A COMPARISON OF ALTERNATIVE ESTIMATORS OF NIGERIA'S MACROECONOMIC MODELS

[CEAR Occasional Paper No 2008-02]

by

Olaomi, J. O. and Olubusoye O.E.
CEAR, University of Ibadan, Ibadan, Nigeria

Abstract

In this work, we have used CEAR Macroeconomic models, the Nigeria's macroeconomic data and some of the FAIR Estimation techniques, OLS, OLSAUTO1, 2SLS and LIML, to determine the best methods to be used in a multi-period prediction of macroeconomic variables. We have shown that a single method, like the usual OLS, should not be used throughout the models and that a particular method is best for just a set of variables.

1.0 Introduction

The Centre for Econometric and Allied Research (CEAR), University of Ibadan, has been actively involved in the development of Perspective Plan for Nigeria. In this connection, a macroeconomic model had been developed and revised on regular basis. The model specifications were based on theoretical postulates and mere knowledge of the structure and workings of the Nigerian economy without serious emphasis on the existing data constraints and exact form of each a priori equation specification for estimation. (Ogunkola and Olofin (1993)).

Although the development of macroeconomic model for Nigeria has reached a stage where the models is now being used occasionally for forecasting purposes, the model do not appear as yet of producing accurate forecast. This is due largely to data problem and the estimation approach which is essentially the ordinary least square (OLS). In the estimation of the CEAR model for instance, no account was taken of both simultaneous-equation bias and first-order serial correlation of the error terms. It is known that time series data are usually autocorrelated. A considerable gain in forecasting accuracy may be achieved by considering these issues.

The focus of this paper is to compare the forecasting accuracy of OLS , LIML and two stage least squares (2SLS) - a more advanced estimation technique. Due to the large size of the model, only the three single equation techniques were used. Ideally, the system estimators are appropriate but the computational burden is enormous; they may lead to solutions that are highly nonlinear in the parameters and therefore often difficult to determine and lastly, if there is a specification error in one or more equations of the system, that error is transmitted to the rest of the system. In practice, the single equation methods are often used. However, they lead to estimates that are consistent but, in general, not asymptotically efficient. The reason for the lack of asymptotic efficiency is the disregard of the correlation of the disturbances across equations. In other words, single equation estimators do not take into account prior restrictions on other equations in the model.

It should be noted from the outset that this study is not a comparison of estimators in terms of the standard properties of unbiasedness, efficiency, and consistency. Rather, this study is an attempt to determine, using and existing model of the Nigerian economy, which estimators lead to the most accurate multi-period predictions. The estimators are compared in terms of the accuracy of the within-sample predictions. Although some of these properties will be noted in the discussion of the estimators. This is done in part so that the ranking of the of the estimators in terms of forecasting accuracy can be compared with the ranking of the estimators in terms of properties like consistency and efficiency. This study is based on the premise that the basic properties of macroeconomic models are similar enough so that the conclusions obtained from the use of one model can be generalized to other models. The CEAR model is large (49 equations and 24 identities), linear and was designed primarily for short-run forecasting purposes.

Fair (1973) compared the performance of ten estimators applied to seven stochastic equations and an identity of the U.S. Macroeconomic model in terms of their multi-period forecasting accuracy, using the root mean squared error (RMSE). First-order serial correlation tended to be fairly pronounced in most of the equations, whereas second-order serial correlation tended to be much less pronounced. The estimators are:

(i) Ordinary Least Squares (OLS)

(ii) Two – Stage Least Squares (2SLS)

(iii) OLS plus first-order serial correlation (OLSAUTO1)

(iv) 2SLS plus first-order serial correlation (2SLSAUTO1)

(v) OLS plus first- and second-order serial correlation (OLSAUTO2)

(vi) 2SLS plus first- and second-order serial correlation (2SLSAUTO2)

(vii) Full-information maximum likelihood (FIML)

(viii) FIML plus first-order serial correlation (FIMLAUTO1)

(ix) FIML plus first-and second-order serial correlation (FIMLAUTO2)

(x) Accounting for the dynamic nature of the model (DYN)

His result shows that for GNP variable, 2SLS estimators perform on average better than their OLS counterparts, that the full-information maximum likelihood estimators perform on average better than their 2SLS counterparts, that the AUTO1 estimators perform on average better than their no-serial-correlation counterparts, and that the AUTO2 estimators perform on average better than their AUTO1 counterparts.

The rest of this paper is divided into four parts. In part two, we discuss the estimation methods. In part three, we describe the CEAR model and the data used for the study. In part four discuss the results and compare the performance of estimators. The summary and conclusion are given in the last part, part five.

2.0 The Estimators

Basically two estimators are considered but each of them is also used to account for the first-order serial correlation of the error terms due to possible serial correlation of the error terms.

The general model

The general model to be estimated is

$$AY + BX = U \qquad (1)$$

where Y is an $h \times T$ matrix of endogenous variables, X is $k \times T$ matrix of predetermined (both exogenous and lagged endogenous) variables, U is an $h \times T$ matrix of error terms, and A and B are $h \times h$ coefficient matrices respectively. T is the number of observations. The *i*th equation of the model will be written as

$$y_i = -A_i Y_i - B_i X_i + u_i, \quad i = 1, 2, ..., h, \quad (2)$$

where y_i is a $1 \times T$ vector of values of y_{it}, Y_i is an $h_i \times T$ matrix of endogenous variables (other than y_i) included in the *i*-th equation, X_i is a $k_i \times T$ matrix of predetermined variables included in the *i*-th equation, u_i is a $1 \times T$ vector of error terms, and A_i and B_i are $1 \times h_i$ and $1 \times k_i$ vectors of coefficients corresponding to the elements of A and B respectively.

The error terms in U are assumed to follow a second-order auto regressive process:

$$U = R^{(1)} U_{-1} + R^{(2)} U_{-2} + E \qquad (3)$$

where the R matrices are $h \times h$ coefficient matrices, E is an $h \times T$ matrix of error terms, and the subscripts denote lagged values of the terms of U. The error terms in E are assumed to have zero expected values, to be contemporaneously correlated but not serially correlated, and to be uncorrelated in the limit with the predetermined, lagged predetermined, and lagged endogenous variables.

Ordinary least squares (OLS)

The first estimator considered is ordinary least squares applied to each equation of (2). Ordinary least squares estimator does not, of course, produce consistent estimates of the coefficients of the model. The estimates are inconsistent both because of the correlation between u_i and Y_i in (2) and because of the correlation between u_i and the lagged endogenous variables in X_i in (2).

Ordinary least Squares plus first-order serial correlation (OLSAUTO1)

The second estimator considered accounts for first-order serial correlation of the error term u_i in (2), but not for simultaneous-equation bias. The estimator is based on the assumption that the error term in each equation is first-order serially correlated:

$$u_i = r_{ii}^{(1)} u_{i-1} + e_i, \qquad i = 1, 2, ..., h, \qquad (4)$$

which means that $R^{(1)}$ in (3) is assumed to be a diagonal matrix and $R^{(2)}$ in (3) to be zero. Under this assumption, equations (2) and (4) can be combined to yield:

$$y_i = r_{ii}^{(1)} y_{i-1} - A_i Y_i + r_{ii}^{(1)} A_i Y_{i-1} - B_i X_i + r_{ii}^{(1)} B_i X_{i-1} + e_i, \quad (5)$$
$$i = 1, 2, \ldots, h.$$

Ignoring the fact that y_i and e_i are correlated, equation (5) is a simple nonlinear equation in the coefficients $r_{ii}^{(1)}$, A_i, and B_i and can be estimated by a variety of techniques. Two of the most common techniques are the Cochrane-Orcutt iterative technique and the Hildreth-Lu scanning technique, but any standard technique for estimating nonlinear equation can be used. The Cochrane-Orcutt technique was used for this study.

Two-stage least Squares (TSLS)

The third estimator considered is two-stage least squares applied to each equation of (2). Two-stage least squares produces consistent estimates if the error term u_i in (2) is not serially correlated or if there are no lagged endogenous variables in X; otherwise not. With a large enough sample, all of the variables in X should be used as regressors in the first-stage regression for each equation. In practice, however, it is usually necessary to use only a subset of variables in X as regressors or to use only certain linear combinations of all of the variables in X as regressors. A necessary condition for TSLS to produce the consistent estimates is that the included predetermined variables in the equation being estimated be in the set of regressors. Otherwise there is no guarantee that TSLS will produce consistent estimates even if the error term is not serially correlated or if there are no lagged endogenous variables among the predetermined variables. However, in the case of instrumental-variable estimator, this condition is not relevant. The predicted values from the first stage regressions merely serve as instruments for the right-hand-side endogenous variables. For this study, therefore, the variables in X_i were always included in the set of regressors when the *i*th equation of (2) was estimated by TSLS.

Two-stage least squares plus first-order serial correlation (TSLSAUT01)

The fourth estimator considered is two-stage least squares applied to each equation of (5). This estimator accounts for both first-order serial correlation and simultaneous-equations bias and produces consistent estimates if $R^{(1)}$ is diagonal and $R^{(2)}$ is zero in (3). This estimator is discussed in Fair Ray C. (1970), where it is shown that the following variables must be included as regressors in the first-stage regressions in order to insure consistent estimates of equation (5): y_{i-1}, Y_{i-1}, X_i and X_{i-1}. For this study, these variables were always included in the set of regressors. Any standard nonlinear technique can be used for the second-stage regression of equation (5), the Cochrane-Orcutt technique was also used here.

3.0 The CEAR Mac-Model and the Data

The CEAR model was designed primarily for short-run forecasting purposes. It consists of 49 stochastic equations and 24 identities. The models are virtually all linear in nature and the parameters estimated using the OLS method. The variables are normally from the Nigerian yearly economic times series data. The SOLVE (using Gauss – Seidel) method in TSP was used to solve the model, in order to produce the periodic forecast of the key variables in the economy. The data used were from sectors like Agriculture, Petroleum, Manufacturing, Building and construction, Transport amongst many others and it consists of twenty-three years (1981 to 2003) period.

4.0 Model Forecasting

In this study, the existing CEAR Mac model was used but the estimation method was not only OLS. We used OLS, OLSAUTO1, 2SLS, and LIML. As in CEAR Mac model, predictions were made for 2004 to 2011. These predicted values are then subjected to statistical analysis using the minimum variance and minimum RMSE criteria. The variance and the RMSE are given respectively as:

$$\text{VARIANCE} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2 \quad \text{and} \quad \text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}$$

The estimator that produces the minimum error is adjudged the best for multi-period forecasting.

We intend to further this research by using the ten Fair estimation methods and many more, in quest of a formidable method of estimating the parameters of the Nigerian macroeconomic model.

The table 1 below shows the performance of the estimators on all the endogenous variables in the model, using both the variance and the RMSE criteria while Table 2 gives the method most suitable to estimate and forecast the variables.

Table 1: The ROOT MEAN SQUARE ERROR and VARIANCE CRITERIA FOR ALL ENDOGENEOUS AND IDENTITY VARIABLES

RMSE				
VAR	OLS	2SLS	LIML	CORC
CFR	11609.14	10586.69	6.54E+12	1119412
CFTR	3315.509	1760.043	20537161	2394.724
DD	1724196	1580804	3682225	1997727
EA	3.61751	2.46023	81.38033	3.223816
ETC	0.227596	0.200832	0.213014	0.227107
GCR	2989.672	2595.666	3105.524	2926.151
M01R	1476.384	484.9881	16160.06	1563.814
M24R	1312.378	153.1119	8640.895	1106.503
M59R	7069.193	1618.14	983937	8523.056
PCR	40336.55	35212.62	38054.6	40151.86
TCR	43324.18	37804.94	41138.61	43068
TYR	6519110	23521.72	5.89E+13	3122336
YAR	13595.65	9649.557	103323.7	15057.96
YCR	744.0251	327.0173	1.88E+10	36734.19
YDR	205312.1	6557.95	6.28E+12	57279.22
YFDR	7924.703	3706.889	95551.29	9847.74
YGR	277.1922	253.0002	13124.11	374.5575
YLMR	10726.45	813.3859	8.08E+10	39992.66
YOR	6349450	91714.75	6.5E+13	3231728
ITTED	425.2731	400.9513	425.2731	400.9513
P	0.00522	0.004922	0.00522	0.004922
TOEL	38349.81	36156.55	38349.81	36156.55
BOCD	61781.98	1523062	1.56E+15	16816.83

VARIANCE				
VAR	LIML	OLS	2SLS	CORC
CFR	4.29E+25	53562828	46947332	1.09E+11
CFTR	4.5E+14	2256145	880010	1437776
DD	6.21E+12	1.08E+12	1E+12	1.76E+12
EA	3311.763	3.561793	2.706021	4.458327
GCR	10093098	9701041	8352024	10378402
M01R	87823551	1791805	121397.8	1944702
M24R	25895330	131469.9	30112.99	88045.35
M59R	6.1E+11	36787218	3036907	48531447
MD	7.07E+10	5.07E+10	4.61E+10	3.98E+11
PCR	1.59E+09	1.76E+09	1.52E+09	2E+09
TCR	1.85E+09	2.03E+09	1.76E+09	2.29E+09
TYR	3.38E+27	6.86E+12	4.58E+08	2.3E+10
YAR	3.95E+09	78495036	43282699	90951095
YCR	3.43E+20	200725	17707.2	16660984
YDR	3.84E+25	8.45E+09	7554484	9.16E+08
YFDR	3.53E+09	33275708	12408332	43952057
YGR	71896763	87792.7	79027.52	125229.3
YLMR	6.16E+21	25971898	128770.9	1.23E+09
YOR	4.1E+27	6.49E+12	2.26E+09	1.79E+10
ETC	0.011679	0.013486	0.011679	0.015914
ITTED	66137.4	66137.4	66137.4	66137.4
P	3.09E-05	3.09E-05	3.09E-05	3.09E-05
TOEL	3.46E+08	3.46E+08	3.46E+08	3.46E+08

BOT	7959940	79092070	2.01E+17	5269415
CHP	2010689	1909871	4328014	1052532
EBC	0.26956	0.240069	148.0923	0.160079
ED	1.653099	1.630375	2.248213	1.0752
EUT	0.059896	0.135851	730316.1	0.016235
FPI	920.4926	983.8814	2349.48	310.3036
M3R	852523.9	368442.2	1.16E+15	180267.6
MGS1	836812.8	346541.8	1.16E+15	167479.5
MS	1537310	1442582	1530090	26329.47
NORR	1287.991	4610.998	4890.702	886.5062
ORD	341.2102	20173.72	21397.46	143.6321
ORR	38425.97	2470426	2620282	16528.64
TE	7.154791	5.949857	730530.7	5.476446
TEE	1.72E+08	4.02E+09	4.65E+18	1.28E+08
TER	7914423	78457349	2.01E+17	4170109
TGRR	37651.48	2467947	2617653	16210.12
TMR	861468.9	370280.3	1.16E+15	190828.7
TPI	1431.997	1471.261	1560.508	396.9428
TXR	7244712	78457349	2.02E+17	5140643
X3R	7243660	79092070	2.02E+17	5140633
X59R	811.0381	817.8942	11724841	798.8778
XGS1	7251287	78457349	2.02E+17	5147453
YFSR	691.348	3596.363	4298.368	184.0184
YLR	806.791	3596.363	4298.368	184.0184
YOCR	5761.24	6705.001	7957.666	5559.403
TCER	54066.73	72224.15	47235.03	159555.2
AFPI	3450.884	6330.904	4932053	3574.412
CFAR	934.8883	18633.15	6.54E+12	1112470
CFBCR	1297.756	6051.736	69185.64	3679.961

BOCD	2.36E+30	7.97E+08	7.49E+11	2.07E+08
BOT	3.92E+34	1.23E+13	5.04E+15	6.42E+10
CHP	6.68E+12	9.34E+11	9.33E+11	4.82E+11
EBC	16855.25	0.024541	0.031198	0.001927
ED	1.346841	0.707827	0.78147	0.004355
EMQ	0.001457	2.68E-05	0.000913	2.53E-05
ES	1.346841	0.237079	0.78147	0.004355
EUT	5.03E+11	0.001655	0.006927	0.000291
FPI	619976.5	57811.76	41211.04	15299.68
INCR	3.92E+34	8.81E+11	5.37E+15	4.61E+11
M3R	1.37E+30	5.05E+11	9.9E+10	2.28E+10
MGS1	1.37E+30	5.15E+11	1E+11	2.54E+10
MS	2.8E+11	2.83E+11	2.8E+11	1.04E+08
NORR	7850917	988392	7850917	739817.9
ORD	1.27E+08	7380.992	1.27E+08	133.8309
ORR	2E+12	79022144	2E+12	9597803
TE	5.03E+11	13.87904	12.52816	6.047427
TER	3.92E+34	1.21E+13	5.06E+15	2.82E+11
TGRR	2.01E+12	87188418	2.01E+12	15150608
TMR	1.37E+30	5.15E+11	1E+11	2.53E+10
TPI	58218.04	100158.3	58218.04	42938.15
TXR	3.97E+34	8.52E+12	5.06E+15	7.12E+10
X3R	3.97E+34	8.52E+12	5.04E+15	7.14E+10
XGS1	3.97E+34	8.52E+12	5.06E+15	7.14E+10
YFSR	5431548	73370.86	3689276	623.8865
YLR	5431548	42348.77	3689276	623.8865
YOCR	13253579	9660364	9782936	8518488
AFPI	2.01E+13	1166563	8838770	1704329
CFAR	4.29E+25	191020.5	1.13E+08	1.07E+11

CFMER	17079.93	18633.15	6.54E+12	1112470	
CPI	86.28826	500.7424	760813.9	740.3999	
EM	0.613932	0.636108	0.674694	0.667807	
EMQ	0.007906	0.059348	0.087607	0.047569	
ES	0.948947	1.630375	2.248213	1.0752	
INCR	1467179	80596767	2.01E+17	7287818	
MD	426372.1	440414.8	779493.7	945225.8	
OPI	161.5286	293.4109	81170.62	218.1691	
OYR	33678.83	98421.06	3.97E+11	111316.5	
TD	306954.6	383719.2	3.13E+08	331633.7	
TGER	66656.13	812491.3	794451.8	204472.1	
TRER	12719.9	744157.1	835919.1	44954.93	
X01R	87.33446	3572.638	1.88E+08	104.9563	
X24R	351.9071	17281.83	5.98E+09	688.6371	
YBCR	349.8388	2698.054	16716.68	415.7602	
YFR	185.3313	337.1296	2.64E+10	6769.603	
YMR	10545.07	15699.49	8.22E+10	40367.85	
YSMR	182.6962	16480.68	7.85E+08	446.6583	
YTAR	13328.9	16116.22	2.61E+10	20818.11	
YTR	311.6566	2446.458	31421856	356.8151	

CFBCR	2.91E+09	862059.7	9581272	3176263
CFMER	4.29E+25	1.03E+08	1.13E+08	1.07E+11
CPI	4.78E+11	8448.218	162508.5	76053.77
EM	0.120727	0.102113	0.120727	0.12517
OPI	5.44E+09	277.1416	8145.165	638.7255
OYR	1.49E+23	2.97E+08	3.19E+09	1.01E+10
TCER	6.56E+08	5.14E+08	2.04E+09	4.35E+09
TD	7.7E+16	3.84E+10	4.48E+10	4.65E+10
TEE	2.1E+37	7.68E+15	1.96E+19	1.24E+16
TGER	1.93E+11	6.83E+08	2.15E+11	7.52E+09
TRER	2.05E+11	12682154	1.82E+11	4.3E+08
X01R	3.54E+16	1307.934	4298902	2025.674
X24R	3.57E+19	16450.57	73325701	30484.5
X59R	1.14E+14	21732.71	33537.08	22924.7
YBCR	1.7E+08	136433.7	2205340	221592
YFR	6.95E+20	16686.58	57132.63	4693471
YMR	6.38E+21	25240496	84575034	1.25E+09
YSMR	6.16E+17	5522.876	90274478	34313.74
YTAR	6.83E+20	77113609	1.07E+08	1.3E+08
YTR	1.05E+15	19935.53	1700264	31920.78

Table 2: Estimation method most suitable to predict variables

METHOD	RMSE	VARIANCE
2SLS	CFR, CFTR, DD, EA, ETC, GCR, M01R, M24R, M59R, PCR, TCR, TYR, YAR, YCR, YDR, YFDR, YGR, YLMR, YOR, ITTED, P, TOEL	CFR, CFTR, DD, EA, GCR, M01R, M24R, M59R, MD, PCR, TCR, TYR, YAR, YCR, YDR, YFDR, YGR, YLMR, YOR, ETC.
OLS	AFPI, CFAR, CFBCR, CFMER, CPI, EM, EMQ, ES, INCR, MD, OPI, OYR, TD, TGER, TRER, X01R, X24R, YBCR, YFR, YMR, YSMR, YTAR, YTR	AFPI, CFAR, CFBCR, CFMER, CPI, EM, OPI, OYR, TCER, TD, TEE, TGER, TRER, X01R, X24R, X59R, YBCR, YFR, YMR, YSMR, YTAR, YTR
CORC	ITTED, P, TOEL, BOCD, BOT, CHP, EBC, ED, EUT, FPI, M3R, MGS1, MS, NORR, ORD, ORR, TE, TEE, TER, TGRR, TMR, TPI, TXR, X3R, X59R, XGS1, YFSR, YLR, YOCR	BOCD, BOT, CHP, EBC, ED, EMQ, ES, EUT, FPI, INCR, M3R, MGS1, MS, NORR, ORD, ORR, TE, TER, TGRR, TMR, TPI, TXR, X3R, XGS1, YFSR, YLR, YOCR
LIML	TCER	ETC
ALL		ITTED, P, TOEL

The results in the tables are fairly self – explanatory, and only a brief description of them will be presented. The result shows that for all variables in the macroeconomic model, a single estimation method should not be used. The best method for each variable is as presented in Table 2.

For instance, Output in Agriculture (YTAR), Output in Manufacturing (YMR), Output in Building and Construction (YBCR), Output in Transport (YTR) and Output in 'Others' should be predicted using the OLS; Output in Petroleum(YOR), Output in Communication (YCR), Output in Distribution (YDR), Total Output (TYR), Total Consumption (TCR), Capital Formation (CFR) and both Government and Private Consumption should be predicted using 2SLS; while OLSAUTO1 (CORC) should be used to predict Total Imports (TMR), Total Government Revenue (TGRR), Oil Revenue (ORR), Non-Oil Revenue (NORR) amongst others.

5.0 Conclusion

We have shown that using the CEAR Macroeconomic Model and the Nigeria's macroeconomic data (between 1981 and 2003), with a forecast from 2004 to 2011, a single estimation method alone should not be used to estimate the parameters of the models and hence, to predict the multi-period values of the variables.

As mentioned in the introduction, the conclusions of this study are based heavily on the premise that the basic properties of macroeconomic models are similar and that the particular model used is a good representative of macroeconomic models. The results give an indication of the relative usefulness of the various estimators for multi-period forecasting purposes. In general, however, the results do indicate that considerable forecasting accuracy can be achieved by using more complicated techniques than simple OLS or 2SLS to estimate macroeconomic models. Because of the increased feasibility of using more advanced estimation techniques, there is now less need for model builders to limit themselves to simpler techniques if more advanced techniques lead to improved results.

References

Fair, C. Ray (1973): A comparison of alternative estimators of macroeconomic Models. International Economic Review, Vol 14,No 2.

Ogunkola O. and Olofin S.O. (1993): CEAR Mac Model IV. CEAR, University of Ibadan.