Transactions of the Nigerian Association of Mathematical Physics
***Vol. 6* (Jan., 2018), pp. 126– 139**
© ***Trans. of NAMP***

In honour of Prof. Ekhaguere at 70

Forecasting Nigeria inflation: application of Bayesian model averaging

O. E. Olubusoye[a*], O. B. Akanbi[b], and O. J. Akintande[c]

[a,b,c]*Department of Statistics, University of Ibadan, Nigeria*

Abstract. The need to understand and forecast inflationary movement will continue to be of interest to the monetary policy authorities. Substantial progress has been made in the development of both theoretical and econometric models but forecasting inflation still remains a herculean task. Recent empirical works on inflation have settled for the use of Bayesian Model Averaging (BMA) approach which makes forecast from variety of existing models by averaging them over rather than using a single model. BMA is increasingly becoming very popular but with little or no application to Nigerian inflation. The focus of this paper is to demonstrate how BMA can be used to determine the good predictors of inflation in Nigeria and make forecasts. Based on the posterior inclusion probability (PIP) of 0.95567 (or 96%), the real interest rate top the list of predictors of inflation and negatively signed. The top 10 models account for 27.6% of the total posterior probability. The posterior model probability is widely scattered across models and no single model dominates. Expectedly, credible intervals predictive values appear more reliable than point forecasts in BMA.

Keywords: Posterior Inclusion Probability (PIP), Posterior Model Probability (PMP), Bayesian Model Selection (BMS) package.

1. Introduction

Inflation will for a long time remain a subject of interest and concern not only to monetary authorities but also econometricians, statisticians and economists. The reasons are quite many including: the macroeconomic effect on savings and investments; the uncertainty effect on fiscal budgeting; and the impact on international competitiveness and trade performance. According to Adewumi and Awosika (1982), inflation is one of the greatest problems plaguing the world economic scene. Thus, households, governments and investors are usually keenly interested in the movement and behaviour of inflation.

The need to ensure an effective conduct of monetary policy by the central bank has made inflation forecasting very essential. The current economic conundrum arising from the fall in the international oil price has threatened the maintenance of price stability, a key function of the central bank. Forecasting inflation generally provides the critical inputs to the conduct of monetary policy. The inflation forecast is not only an important decision-making tool, but also an important communication device.

As a result of the traditional argument of lags in the monetary policy transmission mechanism, inflation forecast plays a crucial role in the conduct of monetary policy. There are long lags between monetary policy actions and their impact on the economy. Since interest rate, the instrument used by monetary policymaker has its strongest impact on inflation several months ahead, then policy should be directed towards targeting the forecast in an appropriate horizon. Policies responding only to the current state of the economy may not prove to be stabilising policies. Therefore, it is generally recognized that central bank policies must be far-sighted.

There is no gainsaying the fact that a lot of progress and development has been made in terms of both theoretical and econometric models in the recent times, forecasting inflation still remains a difficult task requiring more attention and efforts. With the rapid changing in the national economy

*Corresponding author. Email: busoye2001@yahoo.com

often caused by different shocks, the task of forecasting inflation has become even more difficult. The ambiguous and changing structure of the Nigerian economy further complicates this task. However, producing a real-time macroeconomic forecasting particularly in a developing economy like Nigeria is such a complex and challenging problem. In advanced economies, forecasting process often employs a variety of formal models, both structural and purely statistical. The forecasts of inflation are usually developed through an eclectic process that combines model-based projections, anecdotal and other "extra model" information and professional judgement.

The most common tool for inflation forecasting is probably the Phillips curve which uses a single measure of economic slack such as unemployment to predict future inflation. The Phillips curve equation relates the unemployment rate or some other measure of aggregate economic activity to a measure of inflation rate. Some recent specifications of Phillips curve equations relate the current rate of unemployment to future changes in inflation rate. The main idea behind this specification is that there is a baseline rate of unemployment at which inflation tends to remain constant. The thought is that inflation tends to rise over time when unemployment is below this baseline rate, and inflation falls when unemployment is above this baseline rate. The term non-accelerating inflation rate of unemployment (NAIRU) is used to describe this baseline unemployment rate. Hence, modern specifications based on it are referred to as NAIRU Phillips curves.

The NAIRU Phillips curves have become so popular in academic literature on inflation forecasting and among policy making institutions because of the view that inflation forecasts from these models are more accurate than forecasts from other models. Indeed, Blinder (1997) argues that "the empirical Phillips curve has worked amazingly well for a decade" and then concludes on the basis of this empirical success that a Phillips curve should have "a prominent place in the core model" used for macroeconomic policy making purposes. The literature on studies based on the extension of Phillips curve is rather too voluminous but a few representative and prominent ones include Ang et al (2007), Groen et al (2009), Stock and Watson (1999) and Stock and Watson (2008).

However, the usefulness of Phillips curve equation for predicting inflation has been challenged and questioned by several authors. For instance, Atkeson and Ohanian (2001) obtained a short-run version of the Phillips curve by regressing the four-quarter change in the inflation rate on the unemployment rate and a constant. The study was able to show that the short-run Phillips curve does not represent a stable empirical relationship that can be exploited for the purpose of constructing a reliable inflation forecasts. The regression coefficient on the unemployment rate (which measures the slope of the short-run Phillips curve) varies significantly across different sample periods.

Koop and Korobilis (2010) clearly identify four issues that often arise when forecasting inflation based on Phillips curve. First, the coefficient on the predictors can change over time. Second, the model relevant for forecasting can potentially change over time. For example, a set of predictors for inflation in 1990s may have been different from now. Some variables might predict well in expansions but not in recessions. In fact Stock and Watson (2008) find that Phillips curve forecasts are better in some periods but the simpler univariate forecasting models work better at other periods. Third, the number of potential predictors can be large. The existence of several predictors can lead to a large number of models. This definitely raises a serious statistical problem for model selection strategy.

Substantial progress has been made by researchers in the aspect of using large number of predictors in forecasting inflation. Information in this huge number of variables is combined in a sensible manner to prevent the estimation of a large number of unrestricted parameters. Stock and Watson (2001, 2002) submits that the best predictive performance is obtained by averaging forecasts constructed from a large number of models. This popular approach is referred to in literature as the Bayesian Model Averaging (BMA) and the idea was initiated by Leamer (1978). BMA efficiently and systematically evaluates a wide range of predictor variables for inflation and (almost) all possible models that these predictors in combination can give rise to. Using the posterior probabilities of the models, weights are assigned to the different models to obtain a weighted model averaging.

The BMA methodology is increasingly witnessing wide spread applications across an array of disciplines including political science, social research, econometrics, etc. Notable contributions in the area of inflation are indeed numerous. For instance, Jacobson and Karlsson (2002) consider BMA for the purpose of forecasting Swedish consumer price index using a large set of potential indicators, comprising some 80 quarterly time series covering a wide spectrum of Swedish economic activity. The result in terms of out-of-sample-performance suggests that BMA is a useful alternative to other forecasting procedures, in particular recognising the flexibility by which new information can be incorporated.

Similarly, Wright (2003) employs BMA for the prediction of U.S inflation using quarterly series from 1960q1 – 2003q2 for a total number of 93 predictor variables considered as alternative measures of economic slack and several asset prices. The study confirms the superiority in performance and the consistency of BMA forecasts across subsamples of four different measures of inflation CPI inflation, CPI core inflation, GDP deflator inflation and PCE deflator inflation.

Lastly, a more recent study is Gonzalez (2010) which applies BMA to forecast inflation in Colombia. An application of BMA is implemented to construct combined forecasts for the Colombian inflation for the short and medium run. The dataset used for the empirical application consists of 73 monthly macroeconomic Colombian time series from 1999:11 to 2009:12. The series are grouped into three categories: Real Activity (26 series), Prices (23 series), Credit Money and Exchange Rate (24 series). The study finds BMA as a "useful and consistent way to select variables and models with high predictive power.

Given the above background and the dearth of similar study implementing this approach in the case of Nigeria, this study is designed to fill this vacuum. The burning research questions are: what are the factors (monetary and fiscal) determining the path of inflation in Nigeria? How does single model perform relative to averaging of several potential models? Therefore, the overarching objective of this paper is to analyze a wide spectrum of inflation predictors and all possible models that can arise from combining the models using Bayesian Model Averaging. Specifically, the study objectives to be pursued are: (1) Determine the good predictors of inflation in Nigeria; (2) Analyze forecasts from all the models combining these predictors; and, (3) Examine the forecast performance of inflation in Nigeria when simple equal weighted model averaging is used and when Bayesian Model Averaging is also used.

The rest of the paper is structured as follows. In the section following this introduction, we discuss the theory underpinning BMA borrowing largely from Koop (2005). In section 3 we discuss BMA computation and the three packages available in R software for implementing the analysis. Section 4 is devoted to the application of BMA to the analysis of the Nigerian inflation. Finally, Section 5 gives the summary and conclusion.

2. The Bayesian model averaging approach

In situations where many potential explanatory variables exist, alternative models can be defined based on the set of explanatory variables they include. In general, if there are k potential explanatory variables in a study, then 2^k models are possible. It is obvious that as k gets larger, the number of possible models is astronomical. Two major problems are usually confronted. First is how to handle this large number of models. The second relates to prior information about the models. These problems have largely been surmounted in Bayesian literature as would be examined in this section.

Consider a linear regression model with a constant term, β_0, and k potential explanatory variables $x_1, x_2, , x_k$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + u \tag{1}$$

From model (1), 2k different models are feasible based on inclusion and exclusion of each regressor. If we let M_j for $j = 1, 2, 3, , 2^k$ denote the different models under consideration. Having constructed the model space, the posterior distribution of any coefficient, say β_r, given the data D is

$$Pr(\beta_r|D) = \sum_{j:\beta_r \in M_j} Pr(\beta_r|M_j) Pr(M_j|D) \tag{2}$$

The logic of Bayesian inference requires one to obtain result for every model under consideration and average them. The weights in the averaging are the posterior model probabilities, $Pr(M_jD)$. These weights are the key feature for estimation via BMA. The posterior model probability M_j is the ratio of its marginal likelihood to the sum of marginal likelihoods over the entire model space and is given by

$$Pr(M_j|D) = \frac{Pr(D|M_j)Pr(M_j)}{Pr(D)} = \frac{Pr(D|M_j)Pr(M_j)}{\sum_{i=1}^{2^k} Pr(D|M_i)Pr(M_i)} \tag{3}$$

where the marginal likelihood of the jth model is

$$Pr(D|M_j) = \int Pr(D|\beta^j, M_j)Pr(\beta^j|M_j)d\beta^j \tag{4}$$

and β^j is the vector of parameters from model M_j, $Pr(\beta^j M_j)$ is a prior probability distribution assigned to the parameters of model M_j, $Pr(M_j)$ is the probability that M_j is the true model and $Pr(D\beta^j, M_j)$ is the likelihood.

The estimated posterior means and standard deviations of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k)$ are then constructed as

$$E[\hat{\beta}|D] = \sum_{j=1}^{2^k} \hat{\beta} Pr(M_j|D) \tag{5}$$

$$V[\hat{\beta}|D] = \sum_{j=1}^{2^k} (Var[\beta|D, M_j] + \hat{\beta^2})Pr(M_j|D) - E[\hat{\beta}|D]^2 \tag{6}$$

Each Model implies a forecast density $f_1, ..., f_p$, where $p = 2^k$. Similarly, each model produces a point forecast. In reality the true model is unknown thus leading to model uncertainty. In such situation, the point forecast density becomes

$$f^* = \sum_{j=1}^{p} Pr(M_j|D)f_j \tag{7}$$

This implies that the point forecast in (7) weights each of the 2^k forecasts by the posterior density of the model. Thus, the weights in the averaging are the posterior model probabilities. In summary, the BMA approach simply involves the following three steps: (1) specifying the models, (2) specifying the model priors and (3) and specifying parameter priors. The only thing remaining is just computation once the three steps are accomplished.

2.1 *The models*

Even though the models could be of any type but for this paper we assume they are linear regression models. Hence, the jth model is

$$\text{Model } M_j \qquad y_j = X_j\beta_j + U_j \qquad j = 1, 2, ..., p = 2^k \tag{8}$$

Since all the p linear regression models differ only in their explanatory variables and it is standard to assume that all the models have an intercept, equation can be written as

$$\text{Model } M_j \qquad y = \alpha l_N + X_j\beta_j + U \qquad j = 1, 2, ..., p = 2^k \tag{9}$$

where $l_n = N \times 1$ vector of ones, X_j is $N \times k_j$ matrix containing some (or all) the predictor variables. Equation (9) will differ from model to model only in the predictor variables included or excluded. The vector of disturbance terms, U, is assumed to be a multivariate Normal distribution with mean 0_N and covariance $\sigma^2 I_N$, $N(0_N, h^{-1}I_N)$ where $h = \sigma^{-2}$, the error precision. Using the definition of the multivariate Normal density, the likelihood function for the jth model can be written as

$$P(y|\beta_j, h) = \frac{1}{(2\pi)^{N/2}}\{h^{1/2}exp[(\beta_j - \hat{\beta}_j)^{'}X_j^{'}X_j(\beta_j - \hat{\beta}_j)]\}h^{v/2}exp[-\frac{hv}{2s^{-2}}] \tag{10}$$

where $v = N - k$, $\hat{\beta}_j = (X_j^{'}X_j)^{-1}X_j^{'}y$ the ordinary least square estimator and $s^2 = \frac{(\beta_j - \hat{\beta}_j)^{'}(\beta_j - \hat{\beta}_j)}{v}$. The natural conjugate prior for (10) is Normal - Gamma density. Therefore, if we elicit a prior for β_j conditional on h of the form

$$\beta_j|h \sim N(\beta_{*j}, h^{-1}V_{*j}) \tag{11}$$

and a prior for h of the form

$$h \sim G(s_*^{-2}, v_*) \tag{12}$$

Then the posterior will also have Normal-Gamma density of the form

$$\beta_j, h|y \sim NG(\beta_j^*, V_j^*, s^{*-2}, v^*) \tag{13}$$

Where

$$V_j^* = (V_{*j}^{-1} + X_j^{'}X_j)^{-1} \tag{14}$$

$$\beta_j^* = V_j^*(V_{*j}^{-1}\beta_{*j} + X_j^{'}X_j\hat{\beta}_j) \tag{15}$$

$$v^* = v_* + N \tag{16}$$

and s^{*-1} is defined implicitly through

$$v^*s_j^{*2} = v_*s_*^2 + vs^2 + (\hat{\beta}_j - \beta_{*j})^{'}[V_{*j} + (X_j^{'}X_j)^{-1}]^{-1}(\hat{\beta}_j - \beta_{*j}) \tag{17}$$

By integrating out from (13) it gives the marginal posterior distribution for h which is a multivariate t distribution given as

$$\beta_j|y \sim t(\beta_j^*, s_j^{*2}V_j^*, v^*) \tag{18}$$

It follows from the definition of the t distribution that

$$E(\beta_j|y, M_j) = \beta_j^* = V_j^* X_j^{`} y \tag{19}$$

and

$$var(\beta_j|y, M_j) = \frac{v^* s_j^{*2}}{v^* - 2} V_j^* \tag{20}$$

At this stage, the values of the prior hyperparameters β_{*j}, V_{*j}, s_*^{-2}, and v_* must be chosen to reflect the researcher's prior information or belief. The rule of thumb suggested by Koop (2000) is that "when comparing models using posterior odds ratios, it is acceptable to use noninformative priors over parameters which are common to all models. However, informative, proper priors should be used over all other parameters". Thus since error precision, h, and the intercept, α, are common to all models the standard noninformative prior can be used for them. Hence, the prior for h is

$$P(h) \propto \frac{1}{h} \tag{21}$$

and the prior for the intercept is

$$P(\alpha) \propto 1 \tag{22}$$

Fernandez et al (2001b) suggested the need to standardize (subtract off the means) all the predictor variables in the model to be confidently sure that the noninformative prior for the intercept has similar implications for all the models. Now we want to determine the value for β_j. The usual practice is to center priors over the hypothesis that predictor variables are not related or have no effect on the dependent variable. Since there are many variables involved in BMA, we might suspect that some of them are not important. In that case, we set

$$\beta_{*j} = 0_{kj}$$

Finally, we left with choosing the value for V_j^*. This involves setting

$$V_j^* = (g_j X_j^{`} X_j)^{-1} \tag{23}$$

The g_j is a scalar quantity called the g-prior and it was first introduced in Zellner (1986). The g-prior is simply saying that the prior covariance of β_j is proportional to the comparable data-based quantity. In other words, it has similar properties as the information provided by the explanatory variables. The use of g-prior simplifies the problem of eliciting prior covariance matrices by reducing it to choosing just a single hyperparameter. If $g_j = 0$, this implies a perfectly noninformative prior. If $g_j = \frac{1}{2}$, it implies that prior and data information are weighted equally in the posterior covariance matrix. Fernandez et al (2001b) carried out an extensive experimentation which led to the recommendation that g_j can be choosing such that

$$g_j = \begin{cases} \frac{1}{N} & if N > k^2 \\ \frac{1}{k^2} & if N \leq k^2 \end{cases} \tag{24}$$

Finally, each model in the model space is assigned equal probability. That is, $P(M_j) = \frac{1}{2^k}$.

3. BMA computation and packages

Ideally the results of (19), (20) and (23) are needed to be able to implement BMA but the large number of models involved makes it an almost impossible task. Fortunately, literature on BMA has focused on how these quantities can be approximated or computed analytically. Various algorithms have been developed which can compute the posterior probabilities. The most popular of these algorithms is the Markov Chain Monte Carlo Model Composition (MC^3) initially developed by Madigan and York (1995). The most common MC^3 algorithm for sampling from the model space is a Metropolis-Hastings algorithm. For the constraint of space, these algorithms cannot be discussed here. However, there are now packages in R software that can implement the BMA. In the rest of this section we present an overview of three currently available packages that can implement BMA in R software. It is probably the only software that provides a suite of routines to conduct a BMA analysis. Amini and Parmeter (2011) provides similarities of the features of the three packages and a detailed comparison of the options for model sampling and search, construction of model priors, call outputs and diagnostics.

3.1 *The BMS package*

BMS is an acronym for Bayesian Model Selection. The package excels in offering a range of widely used prior structures coupled with efficient MCMC algorithms to sort through the model space. It allows for uniform and binomial-beta priors on the model space as well as informative prior inclusion probabilities. Via these customized model priors one can thus fuse prior beliefs into the otherwise purely agnostic analysis that is prevalent in the applied literature using BMA. The BMS package also provides various specifications for Zellner's g prior including the so-called hyper- g priors advocated in Liang et al. (2008); Ley and Steel (2010); Feldkircher and Zeugner (2009). The sensitivity of BMA results to the specification of Zellner's g prior is well documented in the literature (Feldkircher and Zeugner 2011). The package comes along with numerous graphical tools to analyze posterior coefficient densities, the posterior model size or predictive densities. It also includes a graphical representation of the model space via an image plot.

Since enumerating all potential variable combinations becomes infeasible quickly for a large number of covariates, the BMS package uses Markov Chain Monte Carlo (MCMC) samplers to gather results on the most important part of the posterior distribution when more than 14 covariates exist (Amini and Parmeter 2011).

3.2 *The BAS Package*

BAS is an acronym for Bayesian Adaptive Sampling. The package performs Bayesian Model Averaging in linear models using stochastic or deterministic sampling without replacement from posterior distributions (Clyde 2010). Prior distributions on coefficients are of the form of Zellner's g-prior or mixtures of g-priors. Options include the Zellner-Siow Cauchy Priors, the Liang et al hyper- g priors, Local and Global Empirical Bayes estimates of g, and other default model selection criteria such as AIC and BIC. Sampling probabilities may be updated based on the sampled models. The family of prior distribution on the models is nearly identical to the BMS package allowing uniform, Bernoulli or Beta-Binomial distributions as priors.

If the number of covariates is less than 25, the BAS package enumerates all models; otherwise it implements three different search algorithms to find the models with the highest posterior probability. The first is Bayesian Adaptive Sampling algorithm which samples models from the model space without replacement using the initial sampling probabilities, and will update the sampling probabilities using the estimated marginal inclusion probabilities. The second algorithm is Adaptive MCMC sampler. The last algorithm runs an initial MCMC to calculate marginal inclusion probabilities and then samples without replacement (Amini and Parmeter 2011).

3.3 *The BMA package*

The BMA package, an acronym for Bayesian Model Averaging, performs BMA analysis assuming a uniform distribution on model priors and using a simple BIC (Bayesian Information criterion) approximation to construct the prior probabilities on the regressions coefficients (Raftery et al (2010)). The package is inflexible regarding model priors, assuming all models are on equal footing a *priori*. The model prior is set to $\frac{1}{2^k}$ where k is the total number of variables included in the analysis. Unlike BMS and BAS packages, BMA does not employ Zellner's g-priors as its choice of the prior distributions for the regression coefficients. Instead, it employs the BIC approximation, which corresponds fairly closely to the uniform information priors (UIP).

4. Application of BMA to Nigerian inflation

The data used are yearly which spans through 1970-2011 (42 observations) and were sourced from the 2012 CBN Statistical Bulletin and 2012 World Development Indicators (WDI). The study variable is inflation rate while the predictors are real interest rate (RIR), openness (open) measured as the ratio of total trade to GDP per capita, total public wage (Topbwg), food price for consumer price index (Fdcpi), government expenditure (Gexp), official real exchange rate (oer), term of trade (Tot), government capita formation (Gcf), investment (human capita) (Hc), fiscal deficit (Fisdefc), average rainfall (Rain), GDP per capita (Gdppc), petroleum local price (Pet) and money supply (M2). The MC^3 sampler uses 100000 draws after the burn ins of 20000 with uniform distribution as the prior model and the modified g-prior, $g = BRIC$ for parameters.

In performing Bayesian Model Averaging (BMA) analysis, the first thing to check for after the models sampling is the simulation convergence. The argument, Corr PMP in Table 1 below explains the convergence for the simulation between the Analytical (exact) and the MCMC counts PMPs. And with the 0.9964, the simulation converges almost accurately.

Table 1: Checking for the PMP Convergence

Mean no. regressors ("3.3463")	Draws ("1e+05")	Model space 2K ("32768")	% Top models ("98")	Model Prior "uniform / 7.5"
No. models visited "31966"	Burn ins "20000"	% visited "98"	No. Obs "42"	g-Prior "BRIC"
Shrinkage-Stats "Av=0.9956"	Time "26.85938 secs"	Corr PMP "0.9964"		

Source: Authors

Figure 1b is the plot for the convergence of the model simulation. From the density plot, the correlation value (0.9964) for the model simulation convergence is also confirmed below. Figure 1a shows the average number of the regressors (model size = 3.3463) for the 2000 best models.

Table 2 shows the importance of using the top models for inference instead of the full model (model space). With the best 2000 models, the MC^3 sampler (PMP MCMC) estimate (0.98276) was accurate when compared with the posterior probability for the true (exact) model. This is because the sampler visited about 98% (31966) of the model space (32768) for its simulation. This is realized when all PMP values for the 31966 visited models and the exact models (32768) were summed separately and then compared.

Table 2: Cumulative Model Probabilities for the Exact and the MCMC

PMP (Exact)	PMP (MCMC)
0.98276	0.98276

Source: Authors

Table 3 relates the importance of each regressor with its posterior mean, standard deviation and conditional posterior sign over the plausible models. This relevance of the regressors is measured in BMA by the Posterior Inclusion Probability (PIP). The higher the probability, the more important is the regressor, especially when the PIP value is above 50%. Out of all the regressors, it was only the real interest rate (RIR) that has best PIP (95%) and therefore is the most important variable when

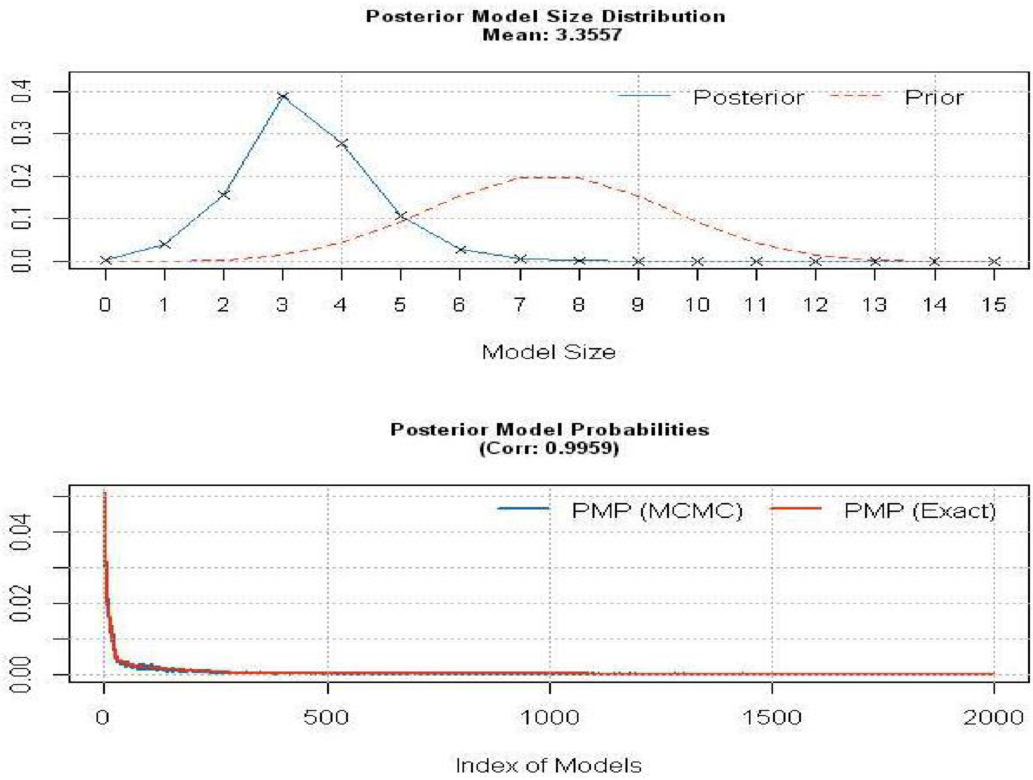

Figure 1a (above): Model size plot & Figure 1b (below): PMP convergence plot

modeling Nigerian Inflation. It is also noticed that the regressor has a negative average posterior mean (-0.5893) and a standard deviation (0.2213). This negative sign is clearly explained by the conditional posterior index in the 5th column of Table 3. In addition, all other regressors have PIPs less than 50% and therefore are not strong determinants of inflation in Nigeria. All these redundant variables have their posterior standard deviations greater than their posterior means, thus suggesting parameter uncertainty (under model uncertainty).

Table 3: Posterior Inclusion Probability for all the regressors using the MCMC simulation

Regressor	PIP	Post Mean	Post SD	Cond.Pos.Sign	Index
RIR	0.94883	-0.5893134	0.2212822	0.000000	15
Topbwg	0.38754	3.8926630	6.2763182	0.98934304	10
Open	0.37942	7.2760745	11.1009453	1.00000000	1
Fdcpi	0.20926	0.9373426	3.9343907	0.75929466	6
Gcf	0.18659	-0.8347824	2.6331984	0.03092341	12
Fisdefc	0.17756	-0.5376326	1.5075153	0.04353458	14
Gexp	0.16992	-0.7068562	2.8030096	0.09539783	11
Totpop	0.15134	-8.4078828	1450.1471	0.18607110	9
Rain	0.13220	-2.4449614	8.4558834	0.00000000	13
Rgdp	0.12777	2.2389171	1449.8909	0.21562182	8
Oer	0.11526	-0.1078729	1.7996881	0.35259413	3
Gdppc	0.11226	-3.7732818	1449.8238	0.33413504	7
Hc	0.09710	0.4316577	3.2055089	0.76138002	4
M2	0.07715	-0.6757781	15.7186686	0.31406351	2
Pet	0.07407	1.2859611	46.3338398	0.35830971	5

Source: Authors

Table 4 is similar to Table 3 except for the PIP values of the regressors which are higher in the

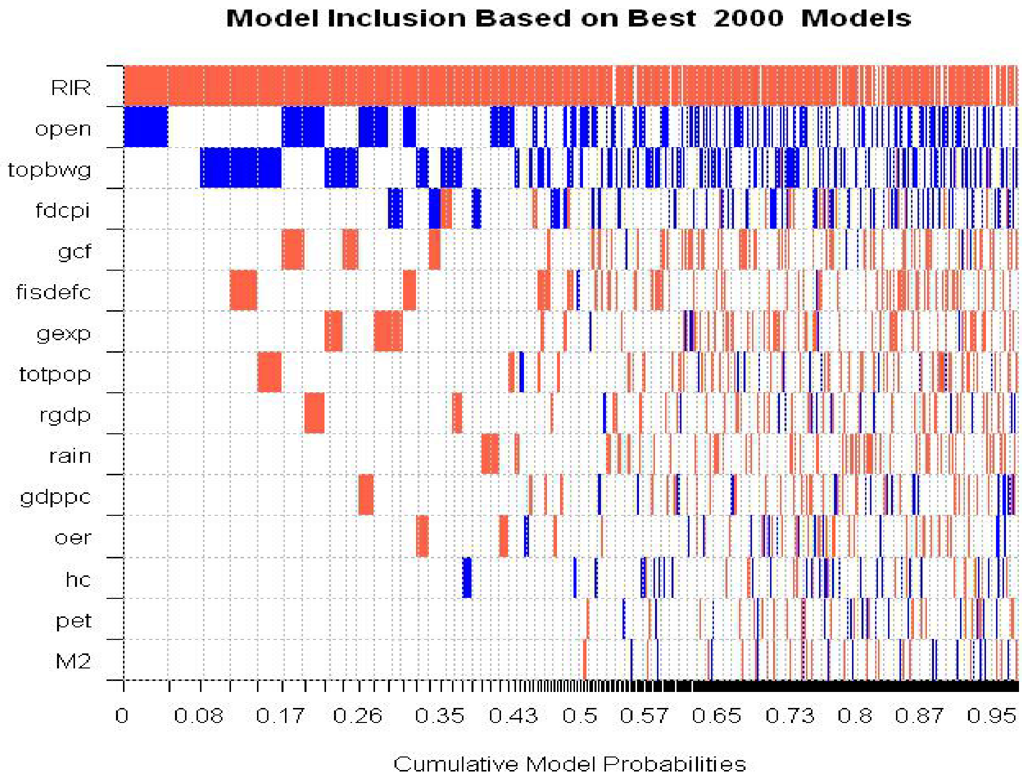

Figure 2: Model inclusion cumulative probability based on best 2000 models

exact simulation. Most Bayesians, like Fernandez et al (2001b) prefer exact simulation for their inference.

Table 4: Posterior Inclusion Probability for the first six regressors using the Exact simulation

Regressor	PIP	Post Mean	Post SD	Cond.Pos.Sign	Index
RIR	0.9556723	-0.5938350	0.2162564	0.00000000	15
Topbwg	0.3907474	7.5136548	11.1517022	1.00000000	10
Open	0.3729707	3.7319986	6.0918903	0.99278141	1
Fdcpi	0.1969269	0.8820093	3.7454258	0.77036304	6
Gcf	0.1764572	-0.7763577	2.4821112	0.03148911	12
Fisdefc	0.1723858	-0.5241394	1.4835624	0.04030597	14

Source: Authors

Figure 2 depicts the inclusion probability for the models. It explains the best 2000 models scaled by their posterior model probability (PMPs). This cumulates the probabilities for the 2000 best models out of the 31966 visited models. The figure also indicates signs of the regressor coefficients in the models. The RIR appears almost in all the 2000 best models (PIP = 96%) and has a negative sign with red while the blue represents a positive coefficient for a regressor in a model.

Table 5 shows the coefficients of the regressors in each of the first best 10 models in the model space along with their PMPs for both the analytical and MCMC counts. From the table, the best model (model1) for the Nigerian Inflation has a PMP value of 0.050 (by exact) and 0.047 (by MCMC) with coefficients, -0.6712 (RIR) and 12.944 (openness). It implies about 5% certainty that the true model is model 1.The probability value confirms that the posterior mass is not concentrated on just a few models. Similarly, Table 5 indicates that the true inflation model (model 1) is always favoured compared to any other model despite low posterior mass allocated to it.

Table 5: Posterior Inclusion Probability for all the regressors using the MCMC simulation

	Model									
Coefficient	1	2	3	4	5	6	7	8	9	10
RIR	-0.67	-0.62	-0.72	-0.60	-0.62	-0.55	-0.54	-0.65	-0.65	-0.61
Topbwg	-	-	4.21	11.12	13.37	-	-	9.85	8.18	-
Open	12.94	-	-	-	-	22.72	23.69	-	-	15.33
Fdcpi	-	-	-	-	-	-	-	-	-	-
Gcf	-	-	-	-	-	-3.18	-	-	-3.50	-
Fisdefc	-	-	-	-4.19	-	-	-	-	-	-
Gexp	-	-	-	-	-	-	-	-3.80	-	-
Totpop	-	-	-	-	-39.61	-	-	-	-	-
Rain	-	-	-	-	-	-	-	-	-	-
Rgdp	-	-	-	-	-	-	-16.1	-	-	-
Oer	-	-	-	-	-	-	-	-	-	-
Gdppc	-	-	-	-	-	-	-	-	-	-27.2
Hc	-	-	-	-	-	-	-	-	-	-
M2	-	-	-	-	-	-	-	-	-	-
Pet	-	-	-	-	-	-	-	-	-	-
PMP(Exact)	0.05	0.035	0.032	0.032	0.027	0.025	0.021	0.019	0.018	0.017
PMP(MCMC)	0.047	0.035	0.032	0.033	0.028	0.026	0.018	0.021	0.018	0.013

Source: Authors

Figure 3 depicts the marginal density of the real interest rate (RIR) along with its PIP value (95%). The middle vertical line indicates the posterior expected value (-0.60) over the model space. The posterior mean of this variable is with a negative sign which conforms to the theoretical expectation. In theory, an interest rate below zero should lower all market rates, thus also reducing borrowing costs for companies and households. In practice, though, theres a risk that the policy might do more harm than good.

Table 6 gives the predicted values for the year 2010 and the year 2011 representing the 41st and 42nd observations respectively. The values are conditional on what we know from other years (1970-2009) on the predictor variables. The predictions are 13.697701 and 1.728355 for 2010 and 2011 respectively. And when compared with their actual values (13.7202, 10.8408), the forecast for 2010 has a good fit (prediction), unlike the forecast for 2011. The forecast for 2011 suggests that it is an outlier or that our predictive model might not perform that well. Thus, other priors settings or data could then be examined in order to improve the out of sample predictive performance of the true model.

Table 6: Predictions for years 2010 and 2011

Year 2010 (41st observation)	Year 2011 (42nd observation)
13.697701	1.728355

Source: Authors

Table 7 shows the 95% credible interval predictive values for years 2010 and 2011. The results show that the predicted values for the years are within the limits.

Table 7: Predictions for years 2010 and 2011

	5%	95%
Year 2010 (41st observation)	-14.04985	41.38725
Year 2011 (42nd observation)	-25.43557	28.83423

Source: Authors

Figure 4 depicts only the expected value for the 2010 predictive density without comparing it with the actual value using 500 models. From the density, the expected predictive value of 13.697701 in table 6 is confirmed for this year. The expected predictive value is the red solid line and while the red break lines are the standard error of the distribution.

Figure 5 compares the predictive density for the year, 2010 between the expected and the actual

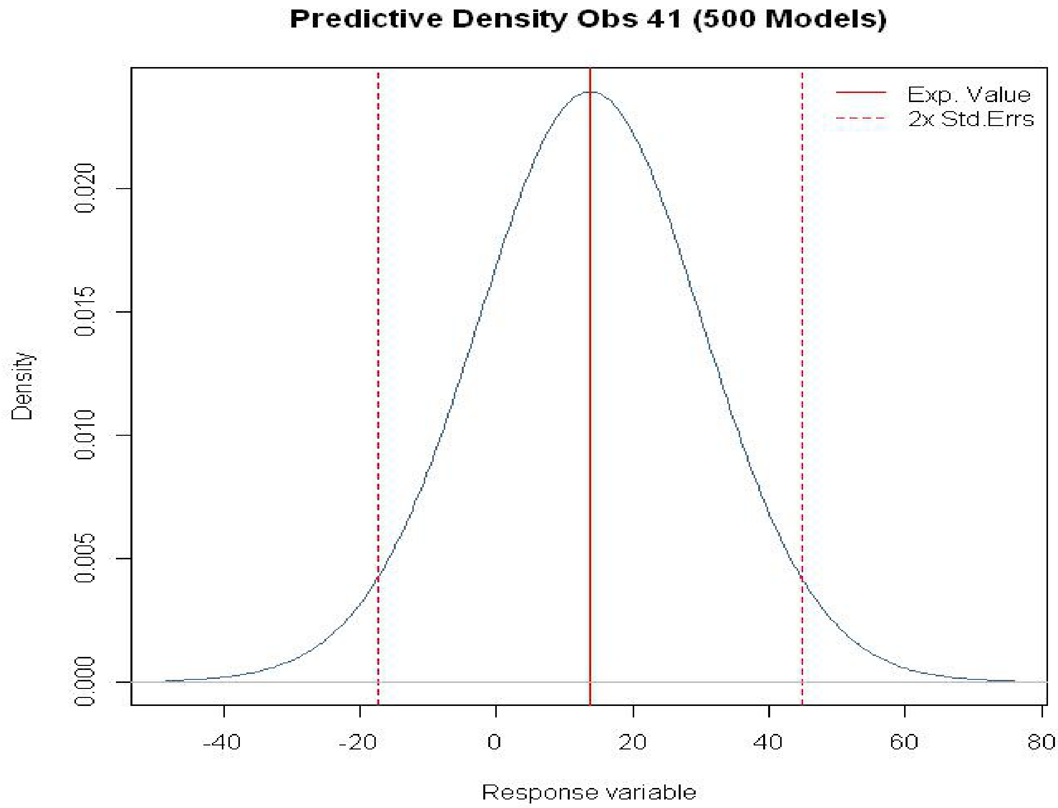

Figure 3: Marginal density of Real Interest Rate (RIR)

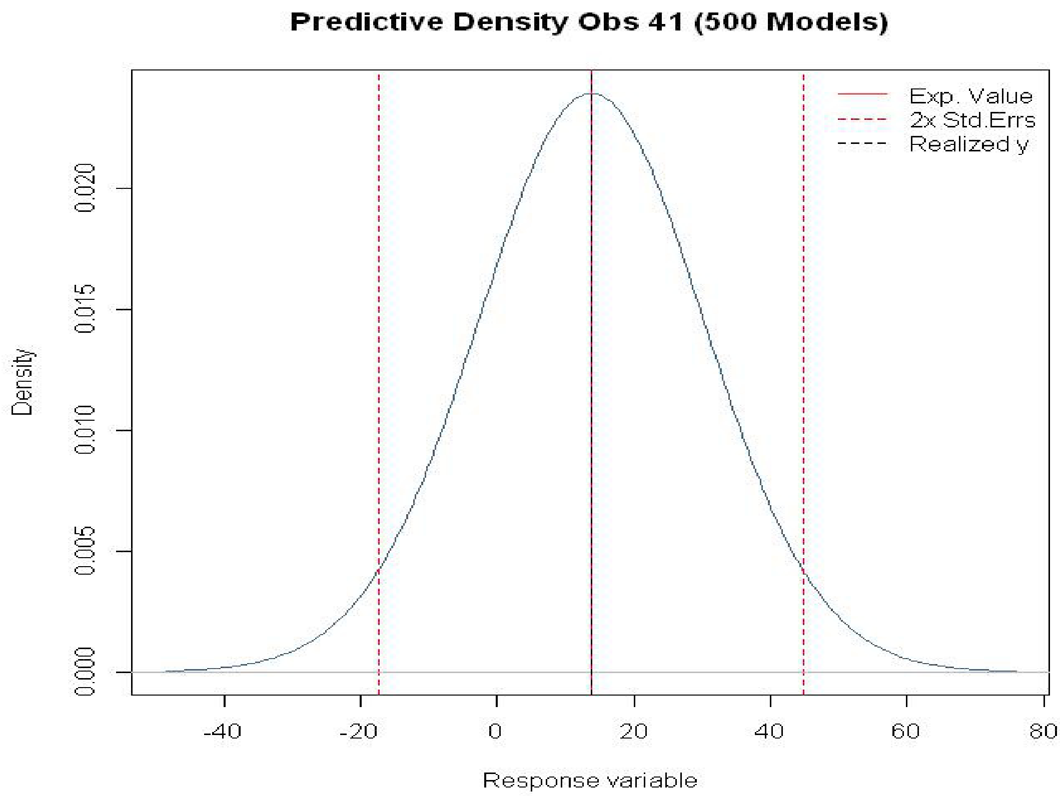

Figure 4: Predictive density for year 2010 over 500 models

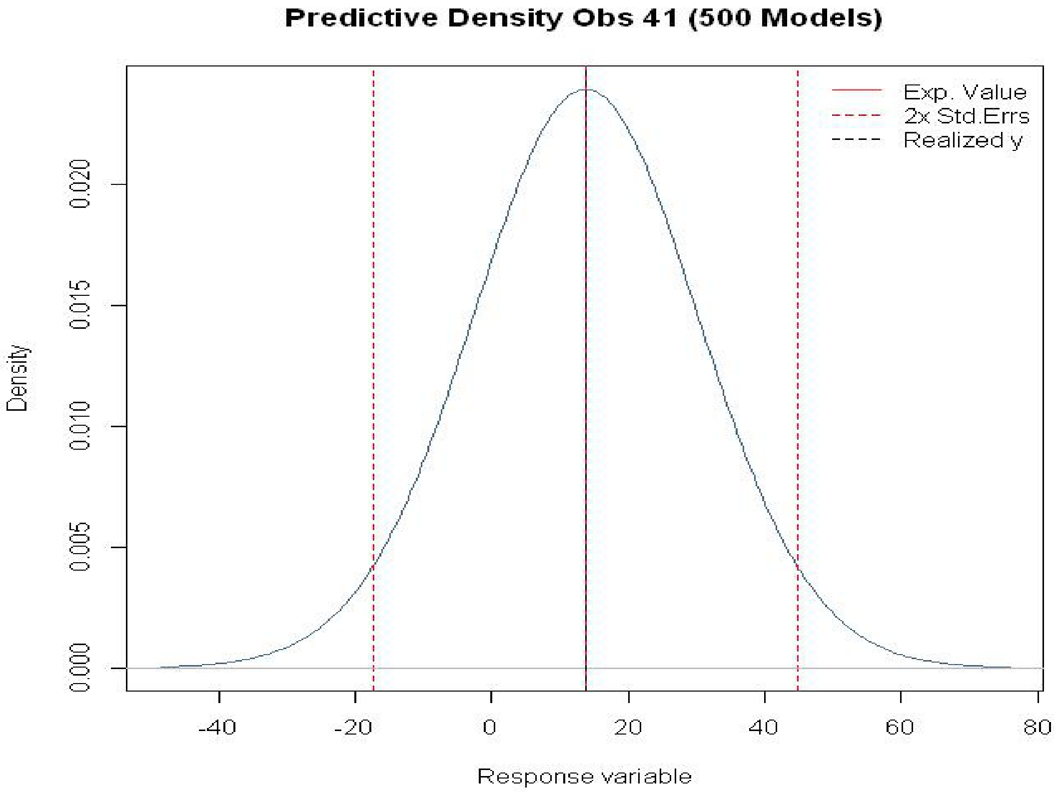

Figure 5: Plot comparison between the actual and the expected predicted values for 500 models.

values for 500 models. The density confirms likely the actual value of 13.7202 when compared with the expected value of 13.697701. This shows that the BMA prediction has good forecast. The dark break line is for the actual value, the expected predictive value is the red solid line and while the red break lines are the standard error of the distribution.

5. Conclusion

The crux of this paper is the practical demonstration of BMA technique to modelling inflation and for determining the predictors of inflation in Nigeria. It has been shown from the application to inflation that judicious pooling of information from a large number of predictors can be very helpful in forecasting inflation in the short term. Implementing BMA can be computationally demanding, since it is quite common for the number of models to be astronomical. We provide an overview of three currently available packages in R that can implement a BMA for empirical exercise. We recognize that this paper is a work in progress and there are still many areas to fine tune. One of such is the data. The quality and quantity of the predictor variables are far from being satisfactory considering the underlying requirements of BMA. The low frequency (yearly) nature of the series is another potential drawback. The least considered in most other studies is quarterly data with number of predictors as high as 70 and even more. The other critical issue is the assessment of the performance of the BMA inflation forecasts when compared with forecasts from other existing models. In sum, BMA could an addition to the suite of models available to forecast inflation by the monetary authorities in Nigeria.

References

[1] Adewumi M.O. and Awosika K. (1982): Conceptual and Statistical Problems in Measurement of Inflation. Proceedings of the National Conference on Inflation in Nigeria, NISER, Ibadan, pg 2 20.

[2] Amini SM, Parmeter CF (2011). Bayesian Model Averaging in R. Journal of Economic and Social Measurement, 36:4, pp. 253-287.

[3] Ang, A. Bekaert, G. and Wei, M., (2007). Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better? Journal of Monetary Economics 54, 1163-1212.

[4] Atkeson A. and Ohanian L.E (2001). Are Phillips Curves Useful for Forecasting Inflation? FRB Minneapolis Quarterly Review (Winter) pp. 2-11. https://www.minneapolisfed.org/research/qr/qr2511.pdf (last accessed June 2016).

[5] Blinder A. S. (1997). Is there a core of practical macroeconomics that we should all believe? American Economic Review 87 (May): 240

[6] Clyde, M.A. (2010), BAS: Bayesian Adaptive Sampling for Bayesian Model Averaging. R.

[7] Feldkircher, M. and Zeugner S. (2011). Bayesian Model Averaging with BMS," R-package.

[8] Feldkircher, M. and Zeugner, S. (2009): Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Efect in Bayesian Model Averaging. IMF Working Paper, WP/09/202.

[9] Fernandez, C., Ley, E. and Steel, M. (2001a): Model Uncertainty in Cross-country Growth Regressions. Journal of Applied Econometrics, 16, 563 -576.

[10] Fernandez, C., Ley, E. and Steel, M. (2001b): Benchmark priors for Bayesian model averaging, Journal of Econometrics, 100, 381 427.

[11] Gary Koop (2005). Bayesian Econometrics. John Wiley & Sons, Ltd.

[12] González, E. (2010). Bayesian Model Averaging. An Application to Forecast Inflation in Colombia (No. 007013). Banco de la Repblica. http://www.banrep.gov.co/sites/default/files/publicaciones/pdfs/borra604.pdf (last accessed June 2016).

[13] Groen J., Paap R and Ravazzolo F (2009). Real-time Inflation Forecasting in a Changing World. Econometric Institute Report, 2009-19, Erasmus University Rotterdam.

[14] Jacobson Tor and Karlsson Sune (2002): Finding Good Predictors for Inflation: A Bayasian Model Averaging Approach. Sveriges Riksbank Working Paper Series No 138.

[15] Jonathan H. Wright (2003): Forecasting U.S. Inflation by Bayesian Model Averaging. International Finance Discussion Papers, No 780.

[16] Koop, Gary and Korobilis, Dimitris (2010). Forecasting Inflation Using Dynamic Model Averaging. SIRE Discussion Papers 2010-113, Scottish Institute for Research in Economics http://repo.sire.ac.uk/bitstream/ handle/10943/242/$SIRE_DP_2010_113.pdf$?sequence=1&isAllowed=y (last accessed June 2016).

[17] Leamer, E.E. (1978): Specification Searches, Wiley, New York.

[18] Ley, E. and Steel, M. F. (2009): On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. Journal of Applied Econometrics, 24:4:651-674.

[19] Liang, F., Paulo, R., Molina, G., Clyde, M. A., and Berger, J. O. (2008): Mixtures of g Priors for Bayesian Variable Selection. Journal of the American Statistical Association, 03:410-423.

[20] Madigan, D. and York, J. (1995): Bayesian Graphical Models for Discrete Data. International Statistical Review, 63, 215–232.

[21] Madigan, D. and York, J. (1995): Bayesian Graphical Models for Discrete Data. International package vs. 0.92. URL: http://CRAN.R-project.org/package=BAS.

[22] Olubusoye, O.E., Akanbi O.B., Kazeem, B.A. (2015). Modelling the Nigerian inflation process using the Bayesian Model Averaging (BMA) approach. In: Seid, N. and Demirew, G. (Eds.). Proceedings of the 19th Annual Conference of the Africa Region Chapter of the Econometric Society, Ethiopia. Addis Ababa: III, pp. 117-140.

[23] Raftery, A., Hoeting, J., Volinsky, C., Painter, I. & Yeung, K. Y. (2010), BMA: Bayesian Model Averaging. R package version 3.13. URL: http://CRAN.R-project.org/package=BMA Statistical Review, 63, 215 232.

[24] Stock J. and Watson M. (2008). Phillips Curve Inflation Forecasts, NBER Working Paper No. 14322.

[25] Stock J. H. and Watson, M. W. (1999). Forecasting inflation. Journal of Monetary Economics 44 (October): 293-3.

[26] Stock J. H. and Watson, M. W. (2001). Forecasting Output and Inflation: The Role of Asset Prices, mimeo.

[27] Stock J. H. and Watson, M. W. (2008). Philips Curve Inflation Forecasts. NBER Working Paper No. 14322.vs. 0.3.0, URL: http://cran.r-project.org/web/packages/BMS/.