# A Modified Least Squares Estimator in Estimating Generated Regressor Models[1]


George P. Mitsopoulos
*Athens University of Economics and Business*
*Department of Economics*
*and*
*Ministry of Economics and Finance*
*Financial and Economic Crime Unit Service*

Michael A. Magdalinos[2]
*Athens University of Economics and Business*
*Department of Economics*


April 2003

ABSTRACT


This paper examines the asymptotic and small sample performance of the most commonly used estimators in generated regressor models, under the hypothesis that the structural and expectation errors are correlated. Furthermore, a modified least squares estimator is proposed, which turns out to be numerically the same as Pagan's double length regression. Our approach, however, has several computational advantages. The structural equation can be transformed and estimated using only ordinary least squares. Moreover, unlike most existing methods, the covariance matrix is consistently estimated by the unmodified least squares covariance matrix.



*Key words:* Generated Regressor; Modified least squares; Expectation formation; Monte Carlo.

*JEL classification:* C13; C15.


---


[1]*The authors are grateful to K. Abadir, A. Chesher, K. Hadri, G. Phillips, and E. Tzavalis for helpful comments on the first draft of the paper.*


[2]*To the memory of professor Michael Magdalinos who was co-author of this paper and died on 18-08-2002*

Contents

# 1. INTRODUCTION

Over the last twenty years the role of expectations formation in both theoretical and applied economics has been of central importance. As the number and scope of empirical applications in this area have expanded, a parallel theoretical econometric literature developed and quickly became very complex [see Cuthbertson et.al. (1992), pp.155-190, and Oxley and McAleer (1995) for relevant surveys]. However, Barro's (1977) simple two-step approach remains very popular, in spite of its several shortcomings [see Pagan (1984), (1986), Gauger (1989), McAleer and McKenzie (1991), (1992), and Smith and McAleer (1994)]. The lists of applied papers in Oxley and McAleer (1995) shows that the two-step least squares method is the preferred option, with few papers bothering to calculate the correct standard errors. It is often asserted that the choice of the two-step method is purely a matter of computational convenience. This is only part of the truth. The two-step approach remains attractive because it provides a natural and intuitively simple framework where a number of interesting economic hypotheses can be tested [see Cuthbertson et.al. (1992), Hoffman et. al. (1984), Murphy and Topel (1985), McAleer and McKenzie (1991), Smith and McAleer (1994) and Oxley and McAleer (1995)].

The present paper seeks to preserve the simple and natural framework of Barro's procedure, while providing fully efficient estimators and asymptotically valid estimates of their covariance matrix. The proposed method is a generalized least squares estimation in a modified version of the structural equation with generated regressor proxies. The estimator of the structural coefficients is numerically identical to Pagan's (1986) double-length regression. Computationally, however, the proposed estimator is much more convenient, as the error covariance matrix can be analytically inverted and the structural equation can be transformed and estimated by ordinary least squares. Moreover, unlike most existing methods, the covariance matrix is consistently estimated by the unmodified least squares covariance matrix.

The rest of the paper is organized as follows. The basic expectation model is introduced in Section 2, and efficiency comparisons are made for the most popular estimation methods. The new estimation method is introduced in Section 3, where its main properties are also derived. The small sample properties of the main estimation procedures are compared in Section 4 by some Monte Carlo experiments. General remarks are presented in the concluding

Section 5. To improve the readability of the paper, all the proofs are relegated in an Appendix.

## 2. ASYMPTOTIC COMPARISON OF ESTIMATORS

Consider the simple specification examined by Pagan (1984) and Turkington (1985), among many others. They examine a two equation model:

$$y = \bar{z}\alpha + X\beta + u \,, \tag{1}$$

$$z = \bar{z} + v \,,\ \bar{z} = Z\gamma \tag{2}$$

where $X$ and $Z$ are matrices of exogenous variables, $y$ and $z$ are vectors of observable endogenous variables, $u$ and $v$ are vectors of disturbances, and $\bar{z}$ is the vector of unobservable expected value of $z$. The scalar $\alpha$ and the vectors $\beta$, $\gamma$ are unknown parameters. It is convenient to write equation (1) as

$$y = Y\delta + u \,,\ Y = \begin{pmatrix} \bar{z} & X \end{pmatrix},\ \delta' = \begin{pmatrix} \alpha & \beta' \end{pmatrix}. \tag{3}$$

The errors $\xi_t' = \begin{pmatrix} u_t & v_t \end{pmatrix}$ are independent and identically distributed with zero mean and finite third order absolute moments. We write

$$\Sigma = E\left(\xi_t \xi_t'\right) = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}, \tag{4}$$

and we assume that the two errors are not perfectly correlated, that is

$$\det(\Sigma) = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2 > 0 \,. \tag{5}$$

Moreover, as the sample size $(T)$ tends to infinity the matrices $X'X/T$, $Z'Z/T$, $X'Z/T$ converge in probability to finite matrices of full column rank. The vector $\bar{z}$ does not belong to the space $R(X)$ spanned by the columns of $X$. For any matrix $X$ we write $P_X$ and $\overline{P}_X$ for the orthogonal projectors into the spaces $R(X)$ and $R(X)^{\perp}$, respectively.

The most popular estimators for this model are the following:

(i) The *two step least squares* (TSLS) estimator proposed by Barro (1977). In (1) the unobservable variable $\bar{z}$ is substituted by $\hat{z} = P_Z z$, and the equation is estimated by least squares (LS). Pagan (1984) shows that the TSLS estimator is consistent, but the conventionally computed LS standard errors are not.

(ii)The *instrumental variables* (IV) estimator, proposed by McCallum (1976). The unobservable $\bar{z}$ in (1) is substituted by the true variable $z$, and the equation is estimated by IV using the set of instruments $F=\begin{pmatrix}X & Z\end{pmatrix}$. Pagan (1984) shows that the IV estimator and the estimated standard errors are consistent.

(iii) The *two step generalized least squares* (TSGL) estimator proposed by Hoffman (1987). The substitution of $\bar{z}$ by $\hat{z}$ produces a new error term in (1) whose variance matrix is used to estimate (1) by generalized least squares, or double length regression [Higgins (1994), Ma and Liu (1996)]. The TSGL estimator and the computed standard errors are consistent.

(iv) The *full information* (FI) estimators, like the full information maximum likelihood (FIML), proposed by Wallis (1980), the three stage least squares (3SLS), proposed by Wickens (1982) , or the double length regression (DLR), proposed by Pagan (1986). The FI estimators are consistent and asymptotically efficient.

Let $V_J$, $(J=TSLS, TSGL, IV, FI)$ be the asymptotic variance matrix of the estimated parameters $\delta'=\begin{pmatrix}\alpha & \beta'\end{pmatrix}$ of equation (1) using the methods (i), (ii), (iii), or (iv) respectively. The explicit form of these matrices are:

*LEMMA 1. The asymptotic variance matrices of the estimators (i)-(iv) are:*

$$V_{TSLS}=\sigma^2\Lambda^{-1}(\Lambda-\lambda_1\Delta)\Lambda^{-1},\ V_{IV}=\sigma^2\Lambda^{-1}, \tag{6}$$

$$V_{TSGL}=\sigma^2(\Lambda+\kappa_1\Delta)^{-1},\ V_{FI}=\sigma^2(\Lambda+\kappa\Delta)^{-1}, \tag{7}$$

*where ,*

$$\sigma^2=\sigma_u^2-2\alpha\sigma_{uv}+\alpha^2\sigma_v^2, \tag{8}$$

$$\kappa=\left(\sigma_{uv}-\alpha\sigma_v^2\right)^2\Big/\left(\sigma_u^2\sigma_v^2-\sigma_{uv}^2\right), \tag{9}$$

$$\kappa_1=\alpha\left(\alpha\sigma_v^2-2\sigma_{uv}\right)\Big/\sigma_u^2,\ \lambda_1=\kappa_1/(\kappa_1+1),$$

*and*

$$\Lambda=\operatorname{plim}\frac{1}{T}\begin{bmatrix} z'z & z'X \\ X'z & X'X\end{bmatrix},\ \Delta=\operatorname{plim}\frac{1}{T}\begin{bmatrix} 0 & 0' \\ 0 & X'\overline{P}_Z X\end{bmatrix}. \tag{10}$$

For any square and symmetric matrices $V_I$, $V_J$ we write $V_I\le V_J$ if the matrix $V_J-V_I$ is a positive semi-definite matrix. The operation $\le$ is a partial ordering. The ordering of the asymptotic variance matrices of the estimators (i)-(iv) is given in the following two theorems:

*THEOREM 1. If* $R(X) \subset R(Z)$ *then the TSLS, IV, and TSGL estimators are identical, and they are fully efficient, that is,*

$$V_{TSLS} = V_{IV} = V_{TSGL} = V_{FI}. \tag{11}$$

*THEOREM 2. If the parameter* $\alpha$ *belongs to the interval* $\left(0, 2\sigma_{uv} / \sigma_v^2\right)$, *then*

$$V_{FI} \le V_{IV} \le V_{TSGL} \le V_{TSLS}. \tag{12}$$

*Otherwise,*

$$V_{FI} \le V_{TSGL} \le V_{TSLS} \le V_{IV}. \tag{13}$$

In many applications the errors in the two equations are assumed to be uncorrelated $(\sigma_{uv} = 0)$. This restriction is rarely justified when the model contains only anticipated variables. Nevertheless, we have the following interesting result:

*COROLLARY 1. If* $\sigma_{uv} = 0$, *then*

$$V_{FI} = V_{TSGL} \le V_{TSLS} \le V_{IV}. \tag{14}$$

## 3. THE MODIFIED LEAST SQUARES ESTIMATOR

Corollary 1 suggests that the inefficiency of the TSGL estimator is caused by the error correlation in the system of equations (1) and (2). The correction of this problem is likely to produce a fully efficient estimator. To do so, we orthogonalize $u$ with respect to $v$, thus obtaining the vector

$$w = u - qv, \; q = \sigma_{uv} / \sigma_v^2 . \tag{15}$$

From construction, the vector $w$ is orthogonal to $v$ and it has variance

$$\sigma_w^2 = \sigma_u^2 - \sigma_{uv}^2 / \sigma_v^2 . \tag{16}$$

Substituting (15) in (1) we find

$$y - qv = \bar{z}\alpha + X\beta + w. \tag{17}$$

Since $v$ and $\bar{z}$ are not observable, we may use the least squares decomposition of $z$ into $R(X)$ and $R(X)^{\perp}$ respectively, that is

$$\hat{z} = P_Z z = \bar{z} + P_Z v \, , \; \hat{v} = \overline{P}_Z z = v - P_Z v \, . \tag{18}$$

From (17) and (18) we find

$$y - q\hat{v} = \hat{z}\alpha + X\beta + e \, , \; e = w + (q - \alpha) P_Z v \, . \tag{19}$$

The variance matrix of new disturbances $e$ is

$$E(ee') = \sigma_w^2 Q \, , \; Q = I_T + \kappa P_Z \, , \tag{20}$$

where $I_T$ is the $T \times T$ identity matrix, and $\sigma_w^2$, $\kappa$ are given in (16) and (9). From (5) we have that $0 \le \kappa < \infty$, so the eigenvalues of $Q$ (either $1$ or $1 + \kappa$) are positive, and $Q$ is positive definite. Therefore, given any consistent estimates of $q$ and $\kappa$, it is natural to define the generalized least squares estimator

$$\hat{\delta} = \left( \hat{Y}' \hat{Q}^{-1} \hat{Y} \right)^{-1} \hat{Y}' \hat{Q}^{-1} (y - \hat{q}\hat{v}), \tag{21}$$

where

$$\hat{Y} = \begin{pmatrix} \hat{z} & X \end{pmatrix}, \; \hat{Q} = I_T + \hat{\kappa} P_Z \, . \tag{22}$$

The estimator becomes computationally feasible because it is not necessary to invert numerically the $T \times T$ matrix $Q$. It is easy to verify that

$$\begin{aligned} \hat{Q}^{-1} &= I_T - \hat{\lambda} P_Z \, , \; \hat{\lambda} = \hat{\kappa} / (\hat{\kappa} + 1), \\ \hat{Q}^{-1/2} &= I_T - \hat{\mu} P_Z \, , \; \hat{\mu} = 1 - (\hat{\kappa} + 1)^{-1/2} . \end{aligned} \tag{23}$$

The matrix $\hat{Q}^{-1/2}$ can be used to transform equation (19) into

$$y - \hat{\mu}\hat{y} - \hat{q}\hat{v} = (1 - \hat{\mu}) \hat{z}\alpha + \left( X - \hat{\mu}\hat{X} \right) \beta + \varepsilon \, , \tag{24}$$

where

$$\hat{y} = P_Z y \; , \; \hat{X} = P_Z X \, . \tag{25}$$

The GLS estimator (21) is the LS estimator in equation (24).

*THEOREM 3. If* $\hat{q}$ and $\hat{\mu}$ are consistent estimates and $\hat{\delta}' = \begin{pmatrix} \hat{\alpha} & \hat{\beta}' \end{pmatrix}$ *is the LS estimator in* (24)*, then as* $T \to \infty$*,*

$$\sqrt{T} \left( \hat{\delta} - \delta \right) \xrightarrow{d} N(0, V_{FI}), \tag{26}$$

*where the matrix* $V_{FI}$ *is given in* (7)*, and it is consistently estimated by the conventionally computed LS covariance matrix.*

In order to estimate $q$ and $\mu$, we need consistent estimates of $\sigma_u^2$, $\sigma_{uv}$ and $\sigma_v^2$. These can be provided by using the residuals from any consistent estimation of equations (1) and (2). In practice, it is easier to use the simplest procedure, that is the TSLS estimation, which also provides the decomposition (18). If the first step of the estimation is the TSLS, then we shall refer to the LS estimator in (24) as the modified least squares (MLS) estimator. Obviously, the MLS estimator is asymptotically fully efficient. Moreover, we can show that it is numerically identical with Pagan's double-length regression estimator.

*THEOREM 4. If the first step of the estimation is the TSLS, then the MLS estimator and the DLR estimator for the parameter vector* $\delta$ *are identical.*

From equation (24) we can see that the MLS estimator can be computed easily with any computer package having a LS subroutine and variable transformations. On the other hand, the DLR estimator is computationally more complicated, requiring some original programming.

An extension of the previous model arises when unanticipated values of $z$, that is, a shock variable, appear in the structural equation:

$$y = v\alpha_1 + \bar{z}\alpha_2 + X\beta + u . \tag{27}$$

In the context of the model of equations (2) and (19) the parameters $q$ and $\alpha_1$ are not simultaneously identified (see Pesaran (1989), p.171). If, however, we can use outside information to fix a value for the parameter $\sigma_{uv}$, then both $q$ and $\alpha_1$ are identified and consistently estimable. The widely used assumption $\sigma_{uv} = 0$ is only one of the possible alternatives. For example, if we can find an exogenous variable correlated with the shock variable $\hat{v}$, we can use it as additional instrument to estimate consistently (27). Then, using the residuals from (2) and (27) we estimate

$$\hat{\sigma}_{uv} = \hat{u}'\hat{v}/T \,,\ \sigma_v^2 = \hat{v}'\hat{v}/T \,,\ \hat{q} = \hat{\sigma}_{uv} \big/ \hat{\sigma}_v^2 . \tag{28}$$

Fixing the value of $\hat{q}$, we can orthogonalize $u$ with respect to $v$ thus obtaining

$$y - \hat{q}\hat{v} = \hat{v}\alpha_1 + \hat{z}\alpha_2 + X\beta + e \ ,\ e = w + \left(\hat{q} + \alpha_1 - \alpha_2\right)P_Z v. \tag{29}$$

As before, we can show that the MLS estimator is the LS estimator in the equation

$$y - \hat{\mu}\hat{y} - \hat{q}\hat{v} = \hat{v}\alpha_1 + (1-\hat{\mu})\hat{z}\alpha_2 + \left(X - \hat{\mu}\hat{X}\right)\beta + \varepsilon, \tag{30}$$

$$\hat{\mu} = 1 - (\hat{\kappa}+1)^{-1/2}, \; \hat{\kappa} = (\hat{q} + \hat{\alpha}_1 - \hat{\alpha}_2)^2 \hat{\sigma}_v^2 / \hat{\sigma}_w^2 . \tag{31}$$

where $\hat{\sigma}_w^2$ can be calculated from (16). Of course, if we assume $\sigma_{uv} = 0$, the same formulae hold with $\hat{q} = 0$, $\hat{\sigma}_w^2 = \hat{\sigma}_u^2$.

## 4. SMALL SAMPLE COMPARISON OF ESTIMATORS

To examine the small sample efficiency of estimators and tests in generated regressor models, several Monte-Carlo experiments have been conducted: See Hoffman (1987), (1991), Hoffman et.al. (1984), Smith and McAleer (1994), among many others. These studies, however, fail to recognize the importance of the position of the parameter $\alpha$ relatively to the interval $\left(0,\ 2\sigma_{uv}/\sigma_v^2\right)$ in the efficiency ordering of competing estimators (see Theorem 2). Consequently, they do not take into account this crucial factor in their experimental design.

The data generating process is the simple model

$$y_t = \alpha\bar{z}_t + 18.9 + 3.7x_{1t} + u_t , \tag{32}$$

$$z_t = \bar{z}_t + v_t,\ \bar{z}_t = 16.4 + 3.5z_{1t} + 2.8z_{2t} \tag{33}$$

where $t = 1,2,\ldots,T$. Theorem 2 suggests that the relative efficiency of the estimators depends on the value taken by the parameter $\alpha$. To account for this factor, we vary the value of $\alpha$. Twenty experiments were conducted, one for each combination of the parameter values

$$T = (15, 30, 50, 100, 200),\ \alpha = (-5.0, 1.1, 5.0, 10.0)$$

For each experiment, the exogenous variables $x_{1t}, z_{1t}, z_{2t}$ were generated[3] as independent uniform $(-5, 5)$ and they are transformed to correlated variables with a theoretical correlation $r^2 = 0.64$. These values were held constant across all the replications of one experiment.

---

[3]*The computations were carried out by a double precision FORTRAN program. The uniform and the normal random numbers were generated via the subroutines G05CAF and G05EZF, respectively, of the NAG library.*

For each replication we produce independently the $N(0,\Sigma)$ vector $\xi_t' = (u_t \quad v_t)$, where

$$\Sigma = \begin{pmatrix} 150 & 135 \\ 135 & 125 \end{pmatrix}$$

and we compute the sampling values of the endogenous variables $y_t$, $z_t$ from (32) and (33), respectively. Then we compute the TSLS, TSGL, MLS/DLR, IV, and 3SLS estimators. Each experiment consists of $10{,}000$ replications and produces Monte-Carlo estimates of the Bias and the Root Mean Square Error (RMSE) of the estimators.

Table 1
Sampling Results for the Estimators of the Parameter $\alpha$

| *T* | *Estimator* | *α = -5.0* | | *α = 1.1* | | *α = 5.0* | | *α = 10.0* | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Bias* | *RMSE* | *Bias* | *RMSE* | *Bias* | *RMSE* | *Bias* | *RMSE* |
| *15* | *TSLS* | 0.424 | 1.254 | -0.004 | 0.317 | -0.277 | 0.845 | -0.628 | 1.810 |
| | *TSGL* | 0.139 | 1.147 | 0.005 | 0.290 | -0.110 | 0.788 | -0.194 | 1.622 |
| | *MLS/DLR* | 0.155 | 1.093 | 0.000 | 0.101 | -0.097 | 0.706 | -0.220 | 1.593 |
| | *IV* | 0.108 | 3.373 | 0.000 | 0.077 | -0.069 | 2.178 | -0.158 | 4.952 |
| | *3SLS* | 1.354 | 20.176 | -0.004 | 0.151 | -0.873 | 12.842 | -1.986 | 29.350 |
| *30* | *TSLS* | 0.052 | 0.704 | 0.001 | 0.125 | -0.032 | 0.464 | -0.074 | 1.024 |
| | *TSGL* | -0.003 | 0.692 | 0.003 | 0.109 | 0.002 | 0.458 | 0.010 | 1.006 |
| | *MLS/DLR* | -0.004 | 0.680 | 0.000 | 0.033 | 0.003 | 0.439 | 0.007 | 0.997 |
| | *IV* | -0.011 | 1.019 | 0.000 | 0.030 | 0.007 | 0.657 | 0.016 | 1.494 |
| | *3SLS* | 0.073 | 0.665 | 0.000 | 0.029 | -0.047 | 0.429 | -0.107 | 0.975 |
| *50* | *TSLS* | 0.115 | 0.645 | -0.002 | 0.167 | -0.077 | 0.434 | -0.173 | 0.928 |
| | *TSGL* | 0.006 | 0.602 | 0.000 | 0.143 | -0.012 | 0.410 | -0.009 | 0.858 |
| | *MLS/DLR* | 0.005 | 0.574 | 0.000 | 0.036 | -0.004 | 0.371 | -0.007 | 0.842 |
| | *IV* | -0.018 | 1.196 | 0.000 | 0.034 | 0.011 | 0.770 | 0.026 | 1.752 |
| | *3SLS* | 0.164 | 0.622 | -0.001 | 0.033 | -0.106 | 0.402 | -0.241 | 0.913 |
| *100* | *TSLS* | 0.060 | 0.501 | 0.000 | 0.116 | -0.038 | 0.335 | -0.087 | 0.726 |
| | *TSGL* | 0.001 | 0.490 | 0.001 | 0.104 | -0.002 | 0.329 | 0.001 | 0.707 |
| | *MLS/DLR* | -0.001 | 0.475 | 0.000 | 0.024 | 0.001 | 0.308 | 0.002 | 0.698 |
| | *IV* | -0.003 | 0.827 | 0.000 | 0.024 | 0.002 | 0.534 | 0.004 | 1.214 |
| | *3SLS* | 0.070 | 0.477 | 0.000 | 0.024 | -0.045 | 0.308 | -0.102 | 0.700 |
| *200* | *TSLS* | 0.027 | 0.279 | -0.001 | 0.074 | -0.019 | 0.188 | -0.042 | 0.401 |
| | *TSGL* | 0.001 | 0.276 | -0.001 | 0.068 | -0.003 | 0.186 | -0.003 | 0.396 |
| | *MLS/DLR* | 0.001 | 0.265 | 0.000 | 0.015 | -0.001 | 0.172 | -0.002 | 0.389 |
| | *IV* | -0.004 | 0.496 | 0.000 | 0.015 | 0.003 | 0.320 | 0.006 | 0.727 |
| | *3SLS* | 0.029 | 0.267 | 0.000 | 0.015 | -0.019 | 0.172 | -0.042 | 0.391 |

From Table 1 we can see that, with few exceptions, the bias is very small and the estimator variance is the main component of the RMSE. For all $\alpha \neq 1.1$, the MLS/DLR estimator is the best, followed by the TSGL estimator for $T \leq 50$ and by the 3SLS estimator for $T > 50$.

The parameter value $\alpha = 1.1$ is the middle point of the interval $\left(0,\ 2\sigma_{uv} / \sigma_v^2\right) = (0,\ 2.16)$ and it is the point of the parameter space where the asymptotic relative efficiency of the TSGL and MLS/DLR estimators with respect to the IV estimator takes its minimum value (see Theorem 2). Hence, it is not surprising that for $T = 15$ the IV estimator is better than the MLS/DLR estimator. The difference becomes negligible for $T \geq 30$.

Table 2.
Sampling Results for the Estimators of the Parameter $\beta_1 = 3.7$

| *T* | *Estimator* | *α = -5.0* | | *α = 1.1* | | *α = 5.0* | | *α = 10.0* | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Bias* | *RMSE* | *Bias* | *RMSE* | *Bias* | *RMSE* | *Bias* | *RMSE* |
| *15* | *TSLS* | -2.983 | 5.714 | .027 | 2.173 | 1.951 | 4.057 | 4.418 | 8.080 |
| | *TSGL* | -.977 | 3.495 | -.038 | 1.961 | .774 | 2.957 | 1.356 | 4.153 |
| | *MLS/DLR* | -1.080 | 2.697 | .001 | .665 | .678 | 1.768 | 1.533 | 3.760 |
| | *IV* | -.576 | 19.666 | .003 | .473 | .374 | 12.685 | .849 | 28.848 |
| | *3SLS* | -8.349 | 120.135 | .027 | .902 | 5.384 | 76.366 | 12.242 | 174.610 |
| *30* | *TSLS* | -.620 | 1.926 | -.006 | 1.079 | .387 | 1.498 | .891 | 2.591 |
| | *TSGL* | -.104 | 1.185 | -.027 | .927 | .074 | 1.158 | .114 | 1.210 |
| | *MLS/DLR* | -.099 | .414 | -.001 | .232 | .059 | .319 | .136 | .537 |
| | *IV* | -.029 | 6.411 | .001 | .190 | .020 | 4.139 | .045 | 9.408 |
| | *3SLS* | -.747 | 1.316 | .005 | .186 | .484 | .864 | 1.098 | 1.924 |
| *50* | *TSLS* | -.826 | 2.133 | .016 | 1.102 | .554 | 1.633 | 1.245 | 2.935 |
| | *TSGL* | -.086 | 1.206 | -.001 | .938 | .110 | 1.198 | .130 | 1.232 |
| | *MLS/DLR* | -.079 | .341 | .005 | .209 | .055 | .275 | .119 | .436 |
| | *IV* | .076 | 6.740 | .003 | .198 | -.043 | 4.338 | -.102 | 9.871 |
| | *3SLS* | -1.109 | 2.219 | .009 | .187 | .718 | 1.443 | 1.629 | 3.255 |
| *100* | *TSLS* | -.431 | 1.171 | -.001 | .771 | .275 | .955 | .628 | 1.499 |
| | *TSGL* | -.032 | .795 | -.002 | .691 | .032 | .794 | .030 | .796 |
| | *MLS/DLR* | -.020 | .151 | .001 | .134 | .013 | .143 | .028 | .168 |
| | *IV* | -.006 | 4.463 | .000 | .134 | .003 | 2.882 | .008 | 6.550 |
| | *3SLS* | -.487 | .900 | .003 | .130 | .314 | .585 | .714 | 1.306 |
| *200* | *TSLS* | -.192 | .679 | .007 | .536 | .134 | .599 | .297 | .812 |
| | *TSGL* | -.003 | .543 | .010 | .495 | .019 | .543 | .015 | .542 |
| | *MLS/DLR* | -.004 | .093 | .001 | .092 | .003 | .093 | .006 | .094 |
| | *IV* | .034 | 3.030 | .000 | .092 | -.022 | 1.955 | -.050 | 4.445 |
| | *3SLS* | -.206 | .373 | .001 | .091 | .133 | .248 | .302 | .534 |

From Table 2 we can see that the alternative estimators of the parameter $\beta = 3.7$, the slope parameter of $x_{1t}$ in (32), have similar properties. Again, for all $\alpha \neq 1.1$, the MLS/DLR estimator is the best, followed by the TSGL estimator for $T \leq 50$ and by the 3SLS estimator for $T > 50$. Again, for $\alpha = 1.1$ and $T = 15$ the IV estimator is slightly better than the MLS/DLR estimator.

Both tables show that the 3SLS estimator performs very poorly in small samples, and that its convergence towards efficiency is rather slow. For example, we can see that in Table 2 and for $\alpha \neq 1.1$, the MLS/DLR estimator is significantly more efficient that the 3SLS estimator even when $T = 200$. This behaviour is rather odd and it is due in the fact that a small percent of the 3SLS estimates are far away from the true parameters value. For example, in the case of the parameter $\alpha = -5.0$, $T = 15$ (Table 1) the RMSE of the 3SLS is 20.176 since a 0.5% (50 from 10.000) of the estimates lie in the interval $(15, 1601)$, also in the case of the parameter $\beta_1 = 3.7$, $T = 15$ (Table 2), 1% (100 from 10.000) of the estimates lie in the interval $(-9821, -59)$ so that the RMSE of the 3SLS estimator is 120.135.

The experiments reveal an interesting pattern in the relative finite sample efficiency of the five competing estimators. This pattern is related not only with the sample size, but also with the position of the parameter $\alpha$ relatively to the interval $\left(0, 2\sigma_{uv} / \sigma_v^2\right)$, as suggested by Theorem 2.

The influence of the parameter $\alpha$ on the distribution of the estimators can be easily seen diagrammatically. Figures 1-9 display the distribution of the estimators under different values of the parameter $\alpha$ in the sample size, $T = 50$.

It is clear that the distribution of all estimators is symmetric around the true parameters value in the middle point $\alpha = 1.1$ (Figures 2, 5, 8). Moreover the distribution of the MLS, IV and 3SLS estimators are more concentrated around the true value followed by TSGL and TSLS estimators.

For the values of the parameter $\alpha \neq 1.1$, not only the concentration of the estimators is affected, but also the shape of their distribution around the true value (Figures 1, 3, 4, 6, 7, 9). Figures 1 and 3 shows that the distribution of the MLS estimator is more concentrated and symmetric around the true values $\alpha = -5.0$ and $\alpha = 10.0$, than the other estimators. Figures 7 and 9 shows that the concentration of the MLS estimator for the parameter $\beta_1 = 3.7$ is very impressive compared with the other estimators. Lastly, Figures 4 and 6 shows the behaviour of the estimators for the parameter $\beta_0 = 18.9$ which is analogous as in the case of the $\beta_1$ parameter.

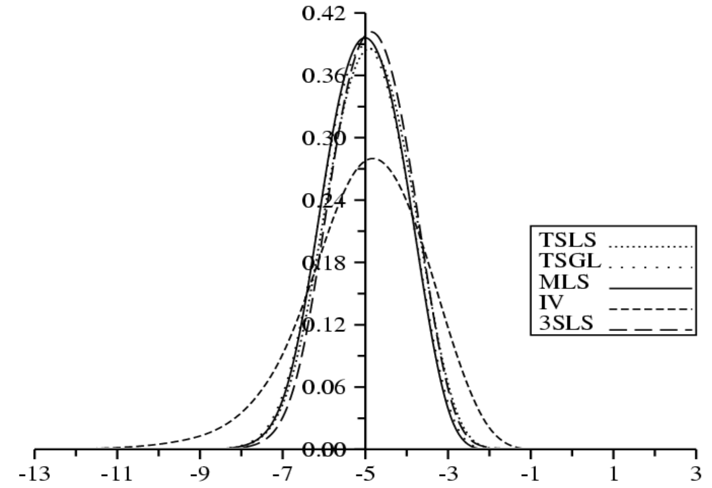


Figure 1. Density of the estimators for the parameter $\alpha_1 = -5.0$, $(T = 50)$

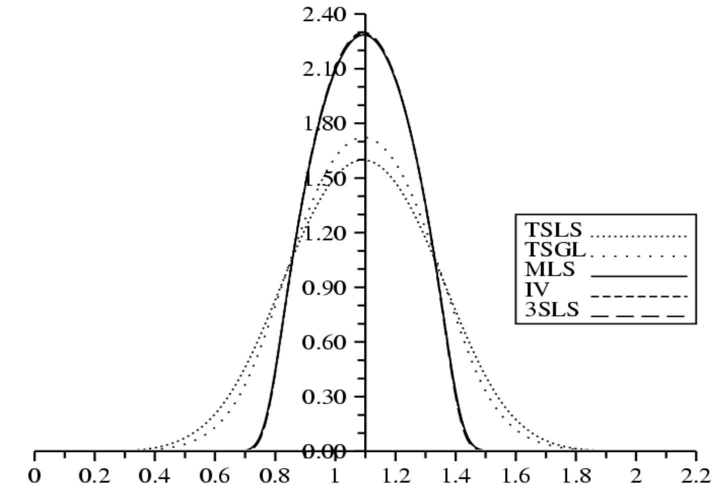


Figure 2. Density of the estimators for the parameter $\alpha_1 = 1.1$, $(T = 50)$

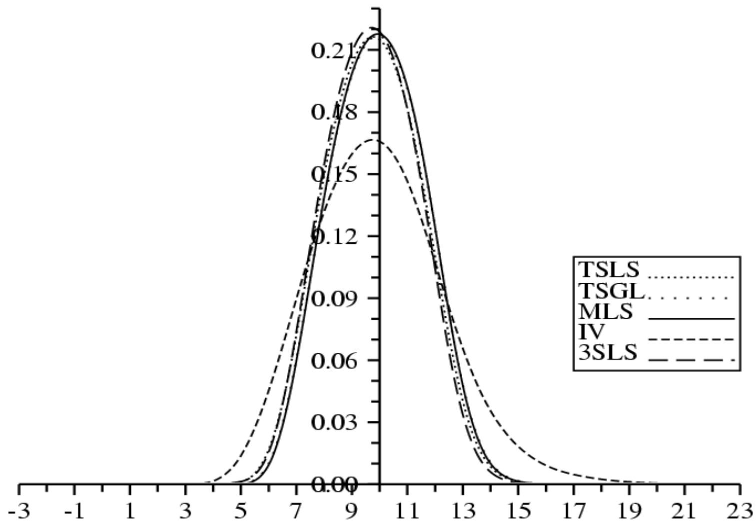


Figure 3. Density of the estimators for the parameter $\alpha_1 = 10.0$, $(T = 50)$

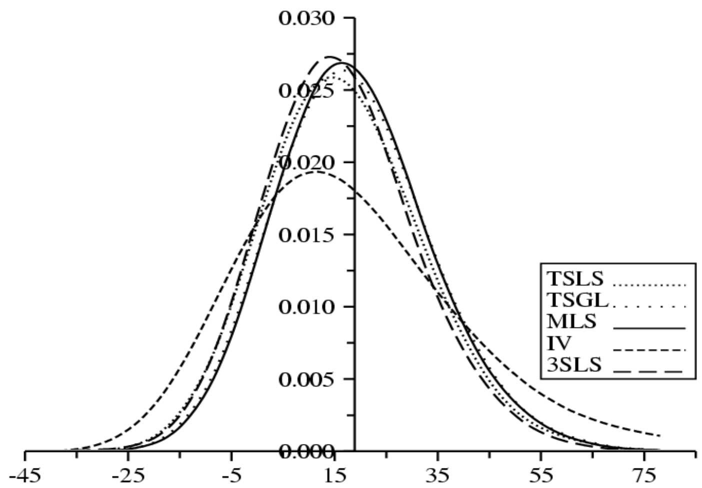


Figure 4. Density of the estimators for the parameter $\beta_0 = 18.9$, $(\alpha = -5.0, T = 50)$

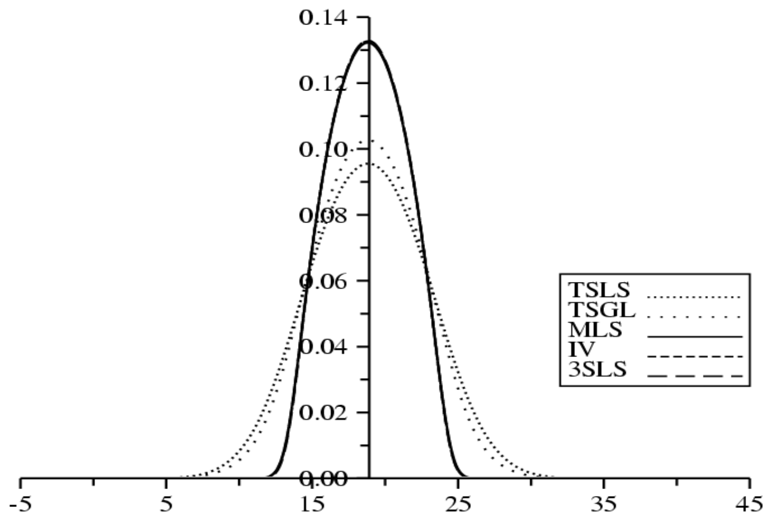


Figure 5. Density of the estimators for the parameter $\beta_0 = 18.9$, $(\alpha = 1.1, T = 50)$

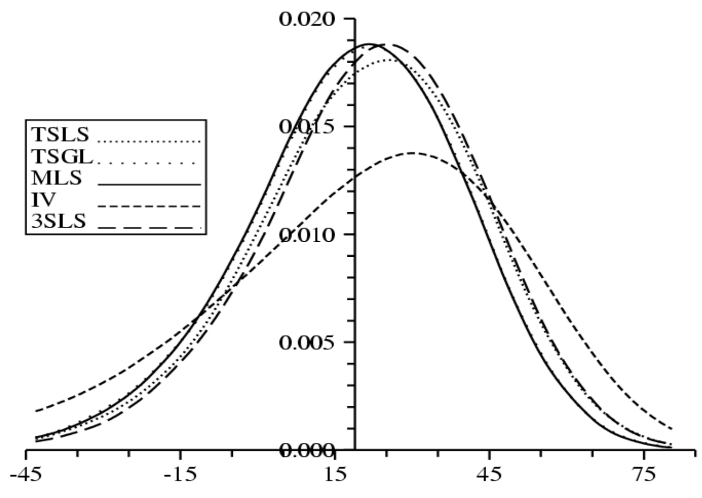


Figure 6. Density of the estimators for the parameter $\beta_0 = 18.9$, $(\alpha = 10.0, T = 50)$

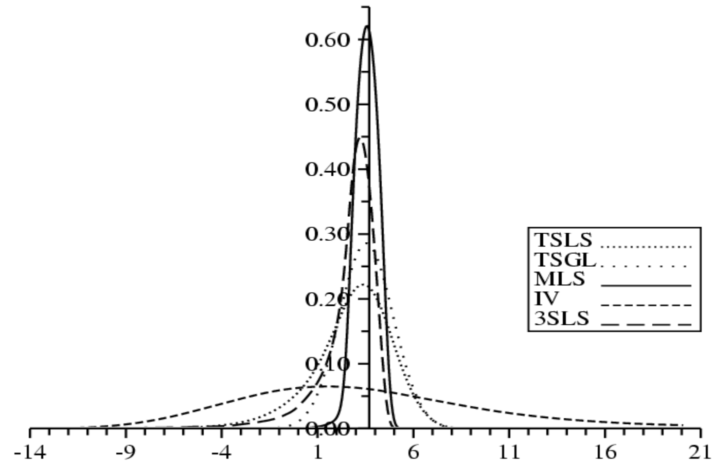


Figure 7. Density of the estimators for the parameter $\beta_1 = 3.7$, $(\alpha = -5.0, T = 50)$

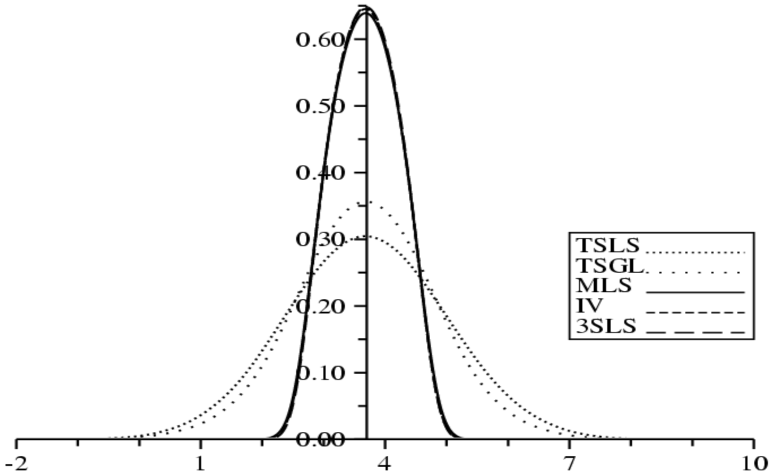


Figure 8. Density of the estimators for the parameter $\beta_1 = 3.7$, $(\alpha = 1.1, T = 50)$

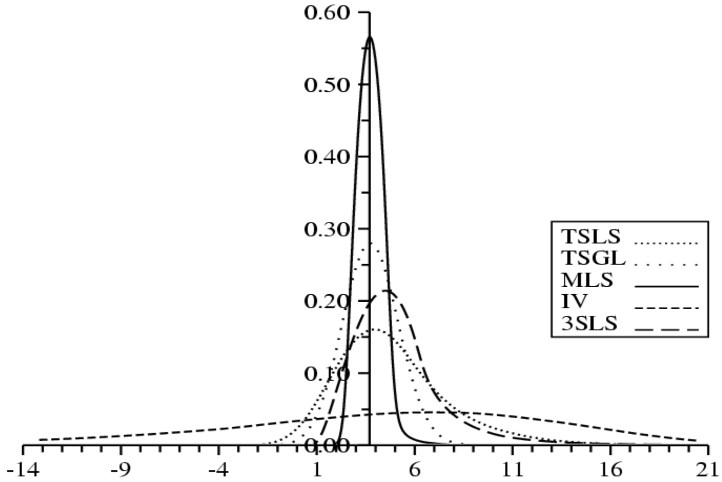


Figure 9. Density of the estimators for the parameter $\beta_1 = 3.7$, $(\alpha = 10.0, T = 50)$

## 5. CONCLUDING REMARKS

This paper proposes a new estimation technique in situations where some regressors are derived as functions of the output from another regression. The main area of application is the expectation literature, but the same techniques can be used in several unobserved variable models, for example in the models developed by Zellner (1970) and Goldberger (1972).

The proposed estimation technique is based on Barro's (1977) two-step approach, and has the same desirable properties: The simple framework is preserved allowing economic hypotheses to be tested in a natural way. The estimator can be easily computed using any computer package having a LS subroutine. Perhaps more importantly, the conventionally computed LS covariance matrix, the standard errors, and the *t* (and *F*) statistics do not require any adjustment, unlike the corresponding quantities from the DLR estimation of Pagan (1986), Higgins (1994) and Ma and Liu (1996). The Monte Carlo experiments show that the small sample properties of the proposed estimator are satisfactory: In 32 out of 40 Monte Carlo experiments the MLS estimator has minimum root mean square error among five competing estimators. Figures 1 and 2 show that the finite sample superiority of the MLS estimator is often quite dramatic.

Several generalizations seem to be possible: Lagged values of expectation and shock variables can be included in the equation, and standard and non-parametric corrections in the case of non-spherical errors can be applied without loss of consistency or asymptotic efficiency. These generalizations will be the subject of a future paper. The computational convenience of the MLS estimator and its satisfactory asymptotic and small sample properties suggest that it might be a useful tool in applied econometric research.

## APPENDIX

*Proof of Lemma 1:* The formulae for $V_{TSLS}$, $V_{IV}$ and $V_{FI}$ are given in Turkington (1985). The proof for $V_{TSGL}$ is very similar to the proof of Theorem 3.

*Proof of Theorem 1:* For any $\kappa > -1$, let $Q = I_T + \kappa P_Z$, $\hat{Y} = (\hat{z} \quad X)$. Since $\hat{z}, X \in R(Z)$ we have $R(\hat{Y}) \subset R(Z)$ and

$$\hat{Y}'Q^{-1} = \hat{Y}'(I_T - \lambda P_Z) = (1-\lambda)\hat{Y}'.$$

Therefore, the TSLS and TSGL estimators are identical. For the IV estimator, the matrix of regressors and of instruments are

$$X_+ = (z \;\; X),\; F = (X \;\; Z),$$

respectively. Since $R(F) = R(Z)$,

$$X_+'P_F = X_+'P_Z = \begin{pmatrix} z'P_Z \\ X' \end{pmatrix} = \begin{pmatrix} \hat{z}' \\ X' \end{pmatrix} = \hat{Y}'.$$

Therefore, the IV and TSLS estimators are identical. The asymptotic efficiency follows from the fact that $\overline{P}_Z X = 0$ implies $\Delta = 0$.

*Proof of Theorem 2:* From the standard GLS theory we have $\boldsymbol{V_{TSGL} \leq V_{TSLS}}$. McAleer and McKenzie (1991) give conditions under which the equality holds. From Lemma 1 we have

$$V_{FI}^{-1} - V_{TSGL}^{-1} = (\kappa - \kappa_1)\big/\sigma^2 \Delta, \tag{34}$$

$$V_{FI}^{-1} - V_{IV}^{-1} = \kappa\big/\sigma^2 \Delta, \tag{35}$$

$$V_{TSGL}^{-1} - V_{IV}^{-1} = \kappa_1\big/\sigma^2 \Delta, \tag{36}$$

$$V_{IV} - V_{TSLS} = \kappa_1 \sigma^2 \big/ (\kappa_1 + 1)\Lambda^{-1}\Delta\,\Lambda^{-1}. \tag{37}$$

It is easy to show that $\kappa \geq 0$, and that $\kappa - \kappa_1 \geq 0$. From (34) we can see that $V_{FI} \leq V_{TSGL}$, and (35) implies that $V_{FI} \leq V_{IV}$. Also notice that $\sigma^2 > 0$ implies $\kappa_1 > -1$. Therefore, if $\alpha \in \left(0, 2\sigma_{uv}\big/\sigma_v^2\right)$ then $-1 < \kappa_1 < 0$ and (36) implies $V_{IV} \leq V_{TSGL}$. If $\alpha \notin \left(0, 2\sigma_{uv}\big/\sigma_v^2\right)$ then $\kappa_1 > 0$ and (37) implies that $V_{TSLS} \leq V_{IV}$.

*Proof of Corollary 1:* When $\sigma_{uv} = 0$, we have $\kappa = \kappa_1 = \alpha^2\sigma_v^2\big/\sigma_u^2 \geq 0$, and the Corollary follows from (34)-(37).

*Proof of Theorem 3:* Equation (24) can be written as

$$y_* = Y_*\delta + \varepsilon,$$

where

$$y_* = y - \hat{\mu}\hat{y} - \hat{q}\hat{v},\; Y_* = \left[(1-\hat{\mu})\hat{z} \quad X - \hat{\mu}\hat{X}\right],\; \varepsilon = \hat{Q}^{-1/2}e.$$

Since $\hat{q}$ and $\hat{\mu}$ are consistent estimates,

$$\sqrt{T}\left(\hat{\delta}-\delta\right)=\left(Y_*'Y_*/T\right)^{-1}Y_*'\varepsilon\big/\sqrt{T}=\left(\hat{Y}'Q^{-1}\hat{Y}\big/T\right)^{-1}\hat{Y}'Q^{-1}e\big/\sqrt{T}+o_p(1)$$

and substituting $\hat{z}=\bar{z}+P_Z v$, we have

$$\sqrt{T}\left(\hat{\delta}-\delta\right)=\left(Y'Q^{-1}Y\big/T\right)^{-1}Y'Q^{-1}e\big/\sqrt{T}+o_p(1) \tag{38}$$

Moreover,

$$Y'Q^{-1}e\big/\sqrt{T}=Y'Q^{-1}w\big/\sqrt{T}+(q-\alpha)(1-\lambda)\,Y'P_Z v\big/\sqrt{T}\ ,$$

which is the sum of two uncorrelated random vectors converging in distribution to normal limits. Therefore, it converges to a normal vector with zero mean and variance the probability limit of the matrix

$$\begin{aligned} S &= \sigma_w^2\,Y'Q^{-2}Y\big/T+\sigma_v^2(q-\alpha)^2(1-\lambda)^2\,Y'P_Z Y\big/T \\ &= \sigma_w^2\,Y'\Big[\left(I_T-\lambda P_Z\right)^2+\kappa\,(1-\lambda)^2 P_Z\Big]Y\big/T \\ &= \sigma_w^2\,Y'\left(I_T-\lambda P_Z\right)Y\big/T=\sigma_w^2\,Y'Q^{-1}Y\big/T\ . \end{aligned}$$

Using the notation (10) it is easy to see that

$$Y'Q^{-1}Y\big/T=(1-\lambda)(\Lambda+\kappa\Delta)+o_p(1) \tag{39}$$

so the vector (38) converges to a normal vector with zero mean and variance

$$\sigma_w^2\left[(1-\lambda)(\Lambda+\kappa\Delta\,)\right]^{-1}=\sigma_w^2(\kappa+1)(\Lambda+\kappa\Delta)^{-1}=\sigma^2(\Lambda+\kappa\Delta)^{-1}=V_{FI}\ .$$

The standard LS covariance matrix estimator is $T\hat{V}$ , where from (39)

$$\begin{aligned} \hat{V} &= s^2\left(Y_*'Y_*/T\right)^{-1}=s^2\left(Y'Q^{-1}Y\big/T\right)^{-1}+o_p(1) \\ &= s^2(\kappa+1)(\Lambda+\kappa\Delta)^{-1}+o_p(1) \end{aligned}$$

and

$$\begin{aligned} s^2 &= \varepsilon'\overline{P}_{Y_*}\varepsilon\big/T=e'Q^{-1}e\big/T+o_p(1) \\ &= w'w/T+o_p(1)=\sigma_w^2+o_p(1) \end{aligned}.$$

Consequently $\hat{V}=V_{FI}+o_p(1)$ .

*Proof of Theorem 4:* Pagan (1986) in Proposition 3.6 defines the DLR estimator of the parameter vector $\theta'=\left(\delta'\quad\gamma'\right)$ as $\widetilde{\theta}=\hat{\theta}+\Delta\theta$, where $\hat{\theta}$ is the TSLS estimator,

$$\Delta\theta = \left(A'\Omega^{-1}A\right)^{-1} A'\Omega^{-1}\xi \qquad (40)$$

evaluated at $\theta = \hat{\theta}$, and

$$\xi' = \begin{pmatrix} u' & v' \end{pmatrix},\ \Omega = E(\xi\xi') = \Sigma \otimes I_T,\ A = \begin{pmatrix} Y & \alpha Z \\ 0 & Z \end{pmatrix}.$$

Since

$$\Omega^{-1} = \Sigma^{-1} \otimes I_T = \frac{1}{d}\begin{pmatrix} \sigma_v^2 I_T & -\sigma_{uv} I_T \\ -\sigma_{uv} I_T & \sigma_u^2 I_T \end{pmatrix},\ d = \sigma_u^2\sigma_v^2 - \sigma_{uv}^2$$

direct multiplication shows that

$$A'\Omega^{-1}A = \frac{1}{d}\begin{bmatrix} \sigma_v^2 Y'Y & \left(\alpha\sigma_v^2 - \sigma_{uv}\right)Y'Z \\ \left(\alpha\sigma_v^2 - \sigma_{uv}\right)Z'Y & \left(\alpha^2\sigma_v^2 - 2\alpha\sigma_{uv} + \sigma_u^2\right)Z'Z \end{bmatrix}.$$

Now, the inverse matrix, and the corresponding vector are

$$\left(A'\Omega^{-1}A\right)^{-1} = d\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},\ A'\Omega^{-1}\xi = A'\Omega^{-1}\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{d}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \qquad (41)$$

where,

$$B_{11} = \frac{1}{\sigma_v^2}\left(Y'Q^{-1}Y\right)^{-1},\ B_{21} = -\frac{\lambda}{\alpha\sigma_v^2 - \sigma_{uv}}(Z'Z)^{-1}Z'Y\left(Y'Q^{-1}Y\right)^{-1},$$

$$b_1 = \sigma_v^2 Y'(u - qv),\ b_2 = \left(\alpha\sigma_v^2 - \sigma_{uv}\right)Z'u + \left(\sigma_u^2 - \alpha\sigma_{uv}\right)Z'v.$$

Substituting (41) in (40) we find that the correction for $\hat{\delta}$ is

$$\Delta\delta = B_{11}b_1 + B'_{21}b_2 = \left(Y'Q^{-1}Y\right)^{-1}\left(Y'Q^{-1}u - Y'\left(qI_T + \lambda\phi P_Z\right)v\right) \qquad (42)$$

where

$$\phi = \left(\alpha\sigma_{uv} - \sigma_u^2\right) / \left(\alpha\sigma_v^2 - \sigma_{uv}\right).$$

To evaluate (42) at $\theta = \hat{\theta}$ notice that $\hat{u} = y - \hat{Y}'\hat{\delta}$, $\hat{v} = \overline{P}_Z z$ and consequently

$$\begin{aligned}\Delta\delta &= -\hat{\delta} + \left(\hat{Y}'\hat{Q}^{-1}\hat{Y}\right)^{-1}\hat{Y}'\hat{Q}^{-1}y - \hat{q}\left(\hat{Y}'\hat{Q}^{-1}\hat{Y}\right)^{-1}\hat{Y}'\overline{P}_Z z \\ &= -\hat{\delta} + \left(\hat{Y}'\hat{Q}^{-1}\hat{Y}\right)^{-1}\hat{Y}'\hat{Q}^{-1}\left(y - \hat{q}\overline{P}_Z z\right)\end{aligned}$$

showing that the DLR and the MLS estimators are identical.

REFERENCES


Barro, R.J. (1977): "Unanticipated Money Growth and Unemployment in the United States," *American Economic Review,* 67, 101-115.

Cuthbertson K., S.G. Hall, and M.P. Taylor (1992): *Applied Econometric Techniques,* New York: Harvester Wheatsheaf.

Gauger, J. (1989): "The Generated Regressor Correction: Impacts upon Inferences in Hypothesis Testing," *Journal of Macroeconomics,* 11, 383-395.

Goldberger, A.S. (1972): "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables," *International Economic Review,* 13, 1-13.

Higgins M.L. (1994): "Computation of the GLS Estimator of a Model with Anticipated and Unanticipated Effects," *Economics Letters,* 45, 125-129.

Hoffman D.L. (1987): "Two-Step GLS Estimators in Multi-Equation Generated Regressor Models," *Review of Economics and Statistics,* 69, 336-46.

Hoffman D.L. (1991): "Two-Step and Related Estimators in Contemporary Rational Expectations Models: An Analysis of Small-Sample Properties," *Journal of Business and Economic Statistics,* 9, 51-61.

Hoffman D.L., S.A. Low, and D.E. Schlagenhauf (1984): "Tests of Rationality, Neutrality, and Market Efficiency: A Monte Carlo Analysis of Alternative Test Statistics," *Journal of Monetary Economics*, 14, 339-363.

Ma Y., and S. Liu (1996): "A Double Length Regression Computation Method for the 2SGLS Estimator of Rational Expectations Models," *Oxford Bulleting of Economics and Statistics,* 58, 423-429.

McAleer M., and C.R. McKenzie (1991): "When are the Two Step Estimators Efficient ?," *Econometric Reviews,* 10, 235-252.

McAleer M., and C.R. McKenzie (1992): "Recursive Estimation and Generated Regressors," *Economics Letters,* 39, 1-5.

McCallum, B.T. (1976): "Rational Expectations and the Natural Rate Hypothesis, Some Consistent Estimates," *Econometrica,* 29, 315-335.

Murphy, K.M. and R.H. Topel (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics*, 3, 370-379.

Oxley, L., and M. McAleer (1995): "Testing the Rational Expectation Hypothesis in Macroeconometric Models with Unobserved Variables," in L. Oxley, D. George, C. Roberts, and S. Sayer (eds), *Surveys in Econometrics*, Oxford: Blackwell.

Pagan, A. (1984): "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review,* 25, 221-247.

Pagan, A. (1986): “Two Stage and Related Estimators and Their Applications,” *Review of Economic Studies,* LIII, 517-538.

Pesaran, M.H. (1989): *The Limits of the Rational Expectations,* Oxford: Basil Blackwell.

Smith, J., and M. McAleer (1994): “Newey-West Covariance Matrix Estimates for Models with Generated Regressors,” *Applied Economics*, 26, 635-640.

Turkington, D.A. (1985): “A Note on Two-Stage Least Squares, Three-Stage Least Squares and Maximum Likelihood Estimation in an Expectation Model," *International Economic Review,* 26, 507-510.

Wallis, K.F. (1980): "Econometric Implications of the Rational Expectation Hypothesis," *Econometrica,* 48, 49-72.

Wickens, M. (1982): "The Efficient Estimation of Econometric Models with Rational Expectations," *Review of Economic Studies,* 49, 55-68.

Zellner, A. (1970): "Estimation of Regression Relationships Containing Unobserved Variables," *International Economic Review,* 11, 441-454.