Scholars Journal of Physics, Mathematics and Statistics **ISSN 2393-8056 (Print)**
Sch. J. Phys. Math. Stat. 2014; 1(1):4-21 **ISSN 2393-8064 (Online)**

(An International Publisher for Academic and Scientific Resources)

A Simulation Study of Effects of Collinearity on Forecasting of Bivariate Time Series Data

Adenomon M. O*[1], Oyejola B. A.[2]
[1]Department of Mathematics and Statistics, The Federal Polytechnic, Bida, Niger State
[2]Department of Statistics, University of Ilorin, Kwara State

***Corresponding Author:**
Adenomon M. O
Email: admonsagie@gmail.com

Abstract: Vector Autoregression (VAR) and Bayesian VAR (BVAR) were first introduced in the early eighties of the last century and have since proven to be practical and effective economic forecasting methodologies. This paper tends to study the forecasting performances of the unrestricted VAR model and four versions of the Sims-Zha BVAR models in the presence of collinearity using Monte Carlo simulation technique. We considered ten (10) collinearity levels (0.8, -0.8, 0.85, -0.85, 0.9, -0.9, 0.95, -0.95, 0.99 and -0.99), the results from our simulation study revealed that the forecasting performances of the models vary as the collinearity levels varied. Furthermore, the values of the criteria increases as the time series length and the collinearity levels increased. We therefore recommend that if VAR modelers know that collinearity is acting upon the model, one can choose the forecasting model that is preferred for the criteria and the time series length selected.
Keywords: Simulation, Collinearity, Forecasting, Time Series, Vector Autoregression (VAR) and Bayesian VAR (BVAR).

INTRODUCTION

The role of forecasting in decision making for economic planning and development cannot be overemphasized, that is why economist, statisticians, econometricians and mathematicians are seeking ways to improve forecast performance. For instance Schiff and Phillips, [1] applied recent time series method to the problem of forecasting New Zealand's real GDP. Robertson and Tallman[2] revealed that the Federal-Fund rates forecast can be effectively improved using the Bayesian VAR model. Clements and Hendry, [3] investigated forecasting with difference-stationary and trend-stationary models where they considered these models in a situation of misspecification and investigated the autocorrelation errors. Caraiani [4] applied the Bayesian VAR model in forecasting Romanian GDP. Stock and Watson [5] studied generalized shrinkage methods for forecasting many predictors; just to mention a few.

Hendry [6] noted that the success of econometric model based forecast depends upon the following, there being regularities to be capture; such regularities being informative about the future; the proposed method capturing those regularities; and excluding non-regularities that swap the regularities.

Clements and Hendry [7] enumerated a number of distinct forecasting methods, including: guessing (which relies on luck); extrapolation (which relies on persistence); leading indicators (which relies on the indicators continuing to lead systematically); surveys (which relies on plans being implemented); analysis 'in the context of an implicit, perhaps informal model' (which relies on the adequacy of the postulated framework); time series models such as the ARIMA class and Vector Autoregressions (VARs) (which relies on the 'continuing of the time series representation); and econometrics systems (which relies on the model capturing the invariants of the economic structure. In addition, Sims[8] throws more light on macroeconomics and its methodology while Phillips[9] revealed the laws and limits of econometrics.

However the focus of this paper is on Vector Autoregression (VAR) models, which is the most common multivariate time series form.

VAR and Bayesian VAR (BVAR) were first introduced in the early eighties of the last century and have since proven to be practical and effective economic forecasting methodologies [10-13]. (On one hand VAR models explain the endogenous variable solely by their own history, apart from the deterministic regressors by incorporating non-statistical a priori information, while on the other hand BVAR model employs mixed estimation, that is, a technique that blends actual data with stochastic prior information. VAR and BVAR models have typically been used in cases where long histories are available for each time series, the techniques are also appropriate when data are limited (that is short terms time series). In particular, BVAR models have been shown in numerous studies to be as accurate, if not more accurate, than large structural models and other time series methodologies [11, 2].

Gujarati, [14] observed that multicollinearity problem usually afflict the VAR models. In a more recent literature, it was reported that correlation coefficients $|r| > 0.7$ was an appropriate indicator for when collinearity begins to severely distort model estimation and subsequent prediction [15].

Therefore in this paper we want to explore the forecasting performances of the reduced form VAR model and the Sims-Zha BVAR model in the presence of high collinearity levels using Monte Carlo simulation technique.

REVIEW OF RELATED LITERATURES

The name “Multicollinearity” was first introduced by Ragnar Frisch. In his original formulation the economic variables are supposed to be composed of two parts, a systematic or “true” and an “error” components. This problem which arises when some or all the variables in the regression equation are highly intercorrelated and it becomes also impossible to separate their influences and obtain the corresponding estimates of the regression coefficient [16]. Multicollinearity is a term that refers to correlation among independent variables in a multiple regression model; it is usually invoked when some correlations are “large” but an actual magnitude test is not well defined [17].

Blalock [18] considered a situation where two independent variables are highly correlated, it was reported that it will be difficult to assess their relative importance in determining some dependent variable. It was also reported that the higher the correlation between independent variables the greater the sampling error of the partials and the implications for social research was also discussed.

Schink and Chiu [19] considered the classical least squares (LS), Limited Information Single equation (LISE) and Two stage least Squares (2SLS) in the presence of multicollinearity and autocorrelation error terms for small sample situation. They reported that using the RMSE or Standard error criteria, either the LISE or 2SLS would give similar results. While on the other hand, using bias criterion, the LISE may be a more accurate method than the 2SLS. They recommended that if econometricians know that multicollinearity and autocorrelation are acting upon the model, one can choose the estimating technique that is best for the criteria that has been selected.

Rama-Sastry,[16] reported some limits in the theory of multicollinearity on the parameters of a multiple regression model.

Harvey,[20] made the following comments on multicollinearity in regression. They includes: Multicollinearity is a problem of degree rather than kind, except when the term is taken to refer to extreme multicollinearity. Although multicollinearity may be present in misspecified regression model, multicollinearity as such should not be regarded as evidence of misspecification, and lastly, in a linear model, prior detrending gives exactly the same results as does multiple regression applied to the full model.

Atkinson, [21] performed a Monte-Carlo experiment to determined whether multicollinearity systematically affects the relative rankings of simultaneous equation techniques. Three single-equation techniques namely Ordinary Least squares (OLS), Two-Stage Least Squares (2SLS), and Limited Information Maximum Likelihood (LIML), two full information techniques namely Three-Stage Least squares (3SLS) and Linearized Maximum Likelihood (LML) were considered under small-sample properties at different levels of multicollinearity. He concluded that since high collinearity is unlikely, 3SLS was suggested for general use.

Mittelhammer et al [22] studied OLS regression with exact linear restriction, mixed estimation (or regression with stochastic linear restriction) and principal components regression (PCR) as an alternative techniques to mitigate the effects of serious multicollinearity. In their work, they applied the techniques to the estimation of an aggregate agricultural production function for Thailand using time series data from 1950 to 1976.

Sharma and James [23] introduced to marketing fields a biased estimation procedure called “Latent root regression”. The procedure, unlike other biased estimation procedures not only provide stable estimates in the presence of multicollinearity, but also provides a measure for determining whether or not a biased estimation was appropriate.

Jagpal, [24] proposed a ridge estimator for the treatment of multicollinearity in structural equation models with unobservable variables. The method was then applied to a simple model of advertising in the multi product firm.

Salinas and Hillmer,[25] investigated the nature of muticollinearity associated with the design matrix **X** associated with the trading-day variation model.

Burt, [26] showed analytically that differencing time series data for the purpose of reducing multicollinearity in the data set for the independent variables of a regression equation cannot possibly succeed when its effect on the disturbance term is taken into account. The work further revealed that the intuitive basis to justify first differencing of multicollinear data contained a flaw, even when the effects of differencing the disturbance terms are ignored.

Mehta, Swamy and Iyengar,[27] proposed a statistic that measure the 'distance' of a cross-product matrix from the diagonal matrix obtained by zeroing its off-diagonal elements and they found it useful in detecting near multicollinearity regression problems. It can also distinguished between apparent and real multicollinearity with positive probability.

Buse,[28] worked out the statistical implications of the orthogonalization procedure in the general linear model. His work demonstrated that orthogonalization can worsen collinearity if measured by its effect on estimated variances.

Clements and Hendry,[7] reported that parameter estimates may be poorly determined in-sample due to the sheer number of variables, perhaps worsened by the high degree of collinearity manifested in the levels of integrated data.

Hendry, [6] investigated the non-uniqueness of collinearity using the static regression model and reported that any collinearity in the explanatory variables is irrelevant to forecasting so long as the marginal process remains constant.

Greenberg and Parks,[29] noted that since non-experimental data in general and economics data in particular, are often highly correlated, then from Bayesian viewpoint model specification is closely related to the problem of multicollinearity. In their approach they compared the predictive densities for an equation with and without the set of variables in question in order to gauge that the set may be safely omitted if the omission has little or no effect on the predictive densities. They concluded that examination of changes in predictive means and of Generalized Variance Ratio (GVR) is a useful method of investigation model specification.

Grewal, Cote and Baumgartner,[30] revealed through their Monte Carlo simulation experiment that multicollinearity can cause problems in theory testing (Type II errors) under certain conditions which includes: when multicollinearity is extreme; when multicollinearity is between 0.6 and 0.8, and when multicollinearity is between 0.4 and 0.5.

Friedman and Wall,[31] studied graphical views of suppressor variables and multicollinearity in multiple linear regression.

Alabi, et al [32] investigated the Type II error rate of the OLS estimators at different levels of multicollinearity and sample sizes through Monte-Carlo studies. Their work revealed that increasing the sample size reduces the type II error rate of the OLS estimator at all levels of multicollinearity.

Ayyangar, [33] considered the available options available to researchers when one or more assumptions of an ordinary Least Squares (OLS) regression model are violated. Ayyangar paid particular attention on the problems of skewness, multcollinearity and heteroskedasticity and autocorrelated error terms on OLS models using SAS with illustration to health care cost data.

Johnson [34] studied the effects of correlation and identification status on methods of estimating parameters of system of simultaneous equations model using Monte Carlo approach. The Monte Carlo approach for the performances of the estimating methods at different levels of correlation, sample sizes and identification status were reported.

Ijomah and Nduka,[35] considered the various performances of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mallow C_p statistic in the presence of a multicollinear regression using simulated data from SAS programme. Their work revealed that the performances of AIC and BIC in choosing the correct model among collinear variables are better when compared with the performances of Mallow's C_p.

Agunbiade, [36] studied the effects of multicollinearity and the sensitivity of the estimation methods in simultaneous equation model for different levels of different levels of multicollinearity. He considered Ordinary Least

Squares (OLS), 2 stage least squares (2SLS) and three Stage Least squares (3SLS) methods of estimation. His result revealed preference of 3SLS over 2SLS and OLS.

Taylor, [37] noted that multicollinearity problems are almost always present in time-series data generated by natural experiments. He also noted that multicollinearity becomes 'harmful' when there is an R^2 in the predictor matrix that is of the same order of magnitude as the R^2 of the overall model.

Olanrewaju,[38] studied the effect of multicollinearity on some estimators (Ordinary Least Squares, Cochran-Orcut (GLS2), Maximum Likelihood Estimator (MLE), Multivariate Regression, Full Information Maximum Likelihood, Seemingly Unrelated Regression (SUR) and Three Stage Least Squares (3SLS)). Results showed that multivariate regression, FIML, SUR and 3SLS estimators are preferred at all levels of sample size.

Yahya and Olaniran, [39] studied the performances of Bayesian Linear regression, Ridge Regression and OLS methods for modeling collinear data. Also Garba et al,[40] studied the effect of multicollinearity and other assumptions as it relate to panel data modeling.

Adenomon and Oyejola,[41] compared the forecasting performances of the Reduced form Vector Autoregression (VAR) and Sims-Zha Bayesian VAR (BVAR) in a situation where the Endogenous variables are collinear at different levels and at different short terms time series lengths. There simulation study revealed that the BVAR forecast seems to be superior.

Sources of multicollinearity includes: the data collection method employed; constraints on the model or in the population being sampled; model specification; and overdetermined model. Multicollinearity especially in time series data, may occur if the regressors included in the model share a common trend - that is, they all increase or decrease over time.

Gujarati, [14] identified some consequences of multicollinearity. They include:

(1) Although BLUE, the OLS estimators have large variances and covariances making precise estimation difficult.
(2) Because of consequence 1, the confidence interval tends to be much wider, leading to the acceptance of the (zero null hypothesis) and the t-ratio of one or more coefficients tends to be statistically insignificant.
(3) Although the t-ratio of one or more coefficients is statistically insignificant, R^2, the overall measure of goodness-of-fit can be very high.
(4) The OLS estimators and their standard errors can be sensitive to small change in the data.

MODEL DESCRIPTION

Vector Autoregression Model (VAR)

VAR methodology superficially resembles simultaneous equation modeling in that we consider several endogenous variables together. But each endogenous variable is explained by its lagged values and the lagged values of all other endogenous variables in the model; usually, there are no exogenous variables in the model [14].

Given a set of *k* time series variables, $y_t = [y_{it}, \ldots \mathrm{y}_{\mathrm{Kt}}]'$, VAR models of the form

$$y_t = A_1 y_{t-1} + \ldots + \mathrm{A}_\mathrm{p} y_{t-p} + u_t \qquad (1)$$

provide a fairly general framework for the Data General Process (DGP) of the series. More precisely this model is called a VAR process of order p or VAR(p) process. Here $u_t = [u_{1t} . \ . \ . \ \mathrm{u}_{\mathrm{Kt}}]'$ is a zero mean independent white noise process with non singular time invariant covariance matrix $\sum_u$ and the A_i are (kxk) coefficient matrices. The process is easy to use for forecasting purpose though it is not easy to determine the exact relations between the variables represented by the VAR model in equation (1) above [42]. Also, polynomial trends or seasonal dummies can be included in the model.

The process is stable if $\det(I_K - A_1 z - \ldots - A_p z^p) \neq 0 \text{ for } |z| \leq 1 \qquad (2)$

In that case it generates stationary time series with time invariant means and variance covariance structure. Therefore To estimate the VAR model, one can write a VAR(p) with a concise matrix notation as

$$Y = BZ + U \tag{3}$$

where $\mathrm{Y} = [\mathrm{y}_1, \ldots, \mathrm{y}_T]$, $\mathrm{Z}_{t-1} = \begin{bmatrix} y_{t-1} \\ . \\ . \\ . \\ y_{t-p} \end{bmatrix}$, $\mathrm{Z} = [\mathrm{Z}_o, \ldots, \mathrm{Z}_{T-1}]$

Then the Multivariate Least Squares (MLS) for B yields

$$\hat{\mathrm{B}} = (ZZ')^{-1} \mathrm{Z}'\mathrm{Y} \tag{4}$$

It can be written alternatively as

$$Vec(\hat{\mathrm{B}}) = ((ZZ')^{-1} Z \otimes I_k) Vec(Y) \tag{5}$$

where $\otimes$ denotes the Kronecker product and Vec the vectorization of the matrix Y. This estimator is consistent and asymptotically efficient. It furthermore equals the conditional Maximum Likelihood Estimator (MLE) [43].

As the explanatory variables are the same in each equation, the Multivariate least squares is equivalent to the Ordinary Least Squares (OLS) estimator applied to each equation separately, as was shown by Zellner [44].

In the standard case, the MLE estimator of the covariance matrix differs from the OLS estimator.

$$\text{MLE estimator} \quad \hat{\Sigma} = \frac{1}{\mathrm{T}} \sum_{t=1}^{T} \hat{\ell}_t \hat{\ell}'_t \tag{6}$$

OLS estimator for a model with a constant, k variables and p lags, in a matrix notation, gives

$$\hat{\Sigma} = \frac{1}{T - kp - 1} (Y - \hat{\mathrm{B}}Z)(Y - \hat{\mathrm{B}}Z)' \tag{7}$$

Therefore, the covariance matrix of the parameters can be estimated as

$$C\hat{o}v = (Vec(\hat{\mathrm{B}})) = (ZZ')^{-1} \otimes \hat{\Sigma} \tag{8}$$

BAYESIAN VECTOR AUTOREGRESSION WITH SIMS-ZHA PRIOR

The most popular BVAR model is that of the Litterman [45], although other priors have been study. For instance, Ni and Sun, [46] explored the properties of Bayesian estimates of Vector Autoregression (VAR) models under several possible choices of sampling distribution of data (normal and Student-t errors), loss functions (for Φ, Σ) and priors (Jeffreys prior, RAT prior, Yang and Bargers prior, Shrinkage prior and constant prior). They concluded that the choice of prior has stronger effect on the Bayesian estimates than the choice of loss function. In this line also, Ni, Sun and Sun [47] investigated the properties of the Bayesian estimates of impulse responses through an information-theoretic approach. They derived Bayesian estimators from an intrinsic entropy loss function and showed that they are distinctly different from the posterior mean. They also proposed an algorithm that uses generated data as latent variables in numerical simulation of Bayesian estimates under loss entropy loss.

However, in recent times, the BVAR model of Sims and Zha [48]) has gained popularity both in economic time series and political analysis. As stated in Brandt and Freeman [49], Litterman proposed BVAR for the reduced form of the model, while Sims-Zha specified prior for the simultaneous equation of the model. They further noted that Sims-Zha has more advantage compared to the BVAR proposed by Litterman. The Sims-Zha BVAR allows for a more general specification and can produce a tractable multivariate normal posterior distribution. Again, for the Litterman BVAR, the estimation of the VAR coefficients is done on an equation-by-equation basis as in the reduced form version while the Sims-Zha BVAR estimates the parameters for the full system in a multivariate regression.

We outline below other advantages of the Sims-Zha BVAR:

i. In a series of experiments with artificial and actual (macroeconomic) data, Sims and Zha showed that their Bayesian shape error bounds are more accurate in terms of location and skewness than the bands produced by other methods. In other words, Bayesians like Sims and Zha prefer 68% (approximately one standard deviation) coverage or posterior probability intervals to the more familiar 95% confidence interval.
ii. Regarding error bands of impulse responses, Sims and Zha's method is more indicative of the relevant range of uncertainty. Also Sims-Zha method is for identified Vector Autoregressions while overidentified models require

a modified approach to construct posterior probabilities for impulse responses. Finally, the method of Sims-Zha can be extended to any analysis in which one must characterize uncertainty about the values of an estimated function of time and uncertainties about the future values of this function are independent.

iii. Sims-Zha found that the BVAR developed by Litterman was incompatible with their beliefs about macroeconomy. Their belief is that macroeconomy is best described by a dynamic simultaneous model in which the belief (prior) are specified for the structural rather than the reduced form parameter. They substituted a Normal-inverse Wishart prior for the whole system of VAR coefficients for the Littterman equation-by-equation prior.

iv. For nonstationarity time series, Sims-Zha prior adds hyperparameters that capture beliefs about the sum of coefficients of lagged dependent variables (the number of unit roots in the system of variables) and about the possibility of cointegrations among these stochastic trends.

v. In terms of forecasting, many econometricians have compared the performance of the Sims-Zha prior to that of the Litterman prior and others, they found out that the Sims-Zha performs as well as or better than models of commercial forecasters.

vi. The idea of theoretical structure also surfaced in Bayesian time series forecasting. Sims-Zha showed how to incorporate fuller theoretically informed structural model of the innovations in the variables in Bayesian forecasting.

vii. Sims-Zha's approach yields a posterior distribution that can be easily sampled while Litterman equation-by-equation construction of the prior on the reduced form representation of the model does not.

viii. Lastly, Brandt and Freeman showed that the error bands computed using the eigenvector decomposition method suggested by Sims and Zha provide a better summary of the shape and likelihood of the responses than the alternatives.

The procedure for BVAR with Sims-Zha prior is as follows. We consider the following (identified) dynamic simultaneous equation model as

$$\sum_{l=0}^{p} \underset{1\times m}{y_{t-l}} \underset{m\times m}{A_l} = \underset{1\times m}{d} + \underset{1\times m}{\varepsilon_t} \qquad ; t = 1,2,\ldots T$$

This is an m-dimensional VAR for a sample of size T with y_t a vector of observations at time t, A_l the coefficient matrix for the l^{th} lag; p the maximum number of lags (assumed known), d a vector of constant and ε_t a vector of i.i.d normal structural shocks such that

$$\mathrm{E}[\varepsilon_t / y_{t-s}, s>0] = \underset{1\times m}{0} \text{ and } \mathrm{E}[\varepsilon_t' / y_{t-s}, s>0] = \underset{m\times m}{I}$$

The structural model can be transformed into a multivariate regression by defining A_0 as the contemporaneous conditions of the series and A_+ as a matrix of the coefficients on the lagged variables by $YA_0 + XA_+ = E$ where Y is Txm, A_0 is mxm, X is Tx(mp+1), A_+ is (mp+1)xm and E is Txm matrices.

To define the VAR in a compact form

$$a_0 = vec(A_0),\ a_+ = vec\begin{bmatrix} -A_1 \\ . \\ . \\ . \\ -A_p \\ d \end{bmatrix}, A = \begin{pmatrix} A_0 \\ A_+ \end{pmatrix},\ a = vec(A)$$

The VAR model can then be written as a linear projection of the residual by letting Z=[Y X], and $A = [A_0 / A_+]'$ is a conformable stacking of the parameters in A_0 and A_+:

$YA_0 + XA_+ = E$

$ZA = E.$

In order to derive the Bayesian estimator for this structural equation model, we have to examine the (conditional) likelihood function for normally distributed residuals

$$L(Y/A) \propto |A_0|^T \exp[-0.5tr(ZA)'(ZA)]$$

$$\propto |A_0|^T \exp[-0.5a'(I \otimes Z'Z)a].$$

The prior overall of the structural parameters has the form

$$\pi(a) = \pi(a_+/a_0)\pi(a_0)$$

$$\pi(a) = \pi(a_0)\phi(\tilde{a}_+, \Psi)$$

$\tilde{a}_+$ denotes the mean parameters in the prior for a_+, Ψ is the prior covariance for $\tilde{a}_+$ and $\phi(\,)$ is a multivariate normal density.

The posterior for the coefficients is then

$$q(A) \propto L(Y/A)\pi(a_0)\phi(\tilde{a}_+, \Psi)$$

$$\propto \pi(a_0)|A_0|^T |\psi|^{-0.5} \times \exp[-0.5(a_0'(I \otimes Y'Y))a_0$$

$$-2a_+'(I \otimes X'Y)a_0 + a_+'(I \otimes X'X)a_+ + \tilde{a}_+'\Psi\tilde{a}_+)]$$

The posterior is conditional multivariate normal, since the prior has a conjugated form. In this case, the posterior can be estimated by a multivariate seeming unrelated regression (SUR) model. The forecast and inferences can be generated by exploiting the multivariate normality of the posterior distribution of the coefficients. The normal conditional prior for the mean of the structural parameters is given by $E(A_+/A_0) = \begin{bmatrix} A_0 \\ 0 \end{bmatrix}$ while $V(A_+/A_0) = \Psi$ is the prior covariance matrix for $\tilde{a}_+$. Though complicated, it is specified to reflect the following general beliefs and facts about the series being model:

1. The standard deviations around the first lag coefficients are proportionate to all the other lags.
2. The weight of each variable's own lags is the same as those of other variables' lags.
3. The standard deviations of the coefficients of longer lags are proportionately smaller than those of earlier lags. (Lag coefficients shrink to zero over time and have smaller variance at higher lags.)
4. The standard deviation of the intercept is proportionate to the standard deviation of the residuals for the equation.
5. The standard deviation of the sums of the autoregressive coefficients should be proportionate to the standard deviation of the residuals for the respective equation (consistent with the possibility of cointegration).
6. The variance of the initial conditions should be proportionate to the mean of the series. These are "dummy initial observations" that capture trends or beliefs about stationarity and are correlated across the equations. The summary of the Sims-Zha prior is given in Table 1.

Table 1: Hyperparameters of Sims-Zha reference prior

Parameter	Range	Interpretation
λ_0	[0,1]	Overall scale of the error covariance matrix
λ_1	>0	Standard deviation around A_1 (persistence)
λ_2	=1	Weight of own lag versus other lags
λ_3	>0	Lag decay
λ_4	≥0	Scale of standard deviation of intercept
λ_5	≥0	Scale of standard deviation of exogenous variable coefficients
μ_5	≥0	Sum of coefficients/Cointegration (long-term trends)
μ_6	≥0	Initial observations/dummy observation (impacts of initial conditions)
v	>0	Prior degrees of freedom

Source: Brandt and Freeman, [49]

Each diagonal element of Ψ therefore corresponds to the variance of the VAR parameters. The variance of each of these coefficients is assumed to have the form

$\psi_{l,j,i} = \left(\dfrac{\lambda_0 \lambda_1}{\sigma_j l^{\lambda_3}}\right)^2$ for the element corresponding to the l^{th} lag of variable j in equation i.

The overall coefficient covariances are scaled by the value of error variances from m univariate AR(p) OLS regressions of each variable on its own lagged values, σ_j^2 for j=1, 2, . . .m. The parameter λ_0 sets an overall tightness across the elements of the prior on $\Sigma = A_0^{-1'} A_0^{-1}$. The hyperparameter λ_1 controls the tightness of the beliefs about the random walk prior or the standard deviation of the first lags. The l^{λ_3} term allows the variance of the coefficients on higher order lags to shrink as the lag length increases. The constant in the model receives a separate prior variance of $(\lambda_0 \lambda_4)^2$ and the prior variance on any exogenous variables is $(\lambda_0 \lambda_5)^2$. The Sims-Zha prior adds dummy observations to account for unit roots, trends, and cointegration which was not possible with the Litterman prior.

Given the reduced form model

$$y_t = c + y_{t-1} B_1 + . . . + y_{t-p} B_p + u_t$$

$$\text{where } c = dA_0^{-1},\ B_l = -A_l A_0^{-1},\ l = 1,2,...p,\ u_t = \varepsilon_t A_0^{-1} \text{ and } \Sigma = A_0^{-1'} A_0^{-1}$$

The matrix representation of the reduced form is given as

$$\underset{T\times m}{Y} = \underset{T\times(mp+1)}{X} \underset{(mp+1)\times m}{\beta} + \underset{T\times m}{U},\ U \sim MVN(0,\Sigma)$$

We can then construct a reduced form Bayesian SUR with the Sims-Zha prior as follows. The prior means for the reduced form coefficients are that B_1=I and B_2, . . . B_p=0. We assume that the prior has a conditional structure that is multivariate Normal-inverse Wishart distribution for the parameters in the model. To estimate the coefficients for the system of the reduced form model with the following estimators

$$\hat{\beta} = (\Psi^{-1} + X'X)^{-1}(\Psi^{-1}\overline{\beta} + X'Y)$$

$$\hat{\Sigma} = T^{-1}(Y'Y - \hat{\beta}'(X'X + \Psi^{-1})\hat{\beta} + \overline{\beta}'\Psi^{-1}\overline{\beta} + \overline{S})$$

where the Normal - inverse Wishart prior for the coefficients is

$$\beta / \Sigma \sim N(\overline{\beta}, \Psi) \text{ and } \Sigma \sim IW(\overline{S}, v)$$

This representation translates the prior proposed by Sims and Zha form from the structural model to the reduced form [49-51].

Since our focus in this research is to compare the forecasting performance of the reduced form VAR and the reduced form BVAR models, we will just mention the error band methods and their interval for constructing the impulse responses. The impulse responses provide a summary of the general trend and shapes of the responses. In Bayesian time series methods provides meaningful error bands for these impulse responses. The error band methods are presented in Table 2.

Table 2: Impulse response error band computation

Error Band Method	Error Band Interval
Gaussian approximation	$\hat{c}_{ij}(t) \pm z_\alpha \sigma_{ij}(t)$
Pointwise quantiles	$[c_{ij.\alpha/2}(t), c_{ij.(1-\alpha)/2}(t)]$
Gaussian Linear eigenvector	$\hat{c}_{ij}(t) \pm z_\alpha W_{.k} \sqrt{\lambda_k}$
Likelihood-based eigenvector	$\hat{c}_{ij} + \gamma_{k,0.16}, \hat{c}_{ij} + \gamma_{k,0.84}$
Likelihood-based stacked eigenvector	$\hat{c}_{ij} + \gamma_{k,0.16}, \hat{c}_{ij} + \gamma_{k,0.84}$
	(with γ_k computed from the stacked covariance)

Source: Brandt and Freeman, [49]

Brandt and Freeman [49] noted the following benefits of impulse responses from Sims-Zha methods.
(i) They provide a better and more intuitive representation of the dynamic of the series in the model than the AR representation.
(ii) The coefficients are a function of time and they provide a good method for seeing how the multivariate process behaves over time.
(iii) Constructing measures of uncertainty for the $C_{ij}(t)$ is difficult. Also, the $C_{ij}(t)$ are high dimensional and thus hard to summarize.

Lastly, we consider an h-step forecast equation for the reduced form VAR model

$$y_{T+h} = cK_{h-1} + \sum_{l=1}^{p} y_{T+1-l} N_l(h) + \sum_{j=1}^{h} \varepsilon_{T+j} C_{h-j}, \ \mathrm{h} = 1,2\ldots$$

where

$$\mathrm{K}_0 = I, K_i = I + \sum_{j=1}^{i} K_{i-j} B_j \ , \mathrm{i} = 1,2,\ldots$$

$$N_l(1) = B_l \ , l = 1,2,\ldots,p$$

$$N_l(h) = \sum_{j=1}^{h-1} N_l(h-j) B_j + B_{h+l-1} \ , l = 1,2\ldots,p, h = 2,3,\ldots$$

$$C_0 = A_0^{-1} \ , C_l = \sum_{j=1}^{i} C_{i-j} B_j \ , i = 1,2,\ldots$$

Where we use the convention that $B_j=0$ for j>p, C_l are the impulse response matrices for lag l, K_i describe the evolution of the constants in the forecasts, and $N_l(h)$ define the evolution of the autoregressive coefficients over the forecast horizon. The h-step forecast equation above gives the dynamic forecasts produced by a model with structural innovations.

Brandt et al. [52] used the structural Bayesian time series approach to evaluate Bystander, Follower, Accountability and credibility in a macropolitical economy with relation to Israel and Palestine conflict and intervention on the part of the United States. In their approach they further addressed the problems of model scale, endogeneity and specification uncertainty. They finally used the reduced Bayesian VAR models to forecast the Israeli-Palestinians conflict. The study revealed the ability of the Bayesian time series model to capture complex dynamic situations. Some literature on BVAR can also be found in Ciccarelli and Alessandro,[53].

SETTING OF HYPERPARAMETERS FOR BVAR MODEL WITH SIMS-ZHA PRIOR

The setting of hyperparameters for BVAR Model has received a lot of attention in Bayesian time series literature. For instance, in the work of Kadiyala & Karlsson, [54], the values of the hyperparameters were chosen based on the forecast performance over a calibration period. Also in Sims & Zha [48,51], and in Leeper, Sims & Zha[55], the Sims-Zha proposed a benchmark prior for empirical macroeconomics with values $\lambda_1 = 0.1, \lambda_3 = 1, \lambda_4 = 0.1, \lambda_5 = 0.07, \mu_5 = \mu_6 = 5$. Brandt & Freeman [49-50] also exploited the Sims-Zha prior in forecasting macro political dynamics. They found that the Sims-Zha prior performed well in forecasting macro political time series data. In addition, Brandt, Colaresi and Freeman [52] set the hyperparameter values for the Sims-Zha prior based on experience with events data and discussing with leading international relations scholars. Suppes[56] recommended that such priors from experience provide better fits. Brandt & Freeman[49] argue that the presence of unit roots and the special nature of the time series sample therefore argue against testing for prior. Brandt & Freeman stated that instead, prior should reflect our beliefs based on past analyses, history and expectations about the future. The prior should not then be estimated from the data, as this is only on realization of the data generation process. In addition, Brandt & Freeman,[52] noted that strictly speaking, empirical estimation of the prior is a violation of Bayesian philosophy: the subsequent prior-to-posterior updating would use the data twice (first in the prior, and again in the likelihood). The resulting inference would thus be ‘overconfident’. Brandt & Freeman also noted another complicating issue about the assessment of prior specification is the nature of the time series data itself. That time series data are not repeated sample. The classical inference is based on inferring something about a population from a sample of data while in time series, the sample is not, and the population contains both the future as well as the past.

In view of this, several authors have recommended in the literature how the hyperparameter can be formed. For instance, Brandt & Freeman (2006) recommended that the choice of the hyperparameters should come from both experience and theory. In addition, the selection of the parameters for the prior must not and should not depend on the data alone, but should be informed by the properties of the data and their dynamics. Brandt & Freeman also noted that if the prior is derived from the data, the resulting forecasts will be too closely mirrored the sample data rather than the population. They further suggest that the prior must be consistent with the data such that it reflect the general beliefs analyst have about the data's variation, dynamic properties, and the general interrelationship of the time series data.

Park, [57] and Brandt,[58] suggested a grid search for setting of the hyperparameters for the BVAR model. The idea behind this, is that BVAR models are fitted for several combinations of the hyperparameters and forecast are generated over some period ahead. The forecast are then compared with the set of data, and the root mean square error and the mean absolute error measures are computed. Then the BVAR model with the minimum RMSE and MAR is considered as the best fitted model. Lutkepohl,[59] also affirmed that in practice, different values for the hyperparameters are sometimes tried. All these also follow the philosophy of the time series analysis which says "Let the data speak" [14]. In line with previous work and recommendation in the literature formed our choice of the hyperparameters used in this work.

SIMULATION PROCEDURE

The simulated data will be generated for time series lengths of 16, 32 and 64. The choice of the length chosen is to be able to study the models in the short run [60]. We also considered ten (10) multicollinearity levels as ρ=(0.8, -0.8, 0.85, -0.85, 0.9, -0.9, 0.95, -0.95, 0.99, -0.99).

The simulation procedure is given in the following steps

Step1: we generated a VAR (2) process that obeys the following form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \begin{bmatrix} 5.0 \\ 10.0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1} + \begin{bmatrix} -0.3 & -0.7 \\ -0.1 & 0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t$$

Our choice for this form model is to obtain a stable process and a VAR process that is not affected by overparameterization[61].

Step2: let the desired correlation matrix be $R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ then the Choleski factor P is $P = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$

and then the simulated data in step 1 is pre-multiplied by the Choleski factor so that the simulated data is scaled to have the desired correlation level[60].

Step 3: the VAR and BVAR models of lag length of 2 was be used for modeling and forecasting simulated data to obtain the RMSE and MAE.

Step 4: Step 1 to step 3 was repeated for 1000 times, and the averages of the criteria were used to access the preferred model.

Table-3: Sample of Generated data for different levels of collinearity for T=16

ρ=0.8			ρ= -0.8		
	y1	y2		y1	y2
1	3.5158326	7.260136	1	6.4389612	1.150354
2	7.9156946	8.196781	2	10.8804266	-6.935628
3	2.8175090	7.797546	3	2.2493752	2.937686
4	3.2759166	5.495379	4	0.8037202	1.696078
5	1.7148837	7.074573	5	1.0537601	5.550753
6	4.4553466	8.248478	6	3.5617301	1.288237
7	-0.5267808	3.506520	7	-0.7925606	4.814620
8	0.2203239	4.865553	8	1.2181963	4.045936
9	5.1875188	10.452788	9	2.4056657	2.566840
10	4.7598151	7.718817	10	2.9355888	3.601462
11	-1.4372751	3.936642	11	0.9108888	2.832962
12	0.7850790	5.145656	12	-0.9623041	6.351088
13	1.8281761	5.690600	13	2.9435275	2.363781
14	0.3119562	5.509374	14	1.3780844	3.186184
15	2.7533802	6.839297	15	1.4095769	3.557838
16	2.6578408	6.709557	16	2.1246557	3.460277
Estimated correlation r= 0.8360603			Estimated correlation r= -0.9109091		

Model Specification

We will model the generated data using a VAR model with lag 2. The choice here is to avoid overfitted VAR model while the VAR and BVAR models of lag length of 2 will be used for modeling and forecasting purpose.

For the BVAR model with Sims-Zha prior, we will consider the following range of values for the hyperparameters given below and the Normal-Inverse Wishart prior.
We consider two tight priors and two loose priors as follows:

$$\begin{aligned}
&\text{The Tight priors are as follows}\\
&\text{BVAR1} = (\lambda_0 = 0.6, \lambda_1 = 0.1, \lambda_3 = 1, \lambda_4 = 0.1, \lambda_5 = 0.07, \mu_5 = \mu_6 = 5)\\
&\text{BVAR2} = (\lambda_0 = 0.8, \lambda_1 = 0.1, \lambda_3 = 1, \lambda_4 = 0.1, \lambda_5 = 0.07, \mu_5 = \mu_6 = 5)\\
&\text{The Loose priors are as follows}\\
&\text{BVAR3} = (\lambda_0 = 0.6, \lambda_1 = 0.15, \lambda_3 = 1, \lambda_4 = 0.15, \lambda_5 = 0.07, \mu_5 = \mu_6 = 2)\\
&\text{BVAR4} = (\lambda_0 = 0.8, \lambda_1 = 0.15, \lambda_3 = 1, \lambda_4 = 0.15, \lambda_5 = 0.07, \mu_5 = \mu_6 = 2)
\end{aligned}$$

where nμ is prior degrees of freedom given as m+1 where m is the number of variables in the multiple time series data. In work nμ is 3 (that is two (2) time series variables plus 1(one)).

Our choice of Normal-Inverse Wishart prior for the BVAR models follow the work of Kadiyala & Karlsson[54], that Normal-Inverse Wishart prior tends to performed better when compared to other priors. Our choice of the overall tightness $\lambda_0 = 0.6 \text{ and } 0.8$ is in line with work of Brandt, Colaresi and Freeman[52].

Methods of Estimation

The methods of estimation of the model are as follows:
The equation-by-equation seemingly unrelated regression (SUR) method is used to estimates the reduced form VAR model.

The multivariate seeming unrelated regression (SUR) method is used in the estimation of the Bayesian VAR models for just identified VARs[48].

Forecast Assessment

The following are the criteria for Forecast assessments used:

(1) Mean Absolute Error or Deviation (MAE or MAD) has a formular $MAE_j = \frac{\sum_{i=1}^{n} |e_i|}{n}$. This criterion measures deviation from the series in absolute terms, and measures how much the forecast is biased. This measure is one of the most common ones used for analyzing the quality of different forecasts.

(2) For the RMSE is given as $RMSE_j = \sqrt{\frac{\sum_{i}^{n} (y_i - y^f)^2}{n}}$ where y_i is the time series data and y^f is the forecast value of y [4].

For the two measures above, the smaller the value, the better the fit of the model[62].

In our simulation study, $RMSE = \frac{\sum_{j}^{N} RMSE_j}{N}$ and $MAE = \frac{\sum_{j}^{N} MAE_j}{N}$ where N=1000. Therefore, the model with the minimum RMSE and MAE result as the preferred model.

Statistical Packages (R)

In this study three procedures in the R package will be used. They are: Dynamic System Estimation (DSE)[63, 58,12].

DATA ANALYSIS AND INTERPRETATION OF RESULTS

We summarize our results using the ranks of the criteria for the level of autocorrelation and collinearity . For instance rank of 1 is used to denote the preferred forecasting model; rank of 2 is used to denote the 2nd the preferred forecasting model and so on. In our work we used two criteria for accessing the preferred forecasting model which includes Roots Mean Square (RMSE) and Mean Absolute Error (MAE). The values of the forecast assessment criteria are presented in Tables 4.1 to 4.10 at the appendix. The result revealed the following:

(1) The values of the criteria increase as a result of increase in the time series length.

(2) The values of the criteria for negative collinearity levels are greater than the values of the criteria for positive collinearity levels.

(3) The values of the criteria increase as result of increase in the collinearity levels (for both positive and negative collinearity levels)

In this section also, we presented the ranks of the forecasting model for each criterion in Tables 5 and 6 while the preference of the models are presented in Tables 5 and 7

Table 4: The Ranks of the models at different levels of collinearity and different Time Series Length using the RMSE Criterion

Time series length (T)		Multicollinearity levels (ρ)									
		Positive					Negative				
	Models	0.8	0.85	0.9	0.95	0.99	-0.8	-0.85	-0.9	-0.95	-0.99
16	VAR(2)	3	2	5	4	1	5	3	3	4	4
	BVAR1	5	1	3	5	3	1	2	1	3	2
	BVAR2	4	3	1	2	5	2	1	4	2	5
	BVAR3	2	5	2	1	2	4	5	5	1	3
	BVAR4	1	4	4	3	4	3	4	2	5	1
32	VAR(2)	3	1	2	4	1	1	4	2	5	1
	BVAR1	4	4	4	3	4	3	2	5	3	5
	BVAR2	5	5	1	1	2	4	1	3	2	4
	BVAR3	2	2	3	5	3	5	3	4	1	3
	BVAR4	1	3	5	2	5	2	5	1	4	2
64	VAR(2)	4	4	2	4	4	4	5	3	4	2
	BVAR1	1	1	1	3	3	1	1	4	1	3
	BVAR2	5	3	3	2	2	2	3	5	3	4
	BVAR3	3	2	4	5	1	3	2	2	5	5
	BVAR4	2	5	5	1	5	5	4	1	2	1

Table 5: Preferred Forecasting Models at Different Levels of Collinearity and Different Time Series using RMSE Criterion

Time series length(T)	Multicollinearity levels (ρ)									
	Positive					Negative				
	0.8	0.85	0.9	0.95	0.99	-0.8	-0.85	-0.9	-0.95	-0.99
16	BVAR4	BVAR1	BVAR2	BVAR3	VAR(2)	BVAR1	BVAR2	BVAR1	BVAR3	BVAR4
32	BVAR4	VAR(2)	BVAR2	BVAR2	VAR(2)	VAR(2)	BVAR2	BVAR4	BVAR3	VAR(2)
64	BVAR1	BVAR1	BVAR1	BVAR4	BVAR3	BVAR1	BVAR1	BVAR4	BVAR1	BVAR4

With reference to Table 5 the preferred models are reported using the RMSE Criterion.

At collinearity level of ρ=0.8, BVAR4 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred. While on the other hand, at collinearity level of ρ= -0.8, the BVAR1 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred.

At collinearity level of ρ=0.85, BVAR1 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred. While on the other hand, at collinearity level of ρ= -0.85, the BVAR2 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred.

At collinearity level of ρ=0.9, BVAR2 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred. While on the other hand, at collinearity level of ρ= -0.9, the BVAR1 model is preferred when T=16, while when T=32 and T=64 the BVAR4 model is preferred.

At collinearity level of ρ=0.95, BVAR3 model is preferred when T=16, and when T=32 the BVAR2 model is preferred, while when T=64 the BVAR4 model is preferred. While on the other hand, at collinearity level of ρ= -0.95, the BVAR3 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred.

At collinearity level of ρ=0.99, VAR(2) model is preferred when T=16 and T=32, while when T=64 the BVAR3 model is preferred. While on the other hand, at collinearity level of ρ= -0.99, the BVAR4 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred.

Table 6: The Ranks of the models at different levels of collinearity and different Time Series Length using the MAE Criterion

Time series length (T)		Multicollinearity levels (ρ)									
		Positive					Negative				
	Models	0.8	0.85	0.9	0.95	0.99	-0.8	-0.85	-0.9	-0.95	-0.99
16	VAR(2)	3	2	5	4	1	5	1	2	4	3
	BVAR1	5	1	3	5	5	1	5	1	1	4
	BVAR2	4	4	1	2	3	2	2	4	3	5
	BVAR3	2	5	4	1	2	4	3	5	2	2
	BVAR4	1	3	2	3	4	3	4	3	5	1
32	VAR(2)	3	1	3	4	1	1	1	4	4	1
	BVAR1	4	4	2	3	2	2	5	5	2	5
	BVAR2	5	5	1	2	3	5	2	3	3	4
	BVAR3	2	2	4	5	4	3	4	1	5	3
	BVAR4	1	3	5	1	5	4	3	2	1	2
64	VAR(2)	4	4	2	4	4	2	4	4	5	3
	BVAR1	1	1	1	3	2	1	1	5	1	2
	BVAR2	5	3	3	2	3	4	3	3	3	4
	BVAR3	3	2	4	5	1	3	2	2	4	5
	BVAR4	2	5	5	1	5	5	5	1	2	1

Table -7: Preferred Forecasting Models at different levels of collinearity and different Time Series Length using the MAE Criterion

Time series length(T)	Multicollinearity levels (ρ)									
	Positive					Negative				
	0.8	0.85	0.9	0.95	0.99	-0.8	-0.85	-0.9	-0.95	-0.99
16	BVAR4	BVAR1	BVAR2	BVAR3	VAR(2)	BVAR1	VAR(2)	BVAR1	BVAR1	BVAR4
32	BVAR4	VAR(2)	BVAR2	BVAR4	VAR(2)	VAR(2)	VAR(2)	BVAR3	BVAR4	VAR(2)
64	BVAR1	BVAR1	BVAR1	BVAR4	BVAR3	BVAR1	BVAR1	BVAR4	BVAR1	BVAR4

With reference to Table 7 the preferred models are reported using the MAE Criterion.

At collinearity level of ρ=0.8, BVAR4 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred. While on the other hand, at collinearity level of ρ= -0.8, the BVAR1 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred.

At collinearity level of ρ=0.85, BVAR1 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred. While on the other hand, at collinearity level of ρ= -0.85, the VAR(2) model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred.

At collinearity level of ρ=0.9, BVAR2 model is preferred when T=16 and T=32, while when T=64 the BVAR1 model is preferred. While on the other hand, at collinearity level of ρ= -0.9, the BVAR1 model is preferred when T=16, when T=32 the BVAR3 is preferred and when T=64 the BVAR4 model is preferred.

At collinearity level of ρ=0.95, BVAR3 model is preferred when T=16, when T=32 and T=64 the BVAR4 model is preferred. While on the other hand, at collinearity level of ρ= -0.95, the BVAR1 model is preferred when T=16 and T=64, while when T=32 the BVAR4 model is preferred.

At collinearity level of ρ=0.99, VAR(2) model is preferred when T=16 and T=32, while when T=64 the BVAR3 model is preferred. While on the other hand, at collinearity level of ρ= -0.99, the BVAR4 model is preferred when T=16 and T=64, while when T=32 the VAR(2) model is preferred.

SUMMARY

The evidences from our simulation study revealed the performances of the forecasting models at different levels of collinearity (both positive and negative) for different time series length.

Using the RMSE Criterion, for a collinearity level of ρ=0.8, the BVAR4 forecast is preferred except when T=64. For a collinearity level of ρ=0.85, the BVAR1 forecast is preferred except when T=32. For a collinearity level of ρ=0.9, the BVAR2 forecast is preferred except when T=64. For a collinearity level of ρ=0.99, the VAR(2) forecast is preferred except when T=64. for a collinearity level of ρ= -0.8, the BVAR1 forecast is preferred except when T=32. For a collinearity level of ρ= -0.85, the BVAR2 forecast is preferred except when T=64. For a collinearity level of ρ= -0.9, the BVAR4 forecast is preferred except when T=16. For a collinearity level of ρ= -0.95, the BVAR3 forecast is preferred except when T=64. For a collinearity level of ρ= -0.99, the BVAR4 forecast is preferred except when T=32.

While on the other hand, using the MAE Criterion, for a collinearity level of ρ=0.8, the BVAR4 forecast is preferred except when T=64. For a collinearity level of ρ=0.85, the BVAR1 forecast is preferred except when T=32. For a collinearity level of ρ=0.9, the BVAR2 forecast is preferred except when T=64. For a collinearity level of ρ=0.95, the BVAR4 forecast is preferred except when T=64. For a collinearity level of ρ=0.99, the VAR(2) forecast is preferred except when T=64. for a collinearity level of ρ= -0.8, the BVAR1 forecast is preferred except when T=32. For a collinearity level of ρ= -0.85, the VAR(2) forecast is preferred except when T=64. For a collinearity level of ρ= -0.95, the BVAR1 forecast is preferred except when T=32. For a collinearity level of ρ= -0.99, the BVAR4 forecast is preferred except when T=32.

Using both criteria, when T=64 BVAR1 forecast is preferred for all the collinearity levels except in few cases when ρ =0.95; 0.99; -0.9 and -0.99.

Using both criteria, The values of the criteria increase as a result of increase in the time series length and in the collinearity levels (for both positive and negative collinearity levels). In addition, the values of the criteria for negative collinearity levels are greater than the values of the criteria for positive collinearity levels.

CONCLUSION AND RECOMMENDATION

The various results reported in this paper revealed that the forecasting performances of the models vary as the collinearity level varies. Also the values of the criteria (RMSE and MAE) increases as the time series and the collinearity levels increased. The values of the criteria (RMSE and MAE) for negative collinearity levels are greater than the values of the criteria for positive collinearity levels.

We therefore recommend that if VAR modelers and econometricians know that collinearity is acting upon the model, one can choose the forecasting that is preferred for the criteria and the desired time series length selected.

REFERENCES

1. Schiff AF, Phillips PCB; Forecasting New Zealand's Real GDP. NewZealand Economic Papers, 2000; 34(2):159-182.
2. Robertson JC, Tallman EW; Improving Federal-Funds Rates Forecasts in VAR Models Used for Policy Analysis. Journal of Business & Economic Statistics, 2011;19(3): 324-330.
3. Clements MP, Hendry DF; Forecasting with Difference-Stationary and Trend-Stationary models. The Econometrics Journal, 2001; 4(1):S1-S19.
4. Caraiani P; Forecasting Romanian GDP using A BVAR model. Romanian Journal of Economic Forecasting. 2010; 4:76-87
5. Stock JH, Watson MW; Generalized Shrinkage Methods for Forecasting Using Many Predictors. Journal of Business & Economic Statistics, 2012;30(4):481-493.
6. Hendry DF; The Econometrics of Macroeconomic Forecasting. The Economic Journal, 1997;107(444):1330-1357
7. Clement MP, Hendry DF; Macro-Economic Forecasting and Modelling. The Economic Journal, 1995;105(431):1001-1013.

8. Sims CA; Macroeconomics and Methodology. The Journal of Economic Perspective. 1996; 10(1):105-120.
9. Phillips PCB; Laws and Limits of Econometrics. The Economic Journal, 2003;113(486):C26-C52.
10. Sims CA; Macro economics and Reality. Econometrica, 1980; 48 :1-48
11. Shoesmith GL, Pinder JP; Potential Inventory Cost Reductions using Advanced Time Series Forecasting Techniques. The Journal of the Operational Research Society, 2001; 52(11):1267-1275
12. Pfaff, B; Analysis of Integrated and Cointegrated Time Sevies with R. New York: Springer. 2008. http://www.springer.com/978-0-387-75966-1.
13. Pfaff B; VAR, SVAR and SVEC Models: Implementation within R Package vars. Journal of Statistical Software , 2008; 27(4):1-32.
14. Gujarati DN; Basic Econometrics (4th ed). New Delhi: The McGraw-Hill Co. 2003.
15. Dormann CF; Elith J, Bacher S, Bachmann C, Carl G, Lagourcade B, *et al;* Collinearity: A Review of Methods to Deal with It and A simulation Study Evaluating their Performance. Ecography,2003; 36:027-046.
16. Rama-Sastry MV; Some Limits in the Theory of Multicollinearity. The American Statistician, 1970; 24(1): 39-40.
17. Gunasekara FI, Carter K. Blakely T; Glossary for Econometrics and Epidemiology. Journal of Epidemiology and Community Health, 2008; 62(10):858-861.
18. Blalock HM Jr; Correlated Independent Variables: The Problem of Multicollinearity. Social Forces, 1963;42(2):233-237.
19. Schink WA, Chiu JSY; A Simulation Study of Effects of Multicollinearity and Autocorrelation on Estimates of Parameters. The Journal of Financial and Quantitative Analysis, 1996; 1(2): 36-67
20. Harvey AC; Some Comments on Multicollinearity in Regression. Appl. Statist, 1977; 26(2):188-191.
21. Atkinson SE; Small-Sample Properties of Simultaneous Equation Estimators with Multicollinearity. Journal of the American Statistical Association, 1978; 73(364):719-723.
22. Mittelhammer RC, Young DL, Tasanasanta D, Donnelly JT; Mitigating the Effects of Multicollinearity using Exact and Stochastic Restrictions: The Case of an Aggregated Agricultural Production Function in Thailand. American Journal of Agricultural Economics, 1980; 62(2):199-210.
23. Sharma S, James WL; Latent Root Regression: An Alternate Procedure for Estimating Parameters in the Presence of Multicollinearity. Journal of Marketing Research, 1981; 18(2):154-161.
24. Jagpal HS; Multicollinearity in Structural Equation Model with Observable Variables. Journal of Marketing Research, 1982;19(4):431-439.
25. Salinas ST, Hillmer SC; Multicollinearity Problems in Modeling Time Series with Trading-Day Variation. Journal of Business & Economic Statistics, 1987; 5(3):431-436.
26. Burt OR; The Fallacy of Differencing to Reduce Multicollinearity. American Journal of Agricultural Economics, 1987; 69(3):697-700.
27. Mehta JS, Swamy PAVB, Iyengar NS; Finite Sample Properties of a Measure of Multicollinearity. Sankhya: The Indian Journal of Statistics B, 1993; 55(3): 353-371.
28. Buse A; Brickmaking and the Collinear Arts: A Cautionary Tale. The Canadian Journal of Economics , 1994;7(2):408-414.
29. Greenberg E, Parks RP; A Predictive Approach to Model Selection and Multicollinearity. Journal of Applied Econometrics, 1997; 12(1):67-75.
30. Grewal R, Cote JA, Baumgartner H; Multicollinearity and Measurement Error in Structural Equation Models: Implication for Theory Testing. Marketing Science, 2004;23(4):519-529
31. Friedman L, Wall M; Graphical Views of Suppression and Multicollinearity in Multiple linear Regression. The American Statistician, 2006; 59(2):127-136
32. Alabi OO, Ayinde K, Oyejola BA; Empirical Investigation of Effect of Multicollinearity on Type II Error Rates of the Ordinary Least Squares Estimators. JNSA, 2007; 19:50-57.
33. Ayyangar L; Skewness, Multicollinearity, Heteroskedasticity- You Name It, Cost Data Have It! Solutions to Violations of Assumptions of Ordinary Least Regression Models Using SAS. SAS Global Forum, 2007.
34. Johnson TL; Effect of Correlation and Identification Status on Methods of Estimating Parameters of System of Simultaneous Equations Model. Unpublished PhD Thesis, University of Ilorin, Kwara State, Nigeria. 2009.
35. Ijomah MA, Nduka EC; Selection in a Multicollinear Data Using Information Criteria. JNSA, 2010; 22:64-75
36. Agunbiade DA; Effect of Multicollinearity and Sensitivity of the Methods in Simultaneous Equation Model. Journal of Modern Mathematics and Statistics, 2011; 5(1):9- 12
37. Taylor LD; A Note on the Harmful Effects of Multicollinearity. 2012. www.econ.arizona.edu/docs/Working_Papers/2012/Econ-WP-12-14.pdf Accessed on 30-04-2014
38. Olanrewaju SO; Effect of Multicollinearity of some Estimators of System of Regression Equations. Book of Abstract of the 32nd Annual Conference of the Nigerian Mathematical Society. , 2013;79.
39. Yahya WB, Olaniran OR; Bayesian Linear Regression, Ridge Regression and Ordinary Least Squares Methods for modeling Collinear data. Book of Abstract of the 32nd Annual Conference of the Nigerian Mathematical Society. 2013; 74

40. Garba MK, Oyejola BA, Yahya WB; Investigations of Certain Estimators for Modelling Panel Data Under Violations of some Basic Assumptions. Mathematical Theory and Modeling, 2013; 3(10):47-53.
41. Adenomon MO, Oyejola BA; Forecasting Multiple Time Series with Reduced Form Vector Autoregression (VAR) and Sims-Zha Bayesian VAR when the Endogenous Variables are Collinear. In the Book of Abstract of the 33rd Annual Conference of the Nigerian Mathematical Society. May 27-30, 2014. P.46
42. Lűtkepohl H, Breitung J; Impulse Response Analysis of Vector Autoregressive Processes: System Dynamic in Economic and Financial Models. 1997.
43. Hamilton JD; Time Series Analysis. Princeton: Princeton University Press. 1994.
44. Zeller A; An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias. Journal of the American Statistical Association, 1962; 57(298):348-368.
45. Litterman RB; A Statistical Approach to Economic Forecasting. Journal of Business and Economic Statistics, 1986; 4(1):1-4.
46. Ni S, Sun D; Bayesian Estimates for Vector Autoregressive Models. Journal of Business & economic Statistics, 2005; 23 (1):105-117.
47. Ni S, Sun D, Sun D, Sun X; Intrinsic Bayesian Estimation of Vector Autoregression Impulse Responses. Journal of Business & Economic Statistics, 2007; 25(2):163-176.
48. Sims CA, Zha T; Bayesian Methods for Dynamic Multivariate Models. International Economic Review, 1998; 39(4):949-968.
49. Brandt PT, Freeman JR; Advances in Bayesian Time Series Modeling and the Study of Politics: Theory, Testing, Forecasting and Policy Analysis. Political Analysis, 2006; 14(1):1-36.
50. Brandt PT, Freeman JR; Modeling Macro-Political Dynamics. Political Analysis,2009; 17(2):113-142.
51. Sims CA, Zha T; Error Bands for Impulse Responses. Econometrica, 1999; 67(5):113-1155.
52. Brandt PT, Colaresi M, Freeman J R; Dynamic of Reciprocity, Accountability and Credibility. Journal of Conflict Resolution, 2008;52(3): 343-374.
53. Ciccaralli M, Rebucci A; Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary system. Technical report IMF Working paper WP/03/102. 2003.
54. Kadiyala KR, Karlsson S; Numerical Methods for Estimation and Inference in Bayesian VAR Models. Journal of Applied Econometrics, 1997; 12(2):99-132.
55. Leeper EM, Sims CA, Zha T; What Does Monetary Policy Do? Brookings Papers on Economic Activit, 1996; 2:1-63
56. Suppes P; Where Do Bayesian Priors come from? Synthese, 2007; 156(3):441-471.
57. Park T; Forecast Evaluation for Multivariate Time Series Models: The US Cattle Market. Western Journal of Agricultural Economies, 1990; 15(1):133-143.
58. Brandt PT; Markov-Switching, Bayesian Vector Autoregression Models-Package 'MSBVAR'. The R Foundation for Statistical Computing. 2012.
59. Lutkepohl H; New Introduction to Multiple Time Series Analysis. New York: Springer Berlin Heidelberg. 2005.
60. Diebold FX, Mariano RS; Comparing Predictive Accuracy. Journal of Business and Economic Statistics. 2002; 20(1):134-144.
61. Cottet R, Smith M; Bayesian Modelling and Forecasting of Intraday Electricity Load. Journal of the American Statistical Association, 2003;98(464):839-849.
62. Cooray T MJA; Applied Time series Analysis and Forecasting. New Delhi: Narosa Publising House. 2008.
63. Gilbert P; Brief User's Guide: Dynamic Systems Estimation (DSE). 2009.

APPENDIX

Table -I: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=0.8

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	5.944947	5.223351	8.681807	8.174890	20.71093	20.52558
BVAR1	5.955027	5.231829	8.686889	8.178201	20.70335	20.51888
BVAR2	5.950172	5.227514	8.689276	8.184529	20.71127	20.52658
BVAR3	5.942064	5.222570	8.680416	8.171187	20.70845	20.52351
BVAR4	5.932330	5.210486	8.677362	8.171151	20.70668	20.52200

Table II: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=-0.8

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	7.301465	6.841694	9.826192	9.528671	21.80770	21.69558
BVAR1	7.288099	6.833397	9.829202	9.529294	21.80545	21.69381
BVAR2	7.294073	6.835263	9.830657	9.530028	21.80692	21.69678
BVAR3	7.297135	6.840034	9.831626	9.529774	21.80744	21.69593
BVAR4	7.296082	6.837071	9.828199	9.529865	21.80858	21.69711

Table III: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=0.85

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	6.090411	5.452392	8.863075	8.418649	20.9421	20.7834
BVAR1	6.079255	5.446204	8.870502	8.429512	20.93501	20.77609
BVAR2	6.094382	5.459815	8.875071	8.431168	20.94113	20.78305
BVAR3	6.112983	5.474805	8.866864	8.423040	20.93950	20.78101
BVAR4	6.096075	5.459584	8.868954	8.426819	20.94490	20.78611

Table IV: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=-0.85

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	7.607295	7.177085	10.131151	9.859581	22.12035	22.02396
BVAR1	7.607116	7.180642	10.128771	9.862367	22.11824	22.02210
BVAR2	7.600942	7.177709	10.127585	9.860571	22.11917	22.02329
BVAR3	7.609343	7.179344	10.129663	9.862353	22.11898	22.02272
BVAR4	7.607440	7.179565	10.131595	9.861678	22.12005	22.02406

Table V: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=0.9

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	6.306290	5.763989	9.117992	8.743905	21.23163	21.10063
BVAR1	6.292205	5.748694	9.118740	8.742548	21.22970	21.09888
BVAR2	6.274667	5.728456	9.111497	8.735942	21.23842	21.10711
BVAR3	6.288587	5.750875	9.118048	8.745025	21.23875	21.10752
BVAR4	6.292849	5.746301	9.123637	8.747928	21.24211	21.11090

Table VI: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=-0.9

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	7.995479	7.583535	10.50547	10.26386	22.50651	22.42261
BVAR1	7.993814	7.582740	10.50744	10.26464	22.50692	22.42281
BVAR2	7.999466	7.585706	10.50635	10.26297	22.50702	22.42243
BVAR3	8.003289	7.588065	10.50672	10.26232	22.50616	22.42190
BVAR4	7.994211	7.583650	10.50469	10.26290	22.50513	22.42137

Table VII: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=0.95

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	6.642924	6.192261	9.486636	9.179957	21.66685	21.56520
BVAR1	6.648193	6.198923	9.485406	9.178486	21.66044	21.55873
BVAR2	6.627917	6.182536	9.479047	9.176038	21.65747	21.55596
BVAR3	6.622997	6.172850	9.494074	9.189341	21.66780	21.56539
BVAR4	6.639005	6.188634	9.480729	9.173462	21.65124	21.54905

Table VIII: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=-0.95

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	8.539304	8.122451	11.03330	10.79439	23.03094	22.95235
BVAR1	8.538264	8.117379	11.03100	10.79334	23.02792	22.94968
BVAR2	8.537425	8.118822	11.03071	10.79433	23.03002	22.95145
BVAR3	8.531616	8.117703	11.02994	10.79445	23.03099	22.95174
BVAR4	8.543170	8.123485	11.03144	10.79329	23.02975	22.95118

Table IX: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=0.99

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	7.153009	6.774761	10.072495	9.823372	22.29085	22.21318
BVAR1	7.167651	6.790131	10.079401	9.825968	22.28772	22.20960
BVAR2	7.168340	6.787391	10.075360	9.826229	22.28757	22.20978
BVAR3	7.155232	6.779545	10.079003	9.828823	22.28578	22.20785
BVAR4	7.168335	6.789046	10.091377	9.841807	22.30184	22.22388

Table X: FORECASTING PERFORMANCES OF THE MODELS WHEN COLLINEARITY LEVEL IS ρ=-0.99

Models	T=16		T=32		T=64	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
VAR(2)	9.298521	8.828682	11.76372	11.50353	23.75086	23.66195
BVAR1	9.295183	8.828971	11.76888	11.50417	23.75158	23.66192
BVAR2	9.299091	8.830511	11.76842	11.50412	23.75191	23.66226
BVAR3	9.297124	8.828097	11.76809	11.50398	23.75259	23.66267
BVAR4	9.292962	8.826009	11.76556	11.50358	23.75073	23.66175