# Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 12 (7): 645-652, 2012
ISSN 1812-5654 / DOI: 10.3923/jas.2012.645.652


# Forecasting Key Macroeconomic Variables of the South African Economy using Bayesian Variable Selection

Mirriam Chitalu Chama-Chiliba, Rangan Gupta, Nonophile Nkambule and Naomi Tlotlego
Department of Economics, University of Pretoria, Pretoria, 0002, South Africa

**Abstract:** This study analyzed the forecasting performances of various multivariate models in predicting 1-8-quarters-ahead of the growth rate of GDP, the consumer price index inflation rate and the three months Treasury bill rate for South Africa over an out-of-sample period of 2000:Q1-2011:Q2, using an in-sample period of 1960:Q1-1999:Q4. The study compared the forecasting performances of the classical and the Minnesota-type Bayesian vector autoregressive (VAR) models with those of linear (fixed-parameter) and nonlinear (time-varying parameter) VARs involving a stochastic search algorithm for variable selection, estimated using Markov Chain Monte Carlo methods. In general, the study finds that variable selection, whether imposed on a time-varying VAR or a fixed parameter VAR, and non-linearity in VARs, play an important part in improving predictions when compared to the linear fixed coefficients classical VAR. However, the results does not indicate marked gains in forecasting power across the different Bayesian models, as well as, over the classical VAR model, possibly because the problem of over parameterization in the classical VAR is not that acute in our three-variable system. Hence, future research would aim to look at VAR models that include over 10 variables.

**Key words:** Forecasting, time varying parameters, variable selection, Bayesian vector autoregression

## INTRODUCTION

The vector autoregressive (VAR) model, though 'atheoretical' is particularly useful for forecasting purposes (Korobilis, 2011). This framework essentially involves a system, whereby equal number of lags of all the dependent variables enters as regressors in the equation of a specific dependent variable. One drawback of VAR models is that many parameters are needed to be estimated, some of which may be insignificant. This problem of over-parameterization, resulting in multicollinearity and loss of degrees of freedom leads to inefficient estimates and large out-of-sample forecasting errors (Gupta and Sichei, 2006; Gupta, 2006, 2007, 2009; Gupta and Das, 2008; Liu and Gupta, 2007; Liu *et al*., 2009, 2010; Balcilar *et al*., 2011). One solution, often adapted, is simply to exclude the insignificant lags based on statistical tests. Another approach is to use near VAR, which specifies unequal number of lags for the different equations (Gupta and Sichei, 2006; Gupta, 2006, 2007, 2009).

However, an alternative approach to overcome this over-parameterization, as described by Litterman (1981, 1986), Doan *et al*. (1984), Todd (1984) and Spencer (1993), is to use a Bayesian VAR (BVAR) model. Instead of eliminating longer lags, the Bayesian method imposes restrictions on these coefficients by assuming that these are more likely to be near zero than the coefficient on shorter lags. However, if there are strong effects from less important variables, the data can override this assumption. The restrictions are imposed by specifying normal prior distributions with zero means and small standard deviations for all coefficients with the standard deviation decreasing as the lags increases. Unless the variable is mean-reverting or stationary, the exception to this is, however, the coefficient on the first own lag of a variable, which has a mean of unity. Generally, following Litterman (1981), a diffuse prior is used for the constant. This is popularly referred to as the 'Minnesota prior' due to its development at the University of Minnesota and the Federal Reserve Bank at Minneapolis.

Not surprisingly, in the literature, the BVAR models have been found to produce the most accurate short- and long-term out-of-sample forecasts relative to both univariate and multivariate unrestricted classical VAR models[1]. In this regard, evidence for South Africa is no different, with a large number of recent studies showing superior forecasting power of BVAR models relative to not only classical VAR models but also, Dynamic Stochastic General Equilibrium (DSGE) models[2], in

**Corresponding Author:** Rangan Gupta, Department of Economics, University of Pretoria, Pretoria, 0002, South Africa

[1]Refer to Banbura *et al.* (2010) and Koop and Korobilis (2010) for further details. [2]An exception to this is the recent study by Gupta and Das (2010) who shows that when one develops a sophisticated DSGE model involving a variety of nominal and real rigidities, it is possible to outperform the BVAR model based on the Minnesota prior

predicting key macroeconomic variables (Gupta and Sichei, 2006; Gupta, 2006, 2007, 2009; Gupta and Das, 2008; Liu and Gupta, 2007; Liu *et al*., 2009, 2010; Alpanda *et al*. (2011) and Balcilar *et al*. (2011)[3].

Now-a-days though, besides the shrinkage approach of the Minnesota-type BVAR models, there are numerous other efficient methods to prevent the proliferation of parameters and eliminate parameter or model uncertainty. For example, variable selection priors (George *et al*., 2008), steady state priors (Villani, 2009), Bayesian model averaging (Anderson and Karlsson, 2008) and factor models (Stock and Watson, 2005), to name a few popular methods. Against this backdrop, following the study of Korobilis (2011), this study compares the forecasting performances of the classical and the Minnesota-type BVAR models with those of linear (fixed-parameter) and nonlinear (time-varying parameter (TVP). VARs involving a stochastic search algorithm for variable selection, estimated using Markov Chain Monte Carlo (MCMC) methods[4]. The term "stochastic search" simply means that if the model space is too large to assess in a deterministic manner, the algorithm will look for only the most probable models. Note that, the two main benefits of using this approach over the shrinkage methods are: First, variable selection is automatic, meaning that along with estimates of the parameters we get associated probabilities of inclusion of each parameter in the best model. This allows one to select among all possible VAR model combinations, without the need to estimate each and every one of these models. Second, this form of Bayesian variable selection is independent of the prior assumptions about the parameters. Note that the decision to use the stochastic search variable selection algorithm proposed by Korobilis (2011) over other available ones, such as those developed by George *et al*. (2008) and Korobilis (2008), is that one can apply the current algorithm to variable selection non-linear (time-varying) VAR models.

Specifically speaking, this study compared the forecasting performances of all these models in predicting one-to eight-quarters-ahead of the growth rate of GDP, the Consumer Price Index (CPI) inflation rate and the three months Treasury bill rate for South Africa over an out-of-sample period of 2000:Q1-2011:Q2, using an in-sample period of 1960:Q1-1999:Q4[5]. While the start and end-points of the sample is determined by data availability, the decision to use 2000: Q1 as the beginning of the out-of-sample period is determined by the fact that South Africa moved to an inflation targeting regime in the February of 2000. Besides, this choice is also consistent with most of the studies, mentioned above, that deals with forecasting in South Africa using BVAR models. The basic idea behind this exercise is to see if we could perform better than the BVAR models in forecasting key macroeconomic variables for South Africa by allowing for stochastic search for variable selection imposed on fixed and time-varying parameter models. In this regard, to the best of our knowledge, this is a first such attempt for South Africa.

## THE ECONOMETRIC METHODS

**Variable selection in VAR and the TVP-VAR model[6]:** A reduced form VAR can be written using following linear regression specification:

$$Y_{i+1} = Bx_i+\varepsilon_{i+1} \tag{1}$$

where, $Y_{i+1}$ is an (m×1) vector of dependent variables at time t = 1, ...,T; is a (k×1) vector, which may include lags of the dependent variables, intercept, dummies, trends and exogenous regressors, B is an (m×k) vector of VAR coefficients and $\varepsilon_t$~N (0, Σ), where Σ is a (m×m) covariance matrix.

Equation 1 can be re-written as a System of Unrelated Regressions (SUR) as follows, thus allowing for different equations in the VAR to have different explanatory variables:

$$Y_{i+1} = z_t\beta+\varepsilon_{i+1} \tag{2}$$

where, $Y_{i+1}$ and $\varepsilon_i$ are defined as above in Eq. 1; $z_t = I_m \otimes x'_t$ is a (m×n) matrix vector; while β = vec (B) is an (n×1) matrix. When there are no parameter restrictions, Eq. 2 can

[3]At the same time, there is also evidence for South Africa that, large scale BVAR models or factor models, which involve over two hundred variables, tend to outperform both classical and small-scale BVAR models, essentially involving three to six variables Gupta and Kabundi (2010, 2011a, b). In addition, allowing for non-linearity in the data generating process through logistic and exponential smooth transition autoregressive models, are also found to forecast better than small-scale VAR and BVAR models, as observed for South Africa by Balcilar *et al*. (2011). [4]For use of other linear, non-linear and nonparametric methods in forecasting various types of variables, by Aydin (2009), Assis *et al*. (2010), Samsudin *et al*. (2010), Khin *et al*. (2011) and Yaziz *et al*. (2011). [5]For use of other linear, non-linear and nonparametric methods in forecasting various types of variables, see the recent studies by Aydin (2009), Assis *et al*. (2010), Samsudin *et al*. (2010), Khin *et al*. (2011) and Yaziz *et al*. (2011). [6]The choice of these three variables are in line with the monetary policy literature. Gungor and Berk (2006), Agbeja (2007), Saibu and Oladeji (2007), Berument *et al*. (2009) and Krishnapillai and Thompson, (2012)

be called an unrestricted VAR model. Bayesian variable selection therefore, will be incorporated in Eq. 2 by embedding indicator variables: $\gamma = (\gamma_1 = \gamma_n)$ such that $\beta_i = 0$ if $\gamma_i = 0$ and $\beta_i \neq 0$ if $\gamma_i = 1$. Note that the indicator variables are treated as random variables by assigning a prior on them and allowing the data likelihood to determine their posterior values. These indicator variables can be explicitly inserted multiplicatively in the VAR model using the form:

$$Y_{i+1} = z_t\beta + \varepsilon_{i+1} \tag{3}$$

where, $\theta = \Gamma\beta$, $\Gamma$ is an (n×n) diagonal matrix with $\Upsilon_{JJ} = \gamma_J$ (j = 1,2...., n) elements on its main diagonal and for $\Upsilon_{JJ} = \gamma_J\ \theta_j = \Gamma_{jj}\beta_j = 0$ where, $\theta_j$ is restricted while for $\Upsilon_{JJ} = 1\ \theta_j = \Gamma_{jj}\beta_j = \beta_j$, so that all possible $2^n$ specifications can be explored and variable selection is equivalent to model selection in this case. Gibbs sampling can be used to estimate these parameters by conditioning on the data and $\Gamma$. Assuming the so-called independent Normal-Wishart prior, the densities of $\beta$ and $\Sigma$ are of standard form. The restriction indices $\gamma$ add one more block to the Gibbs sampler of the unrestricted VAR model, and if needed, for the restriction indicators the n element in the column vector $\gamma = (\gamma_t,...., \gamma_n)'$ is sampled and the diagonal matrix $\Gamma = \text{diag}\{\gamma_t,...., \gamma_n\}$ is recovered. Derivations are however, simplified if indicators $\gamma_j$ are independent of each other. The priors in particular can be defined as below:

$$\beta \sim N_n(b_0, V_0) \tag{4}$$

$$\gamma_j|\gamma_{\backslash j} \sim \text{Bernoulli}(1, \pi_{0j}) \tag{5}$$

$$\Sigma^{-1} \sim \text{Wishart}(\alpha, S^{-}1) \tag{6}$$

where, $b_0$ is (n×1), $V_0$ is (n×n), $\pi_0 = (\pi'_{01},...., \pi'_{0n})$ is (n×1), $\Omega$ is (m×m) matrix and $\alpha$ is a scalar.

It is argued that this form of variable selection may be adopted in many non-linear extensions of the VAR as compared to stochastic variable selection algorithms for VAR models. Adopting variable selection in TVP-VAR model therefore, is a simple extension of the VAR model with constant parameters, where Eq. 7 is replaced with Eq. 3 and variables are as explained in Eq. 3 while priors are as explained by Eq. 4 through 6 (except now $\beta \sim N_n(b_0, V_0)$.

Modern macroeconomic applications increasingly involve the use of VARs with mean regression coefficients and covariance matrices which are time-varying, in the process implying a nonlinear VAR model. Note that, a time-varying parameter VAR with constant variance (Homoscedastic VAR) takes the form:

$$Y_{t+1} = z\beta_t + \varepsilon_{t+1} \tag{7}$$

$$\beta_j = \eta_t \tag{8}$$

where, $z_t$, $x_t'$, $\Sigma$ and $\varepsilon_t$ are defined as before in Eq. 1; $\beta_t$ is an (n×1) vector of t = 1,....T parameters, $\eta_t \sim N(0, Q)$ with Q as a (n×n) covariance matrix. The implied prior for $\beta_1$ to $\beta_t$ are of the form $\beta_t|\beta_{t-1}, Q \sim N(\beta_{t-1}, Q)$ and the covariance matrix Q is considered to be unknown hence will have its own prior of the form $Q^{-1} \sim \text{Wishart}(\xi, R^{-1})$. In order to avoid the explosive behaviour (which might affect forecasting negatively) of the random walk assumption on the evolution of $\beta_t$, it is of importance to restrict its covariance Q. As such, to get a tight prior we subjectively choose the hyper-parameters for the initial condition $\beta_0$ and the covariance matrix Q. It is worth noting that the performance of variable selection is influenced by the hyper-parameters which affect the mean and variance of the mean coefficients $\beta$. For the VAR case, when $\gamma_j = 0$ and $\beta_j$ is restricted, a draw is taken from each prior implying that the prior variance $V_0$ cannot be very large[7] since it would mean no predictors are selected. Variable selection is also affected by the hyper-parameter of the Bernoulli prior of $\gamma_j$.

**Alternative forecasting models and forecast evaluation metric:** Specifically, the priors that we use for the restricted VAR i.e., VAR with variable selection (VAR-VS) are: $\gamma_j|\gamma_{\backslash j} \sim \text{Bernoulli}(1, 0.5)$ for all j = 1,...,N and $\beta_J \sim n(0, 10^2)$ if $\beta_J$ is an intercept, and $\beta_J \sim N(0, 3^3)$ otherwise. For the benchmark VAR, the priors are the same as the VAR-VS, except that we restrict $\gamma_j = 1$ for all j. As far as the BVAR based on the Minnesota prior (VAR-MIN) is concerned, the means and variances of the Minnesota prior for $\beta$ takes the form $\beta \sim N(0, 03^3)$ where:

$$\underset{i,j}{\text{VMIN}} = \begin{cases} g_1/p \\ g_3 \times S_i^2 \\ g2 \times s_i^2/(p \times s_i^2) \end{cases} \tag{9}$$

with $g_1/p$ and $g_3 \times s^2_i$ applying to parameters on own lags and for intercepts, respectively, while $g_3 \times s^2_i/(p \times s^2_i)$ is for parameters j on variable l ≠ i, l, i = 1,.., m. $s^2_i$ is the residual variance from the p-lag univariate autoregression for variable i. After experimenting to produce the best possible forecast, the hyperparameters were set to the following values: $g_1 = 0.09$, $g_2 = 0.0225$, and $g_3 = 100$. Since the variables used in the forecasting exercise is

[7]This section relies heavily on the discussion available by Korobilis (2011). The readers are referred to this paper for further details

transformed to induce stationarity, the prior mean vector $b_{MIN}$ is set equal to zero for parameters on the lags of all variables including the first own lag (Banbura *et al*., 2010).

For the time-varying parameters model with variable selection (TVP-VAR VS), a prior on the initial condition is of the form $\beta \sim N(04^2\ V_{MIN})$, with $\gamma_j|\gamma_{\backslash j} \sim$ Bernoulli (1, 05) The time-varying VAR without variable selection (TVP-VAR) uses prior as in TVP-VAR VS, with the restriction $\gamma_j = 1$ for all j = 1,...,n imposed. The covariance Q of the time-varying coefficients in the TVP-VAR VS has the prior $Q^{-1} \sim$ Wishart ($\xi$, R)where $\xi$ = n+1 and $R^{-1}$, where VMIN is the matrix defined in (9).

To evaluate the forecast accuracy, this study computes and compares the Mean Squared Forecast Error (MSFE) of one-through eight-quarters-ahead recursive out-of-sample forecasts for the period 2000:Q1 to 2011:Q2 in all the models. The covariance was integrated out using an uninformative prior of the form $p(\Sigma) \propto |\Sigma|^{-(m+1)/2}$ which is equivalent to prior defined by equation and an additional restriction is that $\alpha = 0$ and $S^{-1} = 0_{m \times m}$. All models are based on a run of 20,000 draws from the posterior, discarding the first 10,000 draws. The MSFE is computed as:

$$MSFE^h_{i,\tau} = \sqrt{(\hat{y}_{i,t+h|t} + y^0_{i,t+h})^2} \qquad (10)$$

where, $\hat{y}_{i,t+h|t}$ is the time t+h prediction of the variable i created using data available up to time t. $y^0_{i,t+h}$ is the observed value at time t+h. For the TVP models[8], Averages over the full forecasting period 2000:Q1 to 2011:Q2 are presented using the formula:

$$(\widehat{MSFE})^h_i = \frac{1}{\tau_1 - h - \tau_0}\sum_{t=\tau_0}^{\tau_1-h} MSFE^h_{i,j} \qquad (11)$$

where, $\tau_0$ is 2000:Q1 and $\tau_1$ is 2011:Q2.

**Data:** This study estimated the different models for the South African economy using quarterly data for the period 1960:Q2 to 2011:Q2. The macroeconomic variables of interest were: GDP growth rate (quarter on quarter percentage growth rate of the seasonally adjusted Gross Domestic Product at 2005 constant prices), the inflation rate (the quarter-to-quarter percentage change in the consumer price index) and the interest rate (yield on three month treasury bill rate). The data on treasury bill rate and consumer price index were obtained from the International Financial Statistics of International Monetary Fund, while GDP data was obtained from the Quarterly Bulletin of the South African reserve bank. Figure 1 shows the graphs of the three variables used in our forecasting exercise.

Since, the interest rate and inflation rate were found to be non-stationary (based on standard unit root tests)[9], the analysis uses the first difference of these variables, unlike the growth rate of the real GDP. After the transformations, one quarter was lost at the beginning of

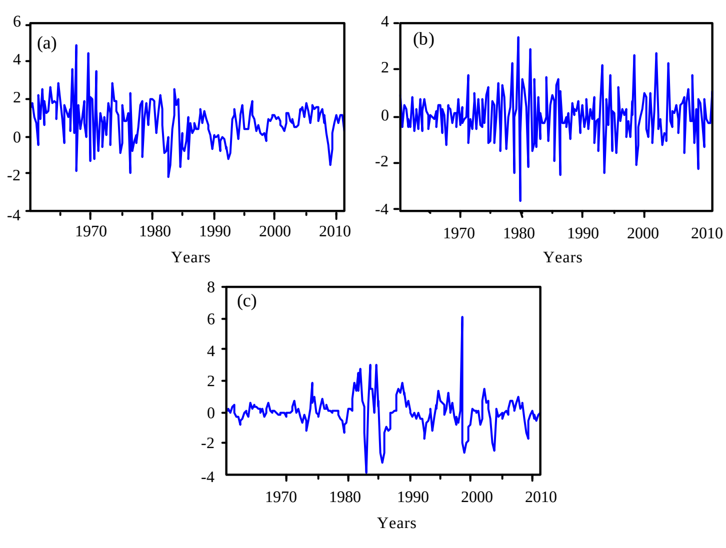


Fig. 1(a-c): Transformed key macroeconomic variables of the South African economy: 1960:Q2 to 2011:Q2 (a) GPD growth rate, (b) Δ inflation rate and (c) Δ interest rate

[8]For a recent empirical application using a TVP-VAR based on stochastic volatility for forecasting key macroeconomic variables of the US economy, see D'Agostino *et al.* (2011)

the sample, and hence, the in-sample contains data from 1960:Q2 to 1999:Q4, while the one-through eight-quarters-ahead out-of sample forecast is obtained from the out-of-sample period of 2000:Q1 to 2011:Q2, by, recursively estimating each of the six models, namely, the random walk (RW), VAR, VAR-MIN, VAR VS, TVP-VAR and TVP-VAR VS. The appropriate lag length was selected using the Akaike information criterion (AIC), which, in turn, yielded 3 lags[10]. Hence, the period 1960:Q2-1961:Q1 was used to feed the lags in the alternative VAR models.

## RESULTS

The findings emanating from the forecasting evaluation exercise, as presented in the Table 1 can be summarized as follows. Note, in line with the series of studies involving BVAR models for the South African economy reported in the introduction, the models are compared in terms of the average relative MSFE, i.e., the MSFE of a specific model with respect to the MSFE of the RW model[11]:

- The results show that with distant forecasts, the naive RW model performs worse than all the models whether restricted or unrestricted. It is always ranked sixth in terms of average MSFE for the three key variables of our concern
- For the GDP growth rate, on average, the VAR-VS model performs the best, followed closely by the TVP-VAR-VS. The TVP-VAR model comes in third, while, the VAR-MIN and the VAR model ends up being the fourth and fifth best performer
- For the change in the inflation rate, the VAR-VS model is again the best performer, as was the case with the growth rate. The TVP-VAR ranks a close second, while, the VAR-MIN, follows closely on the

Table 1: One-to Eight-Quarters-Ahead Out-of-Sample MSFE (2000:Q1-2011:Q2)

| | | RW | VAR | VAR-MIN | VAR-VS | TVP-VAR | TVP-VAR-VS |
|---|---|---|---|---|---|---|---|
| h = 1 | GDP | 0.249 | 1.303 | 1.240 | 1.287 | 1.286 | 1.178 |
| | ΔInflation | 2.171 | 0.444 | 0.413 | 0.424 | 0.425 | 0.446 |
| | ΔInterest rate | 0.458 | 0.868 | 0.768 | 0.831 | 0.835 | 0.780 |
| h = 2 | GDP | 0.498 | 0.578 | 0.600 | 0.566 | 0.567 | 0.606 |
| | ΔInflation | 2.364 | 0.382 | 0.374 | 0.375 | 0.377 | 0.396 |
| | ΔInterest rate | 1.017 | 0.599 | 0.536 | 0.579 | 0.577 | 0.524 |
| h = 3 | GDP | 0.667 | 0.504 | 0.520 | 0.492 | 0.498 | 0.525 |
| | ΔInflation | 2.034 | 0.497 | 0.475 | 0.482 | 0.483 | 0.494 |
| | ΔInterest rate | 1.243 | 0.496 | 0.461 | 0.465 | 0.464 | 0.445 |
| h = 4 | GDP | 0.809 | 0.517 | 0.505 | 0.494 | 0.497 | 0.491 |
| | ΔInflation | 1.692 | 0.556 | 0.546 | 0.535 | 0.535 | 0.544 |
| | ΔInterest rate | 1.307 | 0.456 | 0.433 | 0.428 | 0.426 | 0.430 |
| h = 5 | GDP | 0.889 | 0.488 | 0.479 | 0.463 | 0.469 | 0.471 |
| | ΔInflation | 2.290 | 0.396 | 0.397 | 0.387 | 0.387 | 0.392 |
| | ΔInterest rate | 1.469 | 0.405 | 0.391 | 0.378 | 0.378 | 0.391 |
| h = 6 | GDP | 0.894 | 0.503 | 0.492 | 0.478 | 0.487 | 0.487 |
| | ΔInflation | 2.295 | 0.413 | 0.411 | 0.405 | 0.405 | 0.413 |
| | ΔInterest rate | 1.441 | 0.413 | 0.404 | 0.386 | 0.386 | 0.404 |
| h = 7 | GDP | 0.875 | 0.537 | 0.523 | 0.510 | 0.520 | 0.520 |
| | ΔInflation | 2.298 | 0.419 | 0.418 | 0.402 | 0.402 | 0.410 |
| | ΔInterest rate | 1.152 | 0.548 | 0.525 | 0.508 | 0.507 | 0.514 |
| h = 8 | GDP | 0.778 | 0.617 | 0.599 | 0.590 | 0.602 | 0.604 |
| | ΔInflation | 1.025 | 0.944 | 0.936 | 0.932 | 0.932 | 0.952 |
| | ΔInterest rate | 1.063 | 0.587 | 0.564 | 0.567 | 0.565 | 0.573 |
| Average | GDP | 0.7074 | 0.6308 | 0.6199 | 0.6098 | 0.6156 | 0.6102 |
| | ΔInflation | 2.0211 | 0.5062 | 0.4962 | 0.4927 | 0.4931 | 0.506 |
| | ΔInterest rate | 1.1436 | 0.5465 | 0.5103 | 0.5177 | 0.5173 | 0.5076 |

Models as defined in the text; The third column reports the MSFE from the Random Walk (RW) model, while columns 4 to 8 presents the ratio of the MSFE of a specific model relative to the MSFE of the RW model; h denotes the forecasting horizon; Average denotes the average MSFE of the RW model and the relative MSFE of the other models for h = 1 to 8 for a specific variable. For averages, we report up to four decimal places to distinguish between the models, since some of the models produce the same average at three decimal places. Bold entries indicate the model with the lowest average relative MSFE

[9]The unit root tests, namely the Augmented-Dickey-Fuller (ADF), the Dickey-Fuller with GLS detrending (DF-GLS), the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) and the Phillips-Perron (PP) tests, are available upon request from the authors. [10]Using 2 lags based on the Schwarz information criterion (SIC), did not change our results qualitatively. These results are available upon request from the authors. [11]Results based on the mean absolute forecast error (MAFE) yielded similar conclusions. Also, when a longer out-of-sample period starting in 1981:Q1 was used, we obtained results similar to those reported in Table 1. Both sets of results are available upon request from the authors

heels. The TVP-VAR-VS and the VAR comes in fourth and fifth to round off the list

- As far as the change in the short-term interest rate is concerned, the TVP-VAR-VS outperforms all the other models. The VAR-MIN comes in second followed by the TVP-VAR, VAR-VS and the VAR models
- As observed in the literature (Gupta and Sichei, 2006; Gupta, 2006, 2007, 2009; Gupta and Das, 2008; Liu and Gupta, 2007; Liu *et al*., 2009, 2010; Balcilar *et al*., 2011) of forecasting with BVAR model based on the Minnesota prior (VAR-MIN), the model tends to outperform the classical VAR in our case as well
- Variable selection, whether imposed on a time-varying VAR or fixed parameter VAR, is found to play a role in improving forecast performances. Thus, highlighting that there could be gains in using other forms of efficient methods in solving the overparameterization problem of the classical VAR, besides the standard Minnesota-prior-based shrinkage approach
- Nonlinearity, modelled through the TVP VARs, also clearly play an important role in improving predictions when compared to the linear fixed coefficients classical VAR
- Having said that, the results do not suggest marked gains in terms of the relative average MSFE across the different Bayesian models. In fact, the improvement of the average relative MSFE over the classical VAR model made by its Bayesian counterparts is not significantly large (3.33, 2.67 and 7.11%, respectively for the GDP growth rate and the first differences of the inflation rate and the interest rate). However, note the classical VAR is outperformed by the best performing Bayesian VAR for a specific variable for each of the one- to eight-quarters-ahead forecasts
- One reason behind the result that one does not see significant gains by using Bayesian variants of the VAR model over the classical VAR could be because of the fact that the problem of over parameterization is not that acute for the system in this study. The current system has 30 parameters to be estimated in all, involving one constant and three lags each for the three variables, implying 10 parameters for each of the 3 equations. It is likely that the gains would be bigger for large-scale models involving more than 10 to 15 variables, as observed by Korobilis (2011)

## CONCLUSION

The Vector Autoregressive (VAR) model, though 'atheoretical' is particularly useful for forecasting purposes. One drawback of VAR models is that many parameters are needed to be estimated, some of which may be insignificant. This problem of overparameterization, resulting in multicollinearity and loss of degrees of freedom leads to inefficient estimates and large out-of-sample forecasting errors. One of the most common approaches to overcome this overparameterization is based on using Bayesian shrinkage, popularly called the Minnesota-prior-based Bayesian VAR (BVAR). Not surprisingly, in the literature, the BVAR models have been found to produce the most accurate short- and long-term out-of-sample forecasts relative to both univariate and multivariate unrestricted classical VAR models. In this regard, evidence for South Africa is no different, with a large number of recent studies showing superior forecasting power of BVAR models relative to classical VAR models. Now-a-days though, besides the shrinkage approach of the Minnesota-type BVAR models, there are numerous other efficient methods to prevent the proliferation of parameters and eliminate parameter or model uncertainty, based on stochastic search algorithm for variable selection.

Against this backdrop, this study compared the forecasting performances of the classical and the Minnesota-type BVAR models with those of linear (fixed-parameter) and nonlinear (time-varying parameter [TVP]) VARs involving a stochastic search algorithm for variable selection, estimated using Markov Chain Monte Carlo (MCMC) methods. Specifically speaking, this study analyzes the forecasting performances of all these models in predicting one- to eight-quarters-ahead of the growth rate of GDP, the Consumer Price Index (CPI) inflation rate and the three months Treasury bill rate for South Africa over an out-of-sample period of 2000:Q1-2011:Q2, using an in-sample period of 1960:Q1-1999:Q4.

The results suggest that, the VAR based on variable selection performs the best for forecasting output growth and inflation, while the time-varying VAR is the best model in forecasting the interest rate. In general, the study finds that variable selection, whether imposed on a time-varying VAR or a fixed parameter VAR, is found to play a role in improving forecast performances. Nonlinearity modelled through the TVP-VARs also play an important part in improving predictions when compared to the linear fixed coefficients classical VAR. Similar results were also obtained by Korobilis (2011). However, the results do not indicate marked gains in forecasting power across the different Bayesian models, as well as, over the classical VAR model. One reason behind the result could be because of the fact that the problem of over parameterization in the classical VAR is not that acute for the small-system in this study. It is likely that the gains would be bigger for large-scale models involving more than 10 to 15 variables-an area of research, which is for the future.

## REFERENCES


Agbeja, O., 2007. Central bank independence and inflation targeting: Monetary policy framework for Sub-Saharan Africa. Int. J. Applied Econ. Finance, 1: 29-37.

Alpanda, S., K. Kotze and G. Woglom, 2011. Forecasting performance of an estimated DSGE model for the South African economy. South Afr. J. Econ., 79: 50-67.

Anderson, M.K. and S. Karlsson, 2008. Bayesian Forecast Combination for VAR Models. In: Bayesian Econometrics: Advances in Econometrics, Chib, S., W. Grifitths, G. Koop and D. Terrell (Eds.). Elsevier, Oxford, pp: 501-524.

Assis, K., A. Amran, Y. Remali and H. Affendy, 2010. A comparison of univariate time series methods for forecasting cocoa bean prices. Trends Agric. Econ., 3: 207-215.

Aydin, D., 2009. A comparative study of neural networks and non-parametric regression models for trend and seasonal time series. Trends Applied Sci. Res., 4: 126-137.

Balcilar, M., R. Gupta and Z.B. Shah, 2011. An in-sample and out-of-sample empirical investigation of the nonlinearity in house prices of South Africa. Econ. Model., 28: 891-899.

Banbura, M., D. Giannone and L. Reichlin, 2010. Large Bayesian VARs. J. Applied Econ., 25: 70-92.

Berument, M.H., A. Sahin and B. Saracoglu, 2009. The choice of monetary policy tool(s) and relative price variability: Evidence from Turkey. J. Applied Sci., 9: 2238-2246.

D'Agostino, A., L. Gambetti and D. Giannone, 2011. Macroeconomic forecasting and structural change. J. Applied Econ. 10.1002/jae.1257.

Doan, T., R. Litterman and A.C. Sims, 1984. Forecasting and conditional projection using realistic prior distributions. Econ. Rev., 3: 1-100.

George, E.I., D. Sun and S. Ni, 2008. Bayesian stochastic search for VAR model restrictions. J. Econ., 142: 553-580.

Gungor, C. and A. Berk, 2006. Money supply and inflation relationship in the Turkish economy. J. Applied Sci., 6: 2083-2087.

Gupta, R., 2006. Forecasting the South African economy with VARs and VECMs. South African J. Econ., 74: 611-628.

Gupta, R. and M.M. Sichei, 2006. A BVAR model for the South African economy. S. Afr. J. Econ., 74: 391-409.

Gupta, R., 2007. Forecasting the South African economy with Gibbs sampled BVECMs. S. Afr. J. Econ., 75: 631-643.

Gupta, R. and S. Das, 2008. Spatial Bayesian methods of forecasting house prices in six metropolitan areas of South Africa. S. Afr. J. Econ., 76: 298-313.

Gupta, R., 2009. Bayesian methods of forecasting inventory investment in South Africa. South Afr. J. Econ., 77: 13-24.

Gupta, R. and S. Das, 2010. Predicting downturns in the US housing market: A Bayesian approach. J. Real Estate Finance Econ., 41: 294-319.

Gupta, R. and A. Kabundi, 2010. Forecasting macroeconomic variables in a small open economy: A comparison between small- and large-scale models. J. Forecasting, 29: 168-185.

Gupta, R. and A. Kabundi, 2011a. A large factor model for forecasting macroeconomic variables in South Africa. Int. J. Forecasting, 27: 1076-1088.

Gupta, R. and A. Kabundi, 2011b. Forecasting macroeconomic variables using large datasets: Dynamic factor model versus large-scale BVARs. Indian Econ. Rev., 46: 23-40.

Khin, A.A., M. Zainalabidin and S.M. Nasir, 2011. Comparative forecasting models accuracy of short-term natural rubber prices. Trends Agric. Econ., 4: 1-17.

Koop, G. and D. Korobilis, 2010. Bayesian multivariate time series methods for empirical macroeconomics. Found. Trends Econometrics, 3: 267-358.

Korobilis, D., 2008. Forecasting in Vector Autoregressions with Many Predictors. In: Bayesian Econometrics, Chib, S., W. Griffiths and G. Koop (Eds.). Vol. 23. Emerald Group Publishing, UK., ISBN-13: 9781848553088, pp: 403-431.

Korobilis, D., 2011. VAR forecasting using Bayesian variable selection. J. Applied Econometrics.

Krishnapillai, S. and H. Thompson, 2012. Pass-through effects of exchange rate and monetary policy in Sri Lanka. Int. J. Applied Econ. Fin., 6: 29-36.

Litterman, R.B., 1981. A Bayesian procedure for forecasting with vector autoregressions. Working Paper, Federal Reserve Bank of Minneapolis, Minneapolis, USA.

Litterman, R.B., 1986. Forecasting with Bayesian vector autoregressions: Five years of experience. J. Bus. Econ. Stat., 4: 25-38.

Liu, D.G. and R. Gupta, 2007. A small-scale DSGE model for forecasting the South African economy. S. Afr. J. Econ., 75: 179-193.

Liu, D.G., R. Gupta and E. Schaling, 2009. A new-keynesian DSGE model for forecasting the South African economy. J. Forecasting, 28: 387-404.

Liu, D.G., R. Gupta and E. Schaling, 2010. Forecasting the South African economy: A hybrid-DSGE approach. J. Econ. Stud., 37: 181-195.

Saibu, M.O. and S.I. Oladeji, 2007. Asymmetric policy shocks and real output fluctuations in Nigeria (1960-2004). The Int. J. Applied Econ. Finance, 1: 88-96.

Samsudin, R., A. Shabri and P. Saad, 2010. A comparison of time series forecasting using support vector machine and artificial neural network model. J. Applied Sci., 10: 950-958.

Spencer, D.E., 1993. Developing a Bayesian vector autoregression model. Int. J. Forecasting, 9: 407-421.

Stock, J.H. and M.W. Watson, 2005. Implications of dynamic factor models for VAR analysis. Mimeo, Princeton University.

Todd, R.M., 1984. Improving economic forecasting with Bayesian vector autoregression. Quarterly Review, Federal Reserve Bank of Minneapolis, Fall, 18-29.

Villani, M., 2009. Steady state priors for vector autoregressions. J. Applied Econ., 24: 630-650.

Yaziz, S.R., M.H. Ahmad, L.C. Nian and N. Muhammad, 2011. A comparative study on Box-Jenkins and garch models in forecasting crude oil prices. J. Applied Sci., 11: 1129-1135.