# Evidence of Variance Covariance Matrix Sensitivity to Correlated Normal Deviates

Olusanya E. Olubusoye
Department of Statistics
&
Fellow, Centre for Econometric and Allied Research (CEAR)
University of Ibadan
Ibadan, Oyo-State
Nigeria
234-08056142797
oe.olubusoye@mail.ui.edu.ng

## ABSTRACT

The design of Monte-Carlo experiment to compare alternative estimators of simultaneous equation model requires an absence of correlation between the pairs of normal deviates generated. This is a very difficult requirement except several pairs of normal deviates are generated and then pairs with very low correlation coefficients are drawn. This study designed a 2-equation simultaneous system with a specified covariance matrix and examined the correlation between the pairs of normal deviates generated and the implied covariance matrix.

The result showed that the random disturbances generated from normal deviates with feeble correlation coefficient (negative or positive) reproduced the specified variance-covariance matrix while others failed. The result suggested the need to screen the normal deviates generated for the purpose of Monte-Carlo studies especially in the context of simultaneous equation systems.


**KEY WORDS:** Monte Carlo experiments, Simultaneous Equation System, Correlation Coefficient, Normal Deviates, Disturbances.

## I. **Introduction**

Monte Carlo methods have in the last two decades found extensive use in many fields such as operational research, nuclear physics and econometrics to mention but few, where there are a variety and complexity of problems beyond the available resources of the theoretician. In Econometrics for instance, Wagner, H.M. (1958), Nagar, A.L. (1960), G. Rudebusch (1993), Matyas and Lovrics (1991) e.t.c. are few examples of studies based on Monte Carlo experimentation. By the Monte Carlo approach the econometrician generates data sets and stochastic terms which are free of the problems of multicollinearity, nonspherical disturbances, measurement error and even specification error. In the context of simultaneous equation system, the design of Monte Carlo experiments requires the generation of orthogonal normal deviates or mutually independent sequences distributed as N(0,1). These normal deviates are then transformed to ensure that the disturbance terms are distributed as $N(0,\Sigma)$ which are not serially correlated, where $\Sigma$ is the assumed variance-covariance matrix of the disturbances. This study examined the extent to which the pairs of simulated non-orthogonal or correlated normal deviates affect variance covariance matrix of the random disturbances generated from them..

The rest of the paper is divided into four sections. In section II we discuss the general framework, in section III we discuss the generation of sample values, in section IV we present and discuss the results before rounding off the paper in a concluding section V.

## II. The General Framework of the Study

We assume the following two structural equations:

$$\begin{aligned} Y_{1t} &= b_{12}Y_{2t} + g_{11}X_{1t} + u_{1t} \\ Y_{2t} &= b_{21}Y_{1t} + g_{12}X_{2t} + u_{2t} \end{aligned} \qquad (1)$$

where Y's are the endogenous variables and X's are the predetermined variables.

For an arbitrary period t, let $u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$

$$E(u_t) = 0 \qquad (2)$$

That is,

$$E(u_t) = E\left(\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (3)$$

The Variance-Covariance matrix is defined as:

$$E(u_t u_t') = \Sigma = \Omega \otimes I_T \qquad (4)$$

$$E(u_t u_t') = E\left\{\begin{bmatrix} u_{11} \\ u_{12} \\ . \\ . \\ . \\ u_{1T} \\ u_{21} \\ u_{22} \\ . \\ . \\ . \\ u_{2T} \end{bmatrix} [u_{11} u_{12} \ldots u_{1T} u_{21} u_{22} \ldots u_{2T}]\right\} \quad (5)$$

Estimation of (5) gives a 2T x 2T symmetric matrix whose elements are given by:

$$E(u_{it} u_{i^* t^*}) = \begin{cases} s_i^2, \; for \; i = i^*, \; \text{t=t}^* & \text{(A)} \\ 0, \; for \; i = i^*, \; t \neq t^* & \text{(B)} \\ s_{ii^*}, \; for \; i \neq i^*, \; t = t^* & \text{(C)} \\ 0, \; for \; i \neq i^*, \; t \neq t^* & \text{(D)} \end{cases} \quad (6)$$

where,

(A) implies existence of homoscedasticity

(B) implies absence of autocorrelation

(C) implies that covariances of the same periods are equal to $s_{ii^*}$. This means that contemporaneous covariances in different equations are independent of time.

(D) implies that covariances at different times in different equations or non-contemporaneous covariances are zero.

Now if we consider two mutually independent N(0,1) sequences, say $\{(x_{1t},\ x_{2t}): t = 1,2,...,T\}$ which are generated and are transformed as follows to ensure that the disturbance terms $u_t$ are distributed as $N(0,\Sigma)$ and intertemporally independent.

$$\text{Let } x_t = (x_{1t},\ x_{2t})' \tag{7}$$

then by construction:

$$E(x_t) = 0, \quad Cov(x_{1t},\ x_{2t}) = d_{tt'}\ I \tag{8}$$

Since $\Sigma$ by definition is a positive definite matrix, there exists a non singular matrix P such that

$$\Sigma = PP' \tag{9}$$

Consider now the random vectors:

$$U_t = Px_t \tag{10}$$

By construction, the vectors

$\{U_t : t = 1,2,...,T\}$ have the properties

$$E(U_t) = 0$$

$$\Sigma = Cov(U_t, U_t') = \Omega \otimes I_t \tag{12}$$

where $\otimes$ is the kronecker product or direct product. In this fashion, the desired error terms having the prescribed variance-covariance matrix is obtained.

## III. <u>Generation of Disturbance Terms</u>

The requirement of generating normal deviates $\{(x_{1t},\ x_{2t}): t = 1,2,...,T\}$ such that $Cov(x_{1t},\ x_{2t}) = 0$ is indeed difficult. Consequently, the first step here is to compute the correlation coefficient between the several pairs of normal deviates generated. The correlation coefficients are not necessarily tested for significance. This is because the focus is to determine whether random disturbances generated at a given level of correlation coefficient can still reproduce the specified variance-covariance matrix.

To set up our experiment, we proceed as follows:

(i) The sample size T is specified as T = 15, 25, 40.

(ii) The covariance matrix of the disturbances is specified arbitrarily as follows:

$$\Omega = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 5.0 & 2.5 \\ 2.5 & 3.0 \end{bmatrix} \qquad (13)$$

(iii) A two-stage process was used to generate the values of the random disturbances. The first stage involves drawing independent series of normal deviates of the required length. In the second stage, the series are then transformed into series for the random disturbances in such a way as to guarantee that they have the covariances which are the same as those set out in (13). There are two random disturbances (one for each equation) in our model, hence, we generate two independent series denoted as $x_{1t}, x_{2t}$. Thereafter, each of the series $x_{1t}$ and $x_{2t}$ is standardized to have mean zero and unit variance. To compute random disturbance terms that behave as described in (ii) above, we used the method presented in Nagar (1969) for transforming M independent series of normal deviates, of length T into M series of random variables with zero means and a specified covariance matrix.

From equation (12) and $\Sigma$ being a positive definite matrix, we can decompose it by a non-singular upper triangular matrix such that $\Omega = P_1 P_1^{/}$

Let $P_1 = \begin{pmatrix} P_{11} & P_{12} \\ p_{21} & P_{22} \end{pmatrix}$

Then,

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} P_{11} & P_{21} \\ P_{12} & P_{22} \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$$

Hence, $P_{11}^2 + P_{12}^2 = s_{11}$ (12a)

$P_{11}P_{21} + P_{12}P_{22} = s_{12}$ (12b)

$P_{21}P_{11} + P_{22}P_{12} = s_{21}$ (12c)

$$P_{21}^2 + P_{22}^2 = s_{22} \qquad (12d)$$

Since $s_{12}$ and $s_{21}$ equation (12b) and (12c) are identical. Since $P_{21}$ is not necessarily equal to $P_{12}$ we have only three equations (12a), (12b) or (12c) and (12d) and four unknowns $P_{11}$, $P_{12}$, $P_{21}$ and $P_{22}$. To be able to solve for the $P_{ij}$'s uniquely we have to assure that one of the variable is known. Here we assume that $P_{21} = 0$ and solve for $P_{11}$, $P_{12}$ and $P_{22}$ which is the solution based on the upper triangular matrix $P_1$ or we assume that $P_{12} = 0$ and solve for $P_{11}$, $P_{21}$ and $P_{22}$ which is the solution based on the lower triangular matrix $P_2$. Hence the solution of the $P_{ij}$'s is obtained in term of $s_{ij}$'s from either:

$$\begin{pmatrix} P_{11}^* & P_{12}^* \\ 0 & P_{22}^* \end{pmatrix} \begin{pmatrix} P_{11}^* & 0 \\ P_{12}^* & P_{22}^* \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} = \begin{pmatrix} 5.0 & 2.5 \\ 2.5 & 3.0 \end{pmatrix}$$

which is the solution based on the upper triangular matrix $P_1$ where $P_{21}^* = 0$

or

$$\begin{pmatrix} P_{11}^{**} & 0 \\ P_{12}^{**} & P_{22}^{**} \end{pmatrix} \begin{pmatrix} P_{11}^{**} & P_{12}^{**} \\ 0 & P_{22}^{**} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} = \begin{pmatrix} 5.0 & 2.5 \\ 2.5 & 3.0 \end{pmatrix}$$

which is the solution based on the lower triangular matrix $P_2$ where $P_{12}^{**} = 0$.

The two random disturbance series are therefore constructed using the relation (10) and the values of $s_{11}$, $s_{21}$ and $s_{22}$ derived from the variance-covariance matrix $\Omega$ assumed in (13) above. Using the solution based on $P_1$;

$$\left.\begin{aligned} U_{1t} &= P_{11}^{*}e_{1t} + P_{12}^{*}e_{2t} \\ U_{2t} &= P_{22}^{*}e_{2t} \end{aligned}\right\} (t = 1,2,\ldots,T) \tag{14}$$

Solution based on $P_2$ yields;

$$\left.\begin{aligned} U_{1t} &= P_{11}^{**}e_{1t} \\ U_{2t} &= P_{21}^{**}e_{1t} + P_{22}^{**}e_{2t} \end{aligned}\right\} (t = 1,2,\ldots,T) \tag{15}$$

In the solution based on $P_1$, it can be readily shown that:

$$P_{22}^{*} = +\sqrt{s_{22}} \tag{14a}$$

$$P_{12}^{*} = \frac{s_{12}}{P_{22}^{*}} \tag{14b}$$

$$P_{11}^{*} = +\sqrt{s_{11} - P_{12}^{*2}} \tag{14c}$$

Alternatively using the solution based on $P_2$, it can also be shown that:

$$P_{11}^{**} = +\sqrt{s_{11}} \tag{15a}$$

$$P_{12}^{**} = \frac{s_{12}}{P_{11}^{**}} \tag{15b}$$

$$P_{22}^{**} = +\sqrt{s_{22} - P_{12}^{**2}} \tag{15c}$$

## IV. **Results**

The variances and covariances of the disturbance series generated from equations (14) and (15) when T= 15 and replication is 40 are presented in the table below. Similar results where observed for other sample sizes and replications, hence they are not presented here. The first column in the table gives the normal deviates correlation coefficient. The second column gives the random disturbance correlation. The third column gives the covariance between the disturbances. The value in bracket represents the specified value in the variance covariance matrix. Columns five and six give the variances for each of the disturbances when a non-singular upper triangular matrix is used to decompose the variance covariance matrix. Column seven and eight give the variances for each of the disturbances when non-singular lower triangular matrix is used to decompose the variance covariance matrix.

**Table 1**
**VARIANCES AND COVARIANCES OF PAIRS OF RANDOM DISTURBANCES FROM NORMAL DEVIATES AT VARIOUS LEVELS OF CORRELATION**

| Replication | Normal Deviates Correlation | Random Disturbance Correlation | Cov ($U_1$,$U_2$) (**2.5**) | $P_1$ | | $P_2$ | |
|---|---|---|---|---|---|---|---|
| | | | | Var($U_1$) (**5.0**) | Var($U_2$) (**3.0**) | Var($U_1$) (**5.0**) | Var($U_2$) (**3.0**) |
| 1 | -0.558 | 0.3267 | 0.1661 | 2.2471 | 3.0 | 5.0 | 1.3482 |
| 2 | -0.484 | 0.3817 | 0.304 | 2.6156 | 3.0 | 5.0 | 1.5693 |
| 3 | -0.428 | 0.4193 | 1.2351 | 2.8918 | 3.0 | 5.0 | 1.7351 |
| 4 | -0.396 | 0.4396 | 1.3295 | 3.0492 | 3.0 | 5.0 | 1.8295 |
| 5 | -0.245 | 0.5263 | 1.775 | 3.7916 | 3.0 | 5.0 | 2.275 |
| 6 | -0.203 | 0.5484 | 1.8992 | 3.9987 | 3.0 | 5.0 | 2.3992 |
| 7 | -0.181 | 0.5598 | 1.9653 | 4.109 | 3.0 | 5.0 | 2.4654 |
| 8 | -0.171 | 0.5646 | 1.9935 | 4.1558 | 3.0 | 5.0 | 2.4935 |
| 9 | -0.16 | 0.5703 | 2.0274 | 4.2123 | 3.0 | 5.0 | 2.5274 |
| 10 | -0.157 | 0.5716 | 2.0349 | 4.2249 | 3.0 | 5.0 | 2.5349 |
| 11 | -0.119 | 0.5905 | 2.1491 | 4.4151 | 3.0 | 5.0 | 2.6491 |
| 12 | -0.112 | 0.5938 | 2.1691 | 4.4486 | 3.0 | 5.0 | 2.6691 |
| 13 | -0.086 | 0.606 | 2.2451 | 4.5752 | 3.0 | 5.0 | 2.7451 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 14 | -0.07 | 0.6138 | 2.2944 | 4.6573 | 3.0 | 5.0 | 2.7944 |
| 15 | -0.052 | 0.622 | 1.1061 | 4.7447 | 3.0 | 5.0 | 2.8468 |
| 16 | -0.048 | 0.6236 | 1.1125 | 4.7618 | 3.0 | 5.0 | 2.8571 |
| 17 | -0.038 | 0.6283 | 1.1314 | 4.812 | 3.0 | 5.0 | 2.8872 |
| 18 | -0.016 | 0.6386 | 2.4542 | 4.9236 | 3.0 | 5.0 | 2.9542 |
| 19 | -0.009 | 0.6494 | 2.5258 | 5.0431 | 3.0 | 5.0 | 3.0258 |
| 20 | -0.004 | 0.6474 | 2.5127 | 5.0212 | 3.0 | 5.0 | 3.0127 |
| 21 | -0.003 | 0.6441 | 2.4908 | 4.9846 | 3.0 | 5.0 | 2.9908 |
| 22 | -0.003 | 0.6443 | 2.4919 | 4.9866 | 3.0 | 5.0 | 2.9919 |
| 23 | 0.0008 | 0.6458 | 2.5025 | 5.0042 | 3.0 | 5.0 | 3.0025 |
| 24 | 0.0009 | 0.6459 | 2.5029 | 5.0049 | 3.0 | 5.0 | 3.0029 |
| 25 | 0.0075 | 0.6488 | 2.5221 | 5.0368 | 3.0 | 5.0 | 3.0221 |
| 26 | 0.0119 | 0.6507 | 2.5351 | 5.0585 | 3.0 | 5.0 | 3.0351 |
| 27 | 0.0139 | 0.6517 | 2.5413 | 5.0689 | 3.0 | 5.0 | 3.0414 |
| 28 | 0.0368 | 0.6617 | 1.2704 | 5.1815 | 3.0 | 5.0 | 3.1089 |
| 29 | 0.0434 | 0.6646 | 1.2827 | 5.214 | 3.0 | 5.0 | 3.1284 |
| 30 | 0.0878 | 0.6836 | 1.365 | 5.4327 | 3.0 | 5.0 | 3.2596 |
| 31 | 0.0915 | 0.6851 | 2.7708 | 6.0817 | 3.0 | 5.0 | 3.2708 |
| 32 | 0.2373 | 0.7442 | 3.202 | 6.1699 | 3.0 | 5.0 | 3.702 |
| 33 | 0.2743 | 0.7586 | 3.3115 | 6.3526 | 3.0 | 5.0 | 3.8115 |
| 34 | 0.3696 | 0.7943 | 3.5932 | 6.8220 | 3.0 | 5.0 | 4.0932 |
| 35 | 0.4429 | 0.8208 | 3.8102 | 7.1837 | 3.0 | 5.0 | 4.3102 |
| 36 | 0.4526 | 0.8242 | 2.0419 | 7.2315 | 3.0 | 5.0 | 4.3389 |
| 37 | 0.5475 | 0.8572 | 2.2179 | 7.6992 | 3.0 | 5.0 | 4.6195 |
| 38 | 0.5935 | 0.8727 | 2.3033 | 7.9259 | 3.0 | 5.0 | 4.7556 |
| 39 | 0.6669 | 0.897 | 2.4365 | 8.2878 | 3.0 | 5.0 | 4.9727 |
| 40 | 0.6934 | 0.9056 | 2.4886 | 8.4185 | 3.0 | 5.0 | 5.0511 |

It is evident from Table 1 that the existence of correlation or otherwise between pairs of normal deviates is related to the implied values of the variance-covariance matrix of the random disturbances generated from them. It is observed that when upper triangular matrix, $P_1$ is used for the transformation $Var(U_1)$ and $Cov(U_1,U_2)$ approached thier specified values as correlation approaches zero while $Var(U_2)$ is equal to the specified value of 3.0. However, when the lower triangular matrix $P_2$ is used the $Var(U_2)$ is equal to the specified value of 5.0. Regardless of whether $P_1$ or $P_2$ is used the covariances are unaffected. Also, the correlation coefficients between pairs of random disturbance terms are positive even when the correlation coefficients between the normal deviates from which they are generated are negative. This is due to fact that the random disturbances are linear combination(s) of the normal deviates. Normal deviates with feeble or negligible correlation coefficients were able to reproduce the assumed variance-covariance matrix (see the coloured area) regardless of the triangular matrix; lower or upper.

## V. Conclusion

In this study, the method suggested by Nagar (1969) was used to transform the series of normal deviates to conform with the specified variance covariance matrix $\Sigma$. Since $\Sigma$ is a positive definite matrix, we decomposed it by a non-singular upper triangular matrix ($P_1$) and also by a non-singular lower triangular matrix ($P_2$).

The study found that the random disturbances generated from those which were feebly (negatively or positively) correlated reproduced the specified variance covariance matrix while others failed. When $P_1$ was used $Var(U_1)$ and $Cov(U_1,U_2)$ approached the specified value as correlation approached zero while the $var(U_2)$ was exactly the same as the specified value. In contrast however, when $P_2$ was used $Var(U_2)$ and $Cov(U_1,U_2)$ approached the specified value as correlation approached zero while the $Var(U_1)$ was the same with the specified value.

In conclusion our findings have shown that Monte-Carlo studies on simultaneous equation estimators which unintentionally or otherwise include pairs of inherently correlated normal deviates would not be basing their comparison of the performance of these estimators on equal footing. It is therefore suggested that in comparing the performances of simultaneous equation estimators, pairs of normal deviates to be used should be screened and only those pairs which are feebly correlated should be used.

## References

C. Nelson and R. Startz (1990), "The Distribution of the Instrumental Variables Estimator and its T-Ratio", Journal of Business, 63, 5125-5140.

Cragg, J.G. (1966), "On the Sensitivity of Simultaneous equation Estimators to the Stochastic Assumptions of the Models", Journal of America Statistical Association, 61, 136-151.

Cragg, J.G. (1967), "On the Relative Small Sample Properties of Several Structural Equation Estimators", Econometrica, 35, 89-110.

Cragg, J.G. (1968), "Some Effects of Incorrect Specification on the Small Sampling Properties of Several Simultaneous Equation Estimators", International Economic Review , Vol. 9, pp. 63-86.

D. Hendry (1984), "Monte Carlo Experimentation in Econometrics", Handbook of Econometrics, Vol, 2, Z. Griliches and M.D. Intriligator. Elsevier, 944.

Johnston, J and Dinardo, J., Econometric Methods, Fourth Edition, McGraw-Hill International.

L. Matyas and L. Lovrics (1991), "Missing Observations and panel Data – A Monte Carlo Analysis, "Economic Letters, 37, 39-44.

Nagar, A.L. (1959), "The Bias and Moment Matrix of the General-Class Estimators of the Parameters in Simultaneous Equations", Econometrica, Vol. 27, pp. 575-595.

Nagar, A.L. (1960), "A Monte Carlo Study of Alternative Simultaneous Equation Estimators", Econometrica, 28, 573-590.

Olubusoye, O.E. (2001), "Relative Sensitivity of Estimators of Simultaneous Equations to Violation of Mutual Independence of Random Deviates", Unpublished Ph.D. Thesis, University of Ibadan

Quandt, R.E. (1962), "Some Small Sample Properties of Certain Structural Equation Estimators". Econometric Research Program, Princeton, Research Memorandum, No. 48.

Quandt, R.E. (1965), "On Certain Small Sample Properties of K-Class Estimators" International Economic Review, 6, 92-104.

Summers, R. (1965), "A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators", Econometrica, 33, 1-41.

Wagner, H.M. (1958), "A Monte Carlo Study of Estimators of Simultaneous Linear Structural Equations", Econometrica, 26, 127-133.