American Journal of Applied Mathematics and Statistics, 2016, Vol. 4, No. 3, 67-73
Available online at http://pubs.sciepub.com/ajams/4/3/2

DOI:10.12691/ajams-4-3-2

Missing Values Estimation for a Stable Bivariate Autoregressive Process

I.A.Iwok*

Department of Mathematics/Statistics, University of Port-Harcourt, P.M.B.5323, Port-Harcourt, Rivers State; Nigeria
*Corresponding author: ibywok@yahoo.com

Abstract This work proposed a method for the estimation of missing values in a stable bivariate autoregressive time series process. Missing observations were created at different positions in a stable bivariate series and the method was applied. Despite its ease of implementation, the obtained results suggested good performance of the method. The estimates obtained were compared with those of other existing methods. The result showed that the proposed method provides better estimates than the existing methods.

Keywords: *VAR process, stability condition, selection criteria and missing values.*

Cite This Article: I.A.Iwok, "Missing Values Estimation for a Stable Bivariate Autoregressive Process." *American Journal of Applied Mathematics and Statistics*, vol. 4, no. 3 (2016): 67-73. doi: 10.12691/ajams-4-3-2.

1. Introduction

The problem of missing observations in time series data cannot be overemphasised. It is a common problem usually encountered during data collection. An observation may not be available at the time of need due to faulty equipment used in measurement, lot records of events or mistakes, uncooperative response during data collection etc. In practice, statistical analysis is usually carried on a data set with complete observations. In any situation where an observation is missing, it has to be estimated and replaced in the missing position before a conclusion and inference is drawn from the data set.

Estimation of missing values has gained much grounds in other areas of Statistics such as experiment design, multivariate analysis etc. However, less achievement is recorded in the area of time series; perhaps due to its complicated structure.

Time series is an ordered sequence of observations. The ordering is usually through time. In some cases, it might involve other dimensions such as space. A discrete univariate time series variable is denoted by X_t where t belongs to an index set T. A particular time series X_t is believed to have been generated by an underlying mechanism called a stochastic process. Thus, a stochastic process is a family of time indexed random variables $X(\omega, t)$; where ω belong to a sample space.

Just like in other areas of Statistics, an extension of the univariate case is the multivariate time series models. The simplest of its kind is the bivariate case. A bivariate process consists of two component univariate time series. However, the methods of handling such vector are rather complicated and quite different from the bivariate cases of other areas of Statistics. This is because the indexed set T, the correlation and cross correlation structures at different lags are always taken into consideration. Unfortunately, most of these methods were design merely for predictive purposes and depended on the completeness of the sample in space or time. However, the fact that missing observations are inevitable demands the need to incorporate in these techniques, a method that can accommodate such problems.

2. Review of Literature

In statistical analysis, the model estimation problem in the presence of missing data is usually solved by maximum likelihood approach or imputation methods [1]. [1] considered the case of multidimensional time series with a part of observations that are completely or partially missing. He proposed a model estimation procedure based on composite likelihood combined with a model based imputation method. The proposed method was validated by simulation studies and the results enlightened the effect of imputation strategies on model estimation. The method gave better results at least for the variance covariance parameters.

[9] worked on the estimation of parameters on the dynamic factor analysis using the EM algorithm. The result showed that the dynamic factor analysis can analyse short, non stationary time series containing missing value.

[3] considered the problems of predicting missing observations and forecasting future values in incomplete time series data. In their work, three forecasting models (a dynamic multivariate autoregressive model, a multivariate local trend model and a Gaussian process model) were studied. Each of these models was analysed using air temperature data collected by a network of weather sensors. Comparisons of these models were made and it was discovered that the dynamic linear model coped easily with incomplete or missing observations.

[2] presented a unified approach to the analysis of messy data. The paper examined irregularities such as missing values, outliers, structural breaks and irregular spacing. Here, a missing observation was handled by

introducing a dummy variable into the measurement. By introducing a state space frame work to explanatory variables, [2] discovered that the method of dummy variable has exactly the same effect as skipping a filter update introduced by Kalman filters.

[8] proposed a method for modelling and fitting multivariate spartial time series data based on current spartial methodology coupled with the parameterization of the ARIMAX model. Because of the physical constraints imposed on multivariate data collection in both space and time, the estimation and identification procedures tolerated general patterns of missing or incomplete data.

[5] considered the problem of estimating parametric multivariate density models when unequal amount of data are available on each variable. It was discovered that there exist a significant evidence of time variation in the conditional copula of the exchange rates (Yen-US dollars and Euro-US dollars), and evidence of greater dependence during extreme events than under the normal distribution.

[7] proposed an imputation method to be used with singular spectrum techniques which is based on a weighted combination of the forecasts and hind casts yield by the recurrent forecast method. Despite its ease of implementation, the obtained result suggested an overall good fit of their method. This is because it yielded similar adjustment ability in comparison with the alternative method according to some measures of predictive performance.

According to [6], the commonest way of dealing with missing observations is to replace them with the mean of the data. This is because every observation is expected to be distributed around the mean under normal situation. According to them, any observation that deviate much from the mean has to be tortured to reflect it membership before being used for analysis.

Actually, there is no much direct literature on missing values of a stable vector autoregressive (VAR) process. Some of the literatures cited above are either indirectly or lightly linked with the subject matter. Hence, our comparative study shall base on few related existing methods that are computationally less cumbersome.

3. Methodology

Let y_{1t} and y_{2t} be two univariate time series under consideration. Then, $y_t = (y_{1t}, y_{2t})'$ is said to be a bivariate time series. According to [4], the general vector autoregressive model of order p [$VAR(p)$] for y_t can be expressed as:

$$y_t = c + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \ldots + \Pi_p y_{t-p} + \varepsilon_t; \quad t = 0, \pm 1, \pm 2, \ldots \tag{1}$$

where

$y_t = (y_{1t}, y_{2t})'$ is a (2×1) vector of time series variables,

Π_i are fixed (2×2) coefficient matrices

$c = (c_1, c_2)'$ is a fixed (2×1) vector of intercept terms allowing for the possibility of non zero mean $E(y_t)$.

$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a (2×1) white noise process or innovation process. That is,

$$E(\varepsilon_t) = 0,\ E\left(\varepsilon_t \varepsilon_t'\right) = \Sigma_\varepsilon \ and\ E\left(\varepsilon_t \varepsilon_s'\right) = 0 \ for\ s \neq t$$

Σ_ε = covariance matrix which is assume to be non singular if not otherwise stated.

The model can be written in the matrix form as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^1 & \pi_{12}^1 \\ \pi_{21}^1 & \pi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \pi_{11}^2 & \pi_{12}^2 \\ \pi_{21}^2 & \pi_{22}^2 \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \ldots + \begin{pmatrix} \pi_{11}^p & \pi_{12}^p \\ \pi_{21}^p & \pi_{22}^p \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

3.1. VAR Order Selection

For any n-dimensional time series vector $y_t = (y_{1t}, \ldots, y_{nt})'$ assumed to be generated by a $VAR(p)$ process (where p is the order of the model), three model order selection criteria are usually considered:

(i) ***Akaike Information Criteria (AIC)***

This is given by

$$IC(p) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{2}{N}(number\ of\ estimated\ parameter) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{2pn^2}{N}.$$

The estimate $(\widehat{AIC})$ for p is chosen so that this criterion is minimized. Here the constant in the VAR model may be ignored as freely estimated parameter because counting them would just add a constant to the criterion which does not change the minimizing order.

(ii) ***Hannan-Quin Criterion (HQC)***

This is given as

$$HQ(p) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{2InIn}{N}(freely\ estimated\ parameters) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{2InInN}{N}pn^2.$$

The estimate $(\widehat{HQ})$ is the order that minimizes $HQ\ (p)$ for $p = 0,1, \ldots, P$

(iii) ***Schwarz Criterion (BIC)***

This is given by

$$SC(p) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{InN}{N}(freely\ estimated\ parameters) = In\left|\tilde{\Sigma}_\varepsilon(p)\right| + \frac{2InN}{N}pn^2.$$

The estimate $(\widehat{SC})$ is chosen so as to minimize the value of the criterion.

Where p is the VAR order

$\tilde{\Sigma}_\varepsilon$ is the estimate of white noise covariance matrix Σ_ε

n is the number of time series components of the vector time series

N is the sample size.

3.2. Stable VAR (p) Processes

Any VAR (p) processes with $p > 1$ can be written in VAR (1) form [4]. More precisely, if y_t is VAR (p), a corresponding np-dimensional VAR (1) is given by

$$Y_t = \boldsymbol{c} + \boldsymbol{\Pi} Y_{t-1} + \epsilon_t.$$

where

$$Y_t = \begin{bmatrix} y_t \\ y_{t-p} \\ \vdots \\ y_{t-p+1} \end{bmatrix}; \boldsymbol{c} = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix};$$

$$\boldsymbol{\Pi} = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_{p-1} & \pi_p \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & & I_n & 0 \end{bmatrix}; \epsilon_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The process $\{Y_t\}$ is said to be stable if

$$\det\left(I_{np} - \boldsymbol{\Pi}_z\right) \neq 0 \ \ for \ |z| \leq 1.$$

Specifically, the process (1) is stable if its reverse characteristic polynomial of the VAR (p) has no roots in and on the complex unit circle. That is, y_t is stable if

$$\det\left(I_k - \Pi_1 z - \ldots - \Pi_p z^p\right) \neq 0 \ \ for \ |z| \leq 1.$$

This condition is called the stability condition. A stable VAR (p) process $y_t, t = 0, \pm1, \pm2, \ldots$, is also stationary.

4. The Missing Value Approach

Now, suppose we have N observations for both series; and the missing observations occur at the ith and jth position of y_{1t} and y_{2t} respectively. We estimate these missing observations by the following approach:

Let y_{1i} and y_{2j} be the two missing observations. If $i \geq j$, the first step involves obtaining VAR lag order of the two series $\{y_{1,i+1}, y_{1,i+2}, \cdots, y_{1,N}\}$ and $\{y_{2,i+1}, y_{2,i+2}, \cdots, y_{2,N}\}$. However, if $j \geq i$, we obtain the VAR lag order of the series $\{y_{1,j+1}, y_{1,j+2}, \ldots, y_{1,N}\}$ and $\{y_{2,j+1}, y_{2,j+2}, \cdots, y_{2,N}\}$. Of course this particular step also involves fitting the $VAR(p)$ model to the above series and obtaining the estimate of the parameters as shown:

$$\begin{pmatrix} \widehat{y_{1t}^{(1)}} \\ \widehat{y_{2t}^{(1)}} \end{pmatrix} = \begin{pmatrix} \widehat{c_1} \\ \widehat{c_1} \end{pmatrix}^{(1)} + \begin{pmatrix} \widehat{\pi_{11}^1} & \widehat{\pi_{12}^1} \\ \widehat{\pi_{21}^1} & \widehat{\pi_{22}^1} \end{pmatrix}^{(1)} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} \widehat{\pi_{11}^2} & \widehat{\pi_{12}^2} \\ \widehat{\pi_{21}^2} & \widehat{\pi_{22}^2} \end{pmatrix}^{(1)} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \ldots + \begin{pmatrix} \widehat{\pi_{11}^p} & \widehat{\pi_{12}^p} \\ \widehat{\pi_{21}^p} & \widehat{\pi_{22}^p} \end{pmatrix}^{(1)} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix}$$

The above fitted model is then used to obtain the first estimate of the missing observations $\widehat{y_{1t}^{(1)}}$ and $\widehat{y_{2t}^{(1)}}$.

The second step involves substituting these estimates $\widehat{y_{1t}^{(1)}}$ and $\widehat{y_{2t}^{(1)}}$ in their missing positions in the data and the bivariate analysis is repeated on the complete data to obtain the final model as shown:

$$\begin{pmatrix} \widehat{y_{1t}^{(2)}} \\ \widehat{y_{2t}^{(2)}} \end{pmatrix} = \begin{pmatrix} \widehat{c_1} \\ \widehat{c_1} \end{pmatrix}^{(2)} + \begin{pmatrix} \widehat{\pi_{11}^1} & \widehat{\pi_{12}^1} \\ \widehat{\pi_{21}^1} & \widehat{\pi_{22}^1} \end{pmatrix}^{(2)} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \widehat{\pi_{11}^2} & \widehat{\pi_{12}^2} \\ \widehat{\pi_{21}^2} & \widehat{\pi_{22}^2} \end{pmatrix}^{(2)} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \ldots + \begin{pmatrix} \widehat{\pi_{11}^p} & \widehat{\pi_{12}^p} \\ \widehat{\pi_{21}^p} & \widehat{\pi_{22}^p} \end{pmatrix}^{(2)} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix}.$$

This model is then used to obtain the final estimates of the missing observations $\widehat{y_{1t}^{(2)}}$ and $\widehat{y_{2t}^{(2)}}$.

The superscripts in braces in the above expressions only indicate that we are obtaining the first (1) or second (2) estimate of the missing observations.

5. Illustration and Result

We illustrate the above proposed approach of estimation using 50 monthly cases of hypertension (y_{1t}) and diabetes (y_{2t}) data (see appendix 5). First, we removed the $y_{1,3} = 34$ and $y_{2,4} = 22$ from the data. Next, we conducted VAR analysis on the series $\{y_{1,5}, y_{1,6}, \ldots, y_{1,50}\}$ and $\{y_{2,5}, y_{2,6}, \ldots, y_{2,50}\}$ using gretl software. The selected VAR order (in this case $p = 1$ as indicated by minimum values of the three criteria at lag 1) and the result of the analysis at this first step are displayed on appendix 1 and 2. Thus, the resulting model is:

$$\begin{pmatrix} \widehat{y_{1i}^{(1)}} \\ \widehat{y_{2j}^{(1)}} \end{pmatrix} = \begin{pmatrix} 4.1374 \\ 12.0521 \end{pmatrix}^{(1)} + \begin{pmatrix} 0.85232 & -0.0035431 \\ 0.14153 & 0.41323 \end{pmatrix}^{(1)} \begin{pmatrix} y_{1,i-1} \\ y_{2,i-1} \end{pmatrix} \quad (2)$$

The above model (2) is then used for the estimation of the first estimates of the missing values which gives: $\widehat{y_{1,3}^{(1)}} = 30.4779$ and $\widehat{y_{2,4}^{(1)}} = 25.9438$.

In step two, these estimate ($\widehat{y_{1,3}^{(1)}} = 30.4779$ and $\widehat{y_{2,4}^{(1)}} = 25.9438$) are replaced in their missing positions and the analysis is repeated for the entire series. That is, for the series $\{y_{1,1}, y_{1,2}, \widehat{y_{1,3}^{(1)}}, \ldots, y_{1,50}\}$ and $\{y_{2,1}, y_{2,2}, y_{2,3}, \widehat{y_{2,4}^{(1)}}, \ldots, y_{2,50}\}$. The results of the analysis are displayed in appendix 3 and 4 below; and we have the final model:

$$\begin{pmatrix} \widehat{y_{1i}^{(2)}} \\ \widehat{y_{2j}^{(2)}} \end{pmatrix} = \begin{pmatrix} 4.0884 \\ 11.0435 \end{pmatrix}^{(2)} + \begin{pmatrix} 0.940812 & -0.0026241 \\ 0.15747 & 0.30215 \end{pmatrix}^{(2)} \begin{pmatrix} y_{1,i-1} \\ y_{2,j-1} \end{pmatrix} \quad (3)$$

The above model (3) is then used to compute the final estimates of the missing observations. Thus we have as our final estimates:

$$\widehat{y_{1,3}^{(2)}} = 33.1932 \; and \; \widehat{y_{2,4}^{(2)}} = 22.8745.$$

The absolute deviations (denoted AD) of these estimates from their respective actual values in the data can be computed as errors of the estimates. Thus, we have

$$\left| y_{1,3} - \widehat{y_{1,3}^{(2)}} \right| = 0.8068 \; and \; \left| y_{2,4} - \widehat{y_{2,4}^{(2)}} \right| = 0.8745.$$

We see here that the errors conceived by these estimate are negligible. We all know that the primary practice of dealing with missing observations is to replace them with the mean. Comparatively, however, the error created by our method of estimation is far less than that created by using the means ($\overline{y_{1t}} = 41.760$ and $\overline{y_{1t}} = 24.840$).

5.1. Stability of the VAR (1) Process

For this process, the reverse characteristic polynomial is

$$\det\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.940812 & -0.0026241 \\ 0.15747 & 0.30215 \end{pmatrix} z\right]$$

$$= \det\begin{bmatrix} 1-0.94081z & 0.0026241z \\ -0.15747z & 1-0.30215z \end{bmatrix}$$

which gives the roots: $z_1 = 3.302541336$ and $z_2 = 2.063643218$. These roots are outside the unit circle; and the process is therefore stable.

The table (Table 1) below shows the estimates obtained by creating additional four missing observations at various (i th and j th) positions in the y_{1t} and y_{2t} series. Comparisons of these estimates are made between our proposed new method (PNM) and other methods: the mean (M) and the Stoffer method (SM); based on the absolute deviations (AD). The absolute deviations from the actual values are placed in bracket under their respective estimates.

Table 1. Estimates and Errors of the Different Methods

Position (i ; j)	Actual Values ($y_{1,i}$; $y_{2,j}$)	Mean ($\overline{y_{1t}}, \overline{y_{2t}}$)	SM ($\widehat{y_{1,i}}, \widehat{y_{2,j}}$)	PNM ($\widehat{y_{1,i}}, \widehat{y_{2,j}}$)
(3 ; 4)	(34 ; 22)	(41.76 ; 24.84) (7.76 ; 2.84)	(43.31 ; 26.22) (9.31 ; 4.22)	(33.19 ; 22.87) (0.806 ; 0.875)
(5 ; 5)	(36 ; 18)	(41.76 ; 24.84) (5.76 ; 6.84)	(38.12 ; 23.45) (2.12 ; 5.45)	(36.34 ; 19.02) (0.340 ; 1.020)
(2 ; 7)	(31 ; 23)	(41.76 ; 24.84) (10.76 ; 1.84)	(36.41 ; 20.18) (5.41 ; 2.82)	(30.16 ; 23.91) (0.840 ; 0.910
(6 ; 6)	(38 ; 20)	(41.76 ; 24.84) (3.76 ; 4.84)	(36.31 ; 23.45) (1.69 ; 3.45)	(37.92 ; 20.77) (0.080 ; 0.770)
(4 ; 9)	(29 ; 24)	(41.76 ; 24.84) (12.76 ; 0.84)	(31.52 ; 26.23) (2.52 ; 2.23)	(28.73 ; 24.86) (0.270 ; 0.860)

As seen in the above table, comparison based on AD (errors) shows that the proposed new method provides better estimates than the mean and the Stoffer method.

6. Summary and Conclusion

This work has provided a method of estimating missing values in a stable VAR bivariate process. The approach was illustrated using real life data. The method was found to outperform other methods of estimation with minimum error. In this light, this proposed method has offered a practical framework of dealing with missing observations in a stable VAR process.

References

[1] Giovanni, F. and Grassetti, (2011). Multidimensional time series model estimation in the presence of partially missing observations. Workshop SIG-annual conference paper, Vol.3, No. 3, pp. 28 -35.

[2] Harvey A. and Koopman S.J. (2015). Messy time series ''A unified approach''. Advances in Econometrics; Vol. 13, pp 103-143.

[3] Lee S.M. and Robert S. (2008). Multivariate time series forecasting in incomplete environments. Technical report PARG 08-03. Robostics Research group. Department of Engineering Science, University of Oxford.

[4] Lutkepol H. (2005): New introduction to multiple Time Series Analysis. Springer Berlin Heidebelg New York.

[5] Patton, A.J. (2006). Estimation of multivariate models for time series of possibly different length. Journal of applied Econometrics: 21: 147-173.

[6] Peterson, H. and Pederson, B. (2010). Handling missing values and Outliers in a purely random data. Journal of Mathematical Scieces. Vol.3, No. 7, pp. 37-41.

[7] Rodrignes, P.C. and Carvalho, M. (2014). Spectral Modelling of time series with missing data. Centre for Mathematics and applications. Faculty of Science and Technology. Nova University of Lisbon; 2829-516 Caparica, Portugal.

[8] Stoffer, D.S. (1986). Estimation and Identification of space-time ARMAX models in the presense of missing data. Journal of the American Statistical Association. Vol. 81, No. 395, theory and methods.

[9] Zuur, A.F. *et al* (2015). Estimating common trends in multivariate time series using dynamic factor analysis. J. ana. In Stat. Vol.2, No. 5, pp. 19-24.

Appendix 1. VAR Order Selection with Missing Observations

VAR system, maximum lag order 12

The asterisks below indicate the best (that is, minimized) values of the respective information criteria, AIC = Akaike criterion, BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.

Lags	loglik	p(LR)	AIC	BIC	HQC
1	-195.84598		10.714321*	10.912728*	10.808134*
2	-193.17314	0.21243	10.788452	11.215428	10.848623
3	-191.38435	0.03712	10.768254	11.313462	10.963421
4	-182.27234	0.71023	10.871368	11.642357	11.194232
5	-187.71407	0.51241	11.024157	11.972213	11.365237
6	-185.22362	0.05921	11.016237	12.125132	11.503314
7	-180.43623	0.32145	11.121322	12.334152	11.565781
8	-177.93611	0.65342	11.224136	12.701238	11.752562
9	-171.33414	0.00923	11.114321	12.685643	11.701146
10	-170.20215	0.41242	11.223252	13.014356	11.853217
11	-163.08641	0.02231	11.106753	13.089534	11.883245
12	-161.43124	0.41205	11.202435	13.324512	11.973847

Appendix 2. Analysis Result with Missing Observations

VAR system, lag order 1
OLS estimates, observations 2009:07-2013:02 (T = 43)
Log-likelihood = -234.13517
Determinant of covariance matrix = 117.85316
AIC = 10.7143
BIC = 10.9127
HQC = 10.8081
Portmanteau test: LB(12) = 56.9308, df = 42 [0.0914]

Equation 1: y1t

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	4.1374	3.16324	1.3080	0.04016	**
y1t_1	0.85232	0.0533421	15.9784	<0.00001	***
y2t_1	-0.0035431	0.212332	-0.01669	0.03222	**

Mean dependent var	42.00000	S.D. dependent var	8.044149
Sum squared resid	567.4326	S.E. of regression	4.306623
R-squared	0.725267	Adjusted R-squared	0.720145
F(2, 44)	4326.355	P-value(F)	5.20e-54
rho	-0.352549	Durbin-Watson	2.703658

F-tests of zero restrictions:
All lags of y1t F(1, 44) = 298.74 [0.0000]
All lags of y2t F(1, 44) = 0.50299 [0.4817]

Equation 2: y2t

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	12.0521	4.37116	2.7572	0.00332	***
y1t_1	0.14153	0.0626574	2.2588	0.03112	**
y2t_1	0.41323	0.2245	1.8407	0.03522	**

Mean dependent var	27.08192	S.D. dependent var	4.237256
Sum squared resid	707.2115	S.E. of regression	3.925227
R-squared	0.264962	Adjusted R-squared	0.233004
F(2, 44)	978.4958	P-value(F)	5.01e-39
rho	-0.050675	Durbin-Watson	2.071617

F-tests of zero restrictions:
All lags of y1t F(1, 44) = 15.008 [0.0003]
All lags of y2t F(1, 44) = 22.252 [0.0000]

Appendix 3. VAR Order Selection with Replaced Estimates

VAR system, maximum lag order 12

The asterisks below indicate the best (that is, minimized) values of the respective information criteria, AIC = Akaike criterion, BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.

lags	loglik	p(LR)	AIC	BIC	HQC
1	-197.95396		10.734419*	10.992985*	10.826415*
2	-195.17419	0.23454	10.798642	11.229585	10.951968
3	-190.39440	0.04854	10.757600	11.360921	10.972257
4	-189.39088	0.73446	10.915309	11.691008	11.191297
5	-187.71407	0.50048	11.037582	11.985659	11.374901
6	-183.32577	0.06693	11.017146	12.137599	11.415795
7	-180.96331	0.31670	11.103332	12.396163	11.563312
8	-179.82941	0.68664	11.254179	12.719388	11.775489
9	-173.14203	0.00958	11.112738	12.750325	11.695379
10	-171.14399	0.40654	11.218105	13.028068	11.862076
11	-165.04677	0.01596	11.107725	13.090066	11.813027
12	-163.09897	0.42032	11.215735	13.370454	11.982368

Appendix 4. Analysis Result with Replaced Estimates

VAR system, lag order 1
OLS estimates, observations 2009:02-2013:02 (T = 49)
Log-likelihood = -256.28509
Determinant of covariance matrix = 119.68483
AIC = 10.7344
BIC = 10.9930
HQC = 10.8264
Portmanteau test: LB(12) = 55.5103, df = 44 [0.1144]

Equation 1: y1t

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	4.0884	4.07308	1.0038	0.03326	**
y1t_1	0.940812	0.0643001	14.6316	<0.00001	***
y2t_1	-0.0026241	0.117124	-0.0224	0.02232	**

Mean dependent var	42.00000	S.D. dependent var	8.044149
Sum squared resid	473.3555	S.E. of regression	3.207855
R-squared	0.847600	Adjusted R-squared	0.840974
F(2, 46)	127.9183	P-value(F)	1.62e-19
rho	-0.328626	Durbin-Watson	2.656506

F-tests of zero restrictions:

All lags of y1t F(1, 46) = 214.08 [0.0000]
All lags of y2t F(1, 46) = 0.00050196 [0.9822]

Equation 2: y2t

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	11.0435	3.48019	3.1732	0.00269	***
y1t_1	0.15747	0.0728185	2.1625	0.03581	**
y2t_1	0.30215	0.13264	2.2780	0.02742	**

Mean dependent var	25.04082	S.D. dependent var	4.148088
Sum squared resid	607.0815	S.E. of regression	3.632826
R-squared	0.976547	Adjusted R-squared	0.976048
F(2, 46)	8.290891	P-value(F)	0.000842
rho	0.020217	Durbin-Watson	1.950617

F-tests of zero restrictions:

All lags of y1t F(1, 46) = 4.6764 [0.0358]
All lags of y2t F(1, 46) = 5.1892 [0.0274]

Appendix 5. [Monthly Cases of Hypertension and Diabetes (2009 – 2013)]

S/N	y_{1t}	y_{2t}	S/N	y_{1t}	y_{2t}	S/N	y_{1t}	y_{2t}	S/N	y_{1t}	y_{2t}	S/N	y_{1t}	y_{2t}
1.	30	15	11.	35	29	21.	38	19	31.	46	23	41.	49	25
2.	31	21	12.	36	25	22.	39	25	32.	47	27	42.	50	29
3.	34	23	13.	28	24	23.	41	26	33.	43	29	43.	52	30
4.	29	22	14.	31	27	24.	43	27	34.	42	13	44.	53	31
5.	36	18	15.	33	28	25.	46	28	35.	45	30	45.	51	32
6.	38	20	16.	34	22	26.	39	30	36.	44	28	46.	53	29
7.	35	23	17.	35	19	27.	48	32	37.	46	21	47.	55	23
8.	39	21	18.	32	20	28.	46	21	38.	47	22	48.	56	24
9.	37	24	19.	33	21	29.	43	28	39.	46	23	49.	58	25
10.	31	28	20.	35	23	30.	44	25	40.	48	24	50.	58	30

Source: Liby Hospital, Nigeria.