watson, M. W. (2006). “A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series.” *Journal of Econometrics*, 135(1): 499–526.

Moral-Benito, E. (2015). “Model averaging in economics: An overview.” *Journal of Economic Surveys*, 29(1): 46–75.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). “Bayesian Model Averaging for Linear Regression Models.” *Journal of the American Statistical Association*, 92(437): 179–191.

Strachan, R., van Dijk, H., and Dijk, H. K. V. (2007). “Bayesian Model Averaging in Vector Autoregressive Processes with an Investigation of Stability of the US Great Ratios and Risk of a Liquidity Trap in the.” Technical report, Econometric Institute Research Papers.

Wright, J. H. (2008). “Bayesian Model Averaging and exchange rate forecasts.” *Journal of Econometrics*, 146(2): 329–341.

— (2009). “Forecasting US inflation by bayesian model averaging.” *Journal of Forecasting*, 28(2): 131–144.

Table 1: Forecast Performance with Direct Method using US data

		forecast horizon									
		1		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC											
		Av.Lag=10.78		Av.Lag = 13.02		Av.Lag = 15.67		Av.Lag = 18.55		AvLag = 19.43	
Unemp.	MSFE	0.036	0.015	0.168	0.070	0.681	0.425	1.460	1.050	8.274	1.091
	MAFE	0.136	0.093	0.269	0.192	0.570	0.458	0.878	0.780	1.750	0.745
	Bias Sq.	0.001	0.000	0.013	0.000	0.092	0.010	0.350	0.127	0.612	0.045
Inflation	MSFE	0.605	0.477	0.949	0.603	2.425	0.655	3.025	1.026	12.23	2.673
	MAFE	0.491	0.432	0.644	0.512	0.961	0.568	1.244	0.762	1.751	0.983
	Bias Sq.	0.007	0.002	0.010	0.007	0.057	0.066	0.360	0.312	2.698	0.794
Interest	MSFE	0.038	0.017	0.131	0.058	0.655	0.198	2.513	0.766	2.939	1.357
	MAFE	0.137	0.091	0.262	0.181	0.581	0.357	1.214	0.688	1.432	0.945
	Bias Sq.	0.002	0.001	0.013	0.002	0.084	0.014	0.495	0.311	1.876	0.766
Lag by BIC											
		Av.Lag = 3.000		Av.Lag = 2.000		Av.Lag = 3.893		Av.Lag = 1.009		AvLag = 1.000	
Unemp.	MSFE	0.023	0.015	0.092	0.078	0.417	0.388	1.005	0.965	1.158	1.081
	MAFE	0.106	0.094	0.205	0.202	0.435	0.434	0.677	0.680	0.761	0.748
	Bias Sq.	0.000	0.000	0.003	0.000	0.024	0.003	0.079	0.040	0.082	0.043
Inflation	MSFE	0.499	0.471	0.639	0.571	0.659	0.583	0.786	0.620	0.866	0.636
	MAFE	0.436	0.424	0.508	0.482	0.575	0.511	0.674	0.592	0.684	0.589
	Bias Sq.	0.005	0.002	0.011	0.005	0.080	0.026	0.197	0.126	0.316	0.208
Interest	MSFE	0.024	0.017	0.077	0.058	0.245	0.191	0.673	0.641	1.094	1.153
	MAFE	0.119	0.091	0.216	0.181	0.385	0.346	0.639	0.625	0.807	0.815
	Bias Sq.	0.002	0.000	0.009	0.002	0.050	0.010	0.222	0.200	0.527	0.533

Table 2: Forecast Performance with Iterated Method using US data

		forecast horizon							
		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC (Av.Lag = 13.7)									
Unemp.	MSFE	0.179	0.058	0.670	0.228	1.289	0.655	4.006	0.927
	MAFE	0.284	0.171	0.582	0.332	0.829	0.625	1.168	0.764
	Bias Sq.	0.021	0.000	0.103	0.000	0.228	0.001	0.258	0.001
Inflation	MSFE	1.899	0.665	4.430	0.633	5.968	0.595	2.416	0.506
	MAFE	0.832	0.521	1.088	0.541	1.311	0.564	1.080	0.490
	Bias Sq.	0.016	0.004	0.028	0.010	0.227	0.032	0.953	0.063
Interest	MSFE	0.307	0.060	1.103	0.178	2.631	0.396	1.420	0.516
	MAFE	0.374	0.183	0.606	0.324	0.996	0.509	0.965	0.590
	Bias Sq.	0.014	0.003	0.050	0.011	0.231	0.048	0.736	0.103
Lag by BIC (Av.Lag = 3.00)									
Unemp.	MSFE	0.106	0.059	0.401	0.236	0.913	0.666	1.044	0.902
	MAFE	0.214	0.174	0.440	0.344	0.649	0.514	0.705	0.724
	Bias Sq.	0.005	0.000	0.024	0.001	0.065	0.003	0.054	0.003
Inflation	MSFE	0.679	0.637	0.630	0.579	0.689	0.568	0.698	0.496
	MAFE	0.520	0.506	0.546	0.504	0.618	0.534	0.592	0.471
	Bias Sq.	0.016	0.003	0.043	0.007	0.129	0.021	0.223	0.044
Interest	MSFE	0.098	0.060	0.231	0.175	0.538	0.380	0.741	0.495
	MAFE	0.243	0.183	0.389	0.319	0.592	0.499	0.681	0.578
	Bias Sq.	0.010	0.002	0.038	0.009	0.174	0.043	0.346	0.092

Table 3: Forecast Performance with Direct Method: UK

		forecast horizon									
		1		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC											
		Av.Lag=20.71		Av.Lag = 20.10		Av.Lag = 20.68		Av.Lag = 24.13		AvLag = 26.00	
Unemp.	MSFE	0.004	0.008	0.024	0.022	0.140	0.069	0.725	0.242	6.152	0.726
	MAFE	0.048	0.062	0.111	0.097	0.280	0.178	0.669	0.333	1.655	0.665
	Bias Sq.	0.000	0.000	0003	0.000	0.027	0.004	0.038	0.060	0.195	0.187
Inflation	MSFE	0.401	0.156	0.641	0.133	1.268	0.160	10.84	0.562	9.022	1.178
	MAFE	0.450	0.291	0.563	0.271	0.688	0.313	1.831	0.587	1.543	0.794
	Bias Sq.	0.012	0.001	0.011	0.003	0.002	0.033	1.861	0.110	0.009	0.030
Interest	MSFE	0.059	0.018	0.249	0.059	1.179	0.310	4.480	0.676	6.297	0.780
	MAFE	0.162	0.082	0.314	0.176	0.664	0.407	1.089	0.642	1.562	0.719
	Bias Sq.	0.002	0.003	0.026	0.013	0.275	0.110	0.518	0.290	0.285	0.197
Lag by BIC											
		Av.Lag = 2.094		Av.Lag = 3.000		Av.Lag = 3.385		Av.Lag = 1.080		AvLag = 1.000	
Unemp.	MSFE	0.002	0.007	0.011	0.014	0.062	0.058	0.256	0.193	0.326	0.272
	MAFE	0.033	0.059	0.078	0.083	0.186	0.164	0.403	0.344	0.455	0.428
	Bias Sq.	0.000	0.000	0.002	0.000	0.021	0.006	0.098	0.038	0.101	0.038
Inflation	MSFE	0.269	0.275	0.122	0.129	0.180	0.127	0.374	0.287	0.566	0.447
	MAFE	0.410	0.424	0.276	0.273	0.348	0.294	0.557	0.472	0.699	0.608
	Bias Sq.	0.000	0.001	0.002	0.005	0.070	0.053	0.298	0.197	0.484	0.356
Interest	MSFE	0.017	0.018	0.059	0.056	0.222	0.165	0.681	0.614	1.082	1.017
	MAFE	0.086	0.083	0.172	0.162	0.387	0.307	0.714	0.662	0.959	0.935
	Bias Sq.	0.004	0.003	0.023	0.013	0.132	0.066	0.509	0.437	0.919	0.874

Table 4: Forecast Performance with Iterated Method: UK

		forecast horizon							
		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC (Av.Lag = 21.1)									
Unemp.	MSFE	0.039	0.033	0.222	0.118	1.021	0.321	3.277	0.506
	MAFE	0.136	0.129	0.311	0.254	0.639	0.434	1.061	0.568
	Bias Sq.	0.001	0.000	0.014	0.002	0.054	0.016	0.020	0.055
Inflation	MSFE	1.276	0.153	1.503	0.173	2.980	0.315	2.723	0.373
	MAFE	0.707	0.302	0.731	0.325	0.980	0.460	1.048	0.490
	Bias Sq.	0.003	0.000	0.017	0.005	0.394	0.040	0.826	0.090
Interest	MSFE	0.447	0.060	1.257	0.154	2.923	0.332	2.426	0.499
	MAFE	0.394	0.161	0.673	0.286	1.062	0.421	1.165	0.530
	Bias Sq.	0.029	0.011	0.286	0.041	0.421	0.132	0.530	0.246
Lag by BIC (Av.Lag = 2.84)									
Unemp.	MSFE	0.014	0.031	0.063	0.115	0.187	0.311	0.311	0.484
	MAFE	0.031	0.125	0.115	0.250	0.311	0.428	0.484	0.555
	Bias Sq.	0.001	0.000	0.012	0.001	0.055	0.013	0.051	0.049
Inflation	MSFE	0.522	0.257	0.566	0.356	1.109	0.390	0.901	0.342
	MAFE	0.501	0.436	0.539	0.496	0.706	0.602	0.796	0.699
	Bias Sq.	0.001	0.001	0.004	0.006	0.221	0.026	0.558	0.041
Interest	MSFE	0.083	0.060	0.256	0.155	0.829	0.335	1.379	0.507
	MAFE	0.198	0.162	0.376	0.287	0.745	0.425	1.002	0.537
	Bias Sq.	0.022	0.011	0.118	0.043	0.554	0.136	1.004	0.255

Table 5: Forecast Performance with Direct Method: Japan

		forecast horizon									
		1		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC											
		Av.Lag=26.00		Av.Lag = 24.76		Av.Lag = 25.85		Av.Lag = 26.00		AvLag = 26.00	
Unemp.	MSFE	0.023	0.016	0.076	0.045	0.393	0.148	2.313	0.716	4.417	1.449
	MAFE	0130	0.093	0.219	0.154	0.537	0.286	1.232	0.696	1.768	0.978
	Bias Sq.	0.000	0.000	0.000	0.002	0.000	0.000	0.002	0.007	0.196	0.004
Inflation	MSFE	0.341	0.139	0.366	0.159	0.638	0.163	2.295	0.226	9.053	0.395
	MAFE	0.449	0.274	0.457	0.292	0.656	0.300	0.991	0.377	1.557	0.480
	Bias Sq.	0.002	0.001	0.000	0.001	0.008	0.003	0.014	0.021	1.183	0.064
Interest	MSFE	0.025	0.001	0.104	0.005	0.898	0.048	4.682	0.668	4.958	1.960
	MAFE	0.108	0.026	0.223	0.057	0.582	0.173	1.414	0.588	1.725	1.033
	Bias Sq.	0.003	0.000	0.025	0.002	0.250	0.012	1.525	0.000	2.423	0.001
Lag by BIC											
		Av.Lag = 2.000		Av.Lag = 3.000		Av.Lag = 4.000		Av.Lag = 10.38		AvLag = 13.08	
Unemp.	MSFE	0.015	0.016	0.042	0.044	0.132	0.119	0.579	0.291	2.482	0.868
	MAFE	0.092	0.093	0.152	0.151	0.297	0.265	0.659	0.450	1.239	0.756
	Bias Sq.	0.001	0.000	0.004	0.001	0.020	0.001	0.004	0.010	0.385	0.000
Inflation	MSFE	0.259	0.223	0.267	0.247	0.503	0.404	1.945	0.279	8.839	0.409
	MAFE	0.384	0.350	0.413	0.385	0.563	0.505	0.813	0.439	1.410	0.492
	Bias Sq.	0.003	0.003	0.004	0.006	0.069	0.047	0.049	0.019	1.105	0.052
Interest	MSFE	0.003	0.001	0.010	0.004	0.045	0.015	3.950	0.530	3.893	1.789
	MAFE	0.039	0.025	0.077	0.051	0.173	0.098	1.088	0.474	1.350	0.955
	Bias Sq.	0.000	0.000	0.001	0.001	0.002	0.004	0.911	0.003	1.649	0.001

Table 6: Forecast Performance with Iterated Method: Japan

		forecast horizon							
		2		4		8		12	
Variable		MLE	BMA	MLE	BMA	MLE	BMA	MLE	BMA
Lag by AIC (Av.Lag = 29.0)									
Unemp.	MSFE	0.109	0.051	0.435	0.143	1.949	0.396	4.855	0.613
	MAFE	0.260	0.167	0.534	0.306	1.119	0.567	1.713	0.721
	Bias Sq.	0.005	0.001	0.017	0.006	0.029	0.029	0.042	0.097
Inflation	MSFE	0.766	0.154	0.695	0.158	1.157	0.182	1.287	0.206
	MAFE	0.684	0.301	0.676	0.311	0.847	0.323	0.904	0.346
	Bias Sq.	0.076	0001	0.045	0.001	0.028	0.001	0.017	0.002
Interest	MSFE	0.156	0.004	0.619	0.010	3.587	0.029	9.660	0.052
	MAFE	0.289	0.049	0.575	0.083	1.255	0.154	1.853	0.207
	Bias Sq.	0.011	0.001	0.070	0.004	0.435	0.018	1.322	0.039
Lag by BIC (Av.Lag = 3.35)									
Unemp.	MSFE	0.056	0.050	0.161	0.136	0.469	0.346	0.897	0.489
	MAFE	0.178	0.163	0.333	0.290	0.600	0524	0.863	0.649
	Bias Sq.	0.005	0.000	0.024	0.002	0.097	0.011	0.224	0.045
Inflation	MSFE	0.430	0.255	0.397	0.244	0.806	0.309	0.879	0.346
	MAFE	0.483	0.373	0.501	0.387	0.709	0.450	0.703	0.464
	Bias Sq.	0.014	0.009	0.069	0.023	0.101	0.038	0.171	0.045
Interest	MSFE	0.025	0.004	0.142	0.008	1.438	0.023	5.349	0.040
	MAFE	0.097	0.045	0.189	0.074	0.476	0.136	0.843	0.182
	Bias Sq.	0.001	0.001	0.005	0.003	0.039	0.012	0.182	0.025