

FINITE SAMPLE PROPERTIES OF A MODIFICATION OF THE LIMITED INFORMATION MAXIMUM LIKELIHOOD ESTIMATOR*

By P. A. V. B. SWAMY, J. S. MEHTA and N. S. IYENGAR

Federal Reserve Board, Temple University and Indian Statistical Institute

SUMMARY : In this paper a simple modification of the limited information maximum likelihood method for estimating the co-efficients in a linear economic relationship containing two endogenous variables, is proposed. It is shown that under certain mild assumptions about the sample size, the modified estimators possess finite moments.

1. INTRODUCTION

Consider the k -class estimators of the coefficients of a structural equation as defined by Theil (1971, p. 504). For these estimators k can be either fixed or random. The consistency and asymptotic normality of these estimators are well established. The question of practical importance is whether the size of a currently available econometric sample is large enough to guarantee our reliance on the asymptotic theory. Recently, Anderson and Sawa (1979) have given a partial answer to this question by showing that the desirable asymptotic properties of the "fixed" k -class estimator with $k = 1$ (which is otherwise known as the Theil-Basmann two stage least squares (2SLS) estimator) are not necessarily expected to be relevant to the cases that appear in practice. An interesting result of Anderson and Sawa (1979) is that the distribution of the limited information maximum likelihood (LIML) estimator (which is a member of the "random" k -class estimators with $\lambda \geq 1$ almost surely) approaches normality much faster than that of 2SLS estimator. For this reason we may prefer the LIML estimator to the 2SLS estimator if the sample size is sufficiently large. However, as shown by Mariano and Sawa (1972), the LIML estimator does not possess moments of any positive integer order and thus is inadmissible relative to quadratic loss functions. The

*This paper was presented at the Fourth World Econometric Congress, Aix-en-provence, France, August 28-September 2, 1980. The authors thank Professor Fred Norman for reading several drafts of the introduction to this paper and for providing comments that led to several improvements of the exposition. The views expressed herein are solely those of the authors and do not necessarily represent the views of the organisations to which they belong.

moments of the limiting distribution of an estimator will not coincide with the limits of the moments of the corresponding sequence of distributions unless the necessary and sufficient conditions given in Lukacs (1975, p. 43, Theorem 2.3.4) are satisfied. One of these conditions is that the exact finite sample moments of the estimator exist and are finite. Since the moments of the LIML estimator do not exist, the normal distribution may not provide a good approximation to the distribution of the LIML estimator in finite samples.

The purposes of the present paper are to present a simple modification of the LIML estimators of coefficients in a single equation containing two endogeneous variables and to demonstrate under mild restrictions on the sample size that the modified estimators possess finite moments. This investigation is an extension of an earlier and similar study of the fixed k -class estimators with $0 \leq k \leq 1$ (Mehta and Swamy, 1978). The modified estimators have the same limiting distribution as the unmodified estimators, and the Nagar (1959) type approximations for the moments of modified LIML estimator may be developed. These approximations are useful for obtaining a better approximation than the asymptotic approximation even for quite small samples, at least in some cases.

The model and the corresponding estimators are given in Section 2 and the principal results on the existence of the moments for these estimators are given in Section 3.

2. THE MODEL

Consider the following i -th structural equation of a G -equation model :

$$y_i = Y_i\gamma_i + X_i\beta_i + u_i, \quad \dots \quad (1)$$

where y_i is the $T \times 1$ vector of observations on a left-hand endogenous variable, Y_i is the $T \times G_i$ matrix of observations on the endogenous variables which appear in the equation, X_i is the $T \times K_i$ matrix of observations on the exogenous variables that appear in the equation, u_i is the $T \times 1$ vector of disturbances, and γ_i and β_i are respectively $G_i \times 1$ and $K_i \times 1$ coefficient vectors. Here $G_i \leq G$. In the full model there are $K - K_i$ other exogenous variables whose observations are given in the $T \times (K - K_i)$ matrix X_i^* . We assume that $K_i^* = K - K_i \geq G_i$. In addition it is assumed that u_i is normally distributed, independent of $X = (X_i, X_i^*)$, with mean zero and covariance matrix $\sigma_{ii}I$. We assume that the $T \times K$ matrix X is "fixed" and has full column rank for all $T \geq K$.

We write the reduced form equations for y_t and Y_t as

$$Y_0 = (y_t : Y_t) = X_t(\pi_t : \Pi_t) + X_t^*(\pi_t^* : \Pi_t^*) + (v_t : V_t) \quad \dots (2)$$

and assume that the rank of Π_t^* is G_t . We assume that the rows of $(v_t : V_t)$ are independent drawings from a multivariate normal distribution with zero mean vector and positive definite covariance matrix

$$\Omega = \begin{bmatrix} \omega_{11} & \omega'_{12} \\ \omega_{21} & \Omega_{22} \end{bmatrix}$$

Define

$$M_t = I - X_t(X_t'X_t)^{-1}X_t' \text{ and } M = I - X(X'X)^{-1}X'$$

The LIML estimator of γ_t is

$$g_t(l) = (Y_t Q_l Y_t)^{-1} Y_t' Q_l y_t, \quad \dots (3)$$

where $Q_l = M_t - lM$ and l is the smallest root of $|Y_0' M_t Y_0 - l Y_0' M Y_0| = 0$. It is well known that l is real and ≥ 1 almost surely, see Theil (1971, 503-504). Notice that the definition of the estimator (3) requires the assumption that $\text{rank}(X) = K$. The LIML estimator of β_t is

$$b_t(l) = (X_t' X_t)^{-1} X_t' (y_t - Y_t g_t(l)). \quad \dots (4)$$

Mariano and Sawa (1972) show that when $X'X = T1$, and $G_t = 1$, the estimators (3) and (4) do not possess finite moments of any positive integer order. We consider the Mehta and Swamy (1978) type modifications of the estimators (3) and (4) where the modifications guarantee the existence of the moments for certain sample sizes. The modification of the estimator (3),

$$g_t(l, \mu) = (Y_t' Q_{l\mu} Y_t + \mu I)^{-1} Y_t' Q_{l\mu} y_t \quad \dots (5)$$

where $Q_{l\mu} = I - X_t(X_t'X_t + \mu I)^{-1}X_t' - lM$ and $\mu > 0$ is an arbitrary constant. Similarly, the modification of the estimator (4),

$$b_t(l, \mu) = (X_t' X_t + \mu I)^{-1} X_t' (y_t - Y_t g_t(l, \mu)). \quad \dots (6)$$

Combining the estimators (5) and (6), we can write the modified LIML estimator of $\delta_t = (\gamma_t', \beta_t')$ as

$$d_t(l, \mu) = [Z_t'(I - lM)Z_t + \mu I]^{-1} Z_t'(I - lM)y_t, \quad \dots (7)$$

where $Z_t = (Y_t : X_t)$.

The argument of Mehta and Swamy (1978, 3-4) can be utilized here to show that the estimator (7) approximates the restricted least squares estimator of δ_i subject to a boundedness restriction on the squared length of δ_i . Also, the representation (5) may be utilized to emphasize the errors-in-variables-nature of the problem, see Zellner (1970) and Anderson (1976). On postmultiplying both sides of eq. (2) by $(1 : \gamma_i)'$ and equating the resulting coefficients with those appearing in eq. (1) we obtain

$$\pi_i^* = \Pi_i^* \gamma_i. \quad \dots (8)$$

If π_i^* and Π_i^* were known an estimator of γ_i^* is

$$(\Pi_i^* X_i^* M_i X_i^* \Pi_i^*)^{-1} \Pi_i^* X_i^* M_i X_i^* \pi_i^*. \quad \dots (9)$$

Since π_i^* and Π_i^* are unobservable, we can approximate the estimator (9) by the estimation (5) because $(Y_i' Q_{1\mu} Y_i + \mu I)$ is an estimator of $\Pi_i^* X_i^* M_i X_i^* \Pi_i^*$ and $Y_i' Q_{1\mu} y_i$ is an estimator of $\Pi_i^* X_i^* M_i X_i^* \pi_i^*$.

Thus Zellner (1970) and Anderson (1976) are able to establish a connection between the formulae defining estimators for errors-in-variables models and simultaneous equations. As pointed out by Anderson (1976, p. 8), the estimator (5) is consistent as $T \rightarrow \infty$ for a fixed value of K_i^* . This clarifies the conditions under which the estimators for errors-in-variables models are consistent.

To prove our results on the existence of the moments for the estimators (5) and (6), we need the following results. We define

$$k^* = \min_{c \neq 0} \frac{c' Y_i' M_i Y_i c}{c' Y_i' M_i Y_i c}. \quad \dots (10)$$

Kadiyala (1970) has shown that with probability 1

$$k^* > l \geq 1 \quad \dots (11)$$

and the matrix $Z_i'(I - kM)Z_i$ is positive definite for $0 \leq k < k^*$. Also, k^* and l are real. We may rewrite the matrix $Q_{1\mu}$ in (5) as

$$Q_{1\mu} = I - X_i(X_i' X_i + \mu I)^{-1} X_i' - M + (I - l)M = Q_{1\mu} + (I - l)M \quad \dots (12)$$

where $Q_{1\mu} = I - X_i(X_i' X_i + \mu I)^{-1} X_i' - M$.

3. PRINCIPAL RESULTS

We now specialize (5) to situations where there are two endogenous variables, i.e., $G_i = 1$. It is convenient to reparameterize eq. (1), in order that y_i and Y_i are independent. We let

$$y_i^* = Y_i^* \gamma_i^* + X_i^* \beta_i^* + u_i^* \quad \dots (13)$$

where

$$\Omega = \Psi' \Psi, \quad \Psi = \begin{pmatrix} \xi & 0 \\ \omega\rho & \omega \end{pmatrix}, \quad \omega = \sqrt{\Omega_{22}},$$

$$\rho = \omega_{12}/\Omega_{22}, \quad \xi^2 = (\omega_{11} - \omega_{12}^2/\Omega_{22}),$$

$$(\mathbf{y}_i^* : \gamma_i^*) = (\mathbf{y}_i : Y_i)\Psi^{-1} = (\xi^{-1}\mathbf{y}_i - \xi^{-1}\rho Y_i : \omega^{-1}Y_i),$$

$$\gamma_i^* = \omega(\gamma_i - \rho)/\xi, \quad \beta_i^* = \xi^{-1}\beta_i \text{ and } \mathbf{u}_i^* = \xi^{-1}\mathbf{u}_i;$$

\mathbf{y}_i^* is independent of Y_i^* . The estimator (5) can now be written as

$$g_i^*(l, \mu) = \frac{\rho Y_i^* Q_{1\mu} Y_i^* + \xi^* Y_i^* Q_{1\mu} \mathbf{y}_i^*}{Y_i^* Q_{1\mu} Y_i^* + \mu^*} \quad \dots \quad (14)$$

where $\xi^* = \xi/\omega$ and $\mu^* = \mu/\omega^2$, see Mehta and Swamy (1978, p. 6).

Theorem : Given the model assumptions, if $G_i = 1$ and $T - K > 4r$, where r is an arbitrary positive integer, then the $2r$ -th moment of the estimator (14) exists.

Proof : Inserting (12) into (14) gives

$$g_i^*(l, \mu) = \frac{(1-l)(\rho Y_i^* M Y_i^* + \xi^* Y_i^* M \mathbf{y}_i^*) + (\rho Y_i^* Q_{1\mu} Y_i^* + \xi^* Y_i^* Q_{1\mu} \mathbf{y}_i^*)}{(1-l)Y_i^* M Y_i^* + Y_i^* Q_{1\mu} Y_i^* + \mu^*} \quad \dots \quad (15)$$

From the inequality in Rao (1973, p. 149, Problem 8(a)) we find that

$$E |g_i^*(l, \mu)|^r \leq C_r (E |R_1|^r + E |R_2|^r) \quad \dots \quad (16)$$

where $C_r = 1$ for $r \leq 1$ and $= 2^{r-1}$ for $r \geq 1$,

$$R_1 = \frac{(1-l)(\rho Y_i^* M Y_i^* + \xi^* X_i^* M \mathbf{y}_i^*)}{(1-l)Y_i^* M Y_i^* + Y_i^* Q_{1\mu} Y_i^* + \mu^*} \quad \dots \quad (17)$$

and

$$R_2 = \frac{(\rho Y_i^* Q_{1\mu} Y_i^* + \xi^* Y_i^* Q_{1\mu} \mathbf{y}_i^*)}{(1-l)Y_i^* M Y_i^* + Y_i^* Q_{1\mu} Y_i^* + \mu^*}. \quad \dots \quad (18)$$

Consequently, the r -th absolute moment of $g_i^*(l, \mu)$ is finite if the r -th absolute moments of R_1 and R_2 are finite. It follows by the inequality (11) that

$$E |R_1|^r \leq E \left| \frac{(1-l)(\rho Y_i^* M Y_i^* + \xi^* Y_i^* M \mathbf{y}_i^*)}{(1-l)Y_i^* M Y_i^* + Y_i^* Q_{1\mu} Y_i^* + \mu^*} \right|^r \quad \dots \quad (19)$$

and

$$E |R_2|^r \leq E \left| \frac{(\rho Y_i^* Q_{1\mu} Y_i^* + \xi^* Y_i^* Q_{1\mu} \mathbf{y}_i^*)}{(1-l)Y_i^* M Y_i^* + Y_i^* Q_{1\mu} Y_i^* + \mu^*} \right|^r \quad \dots \quad (20)$$

The problem of showing the finiteness of the r -th absolute moment of $g_i^*(l, \mu)$ thus resolves into the problem of showing the finiteness of the quantities on the right-hand side of the inequalities (19) and (20). Since

$$k^* = (Y_i^* M Y_i)^{-1} Y_i^* M_i Y_i = (Y_i^{*'} M Y_i^*)^{-1} Y_i^{*'} M_i Y_i^*$$

for $G_i = 1$ (from (10) and (13)), the right-hand side quantities of (19) and (20) can be written, respectively, as

$$E |R_3|^r = E \left| \frac{(1-l)(\rho Y_i^{*'} M Y_i^* + \xi^* Y_i^{*'} M y_i^*)}{Y_i^{*'} [X_i (X_i X_i)^{-1} X_i' - X_i (X_i X_i + \mu I)^{-1} X_i'] Y_i^* + \mu^*} \right|^r \dots (21)$$

and

$$E |R_4|^r = E \left| \frac{(\rho Y_i^{*'} Q_{1\mu} Y_i^* + \xi^* Y_i^{*'} Q_{1\mu} y_i^*)}{Y_i^{*'} [X_i (X_i X_i)^{-1} X_i' - X_i (X_i X_i + \mu I)^{-1} X_i'] Y_i^* + \mu^*} \right|^r \dots (22)$$

It now follows by application of the Cauchy-Schwarz inequality (Lukacs, 1975, p. 12) that

$$(E |R_3|^r)^2 \leq E(1-l)^{2r} E \left[\frac{(\rho Y_i^{*'} M Y_i^* + \xi^* Y_i^{*'} M y_i^*)}{Y_i^{*'} [X_i (X_i X_i)^{-1} X_i' - X_i (X_i X_i + \mu I)^{-1} X_i'] Y_i^* + \mu^*} \right]^{2r} \dots (23)$$

Since $M X_i = 0$ and Y_i^* and y_i^* are independent normal variables, the second factor on the right-hand side of the inequality (23) can be shown to be finite by recognizing that

- (i) $\rho Y_i^{*'} M Y_i^* + \xi^* Y_i^{*'} M y_i^*$ is independent of $Y_i^{*'} [X_i (X_i X_i)^{-1} X_i' - X_i (X_i X_i + \mu I)^{-1} X_i'] Y_i^* + \mu^*$,
- (ii) $E(\rho Y_i^{*'} M Y_i^* + \xi^* Y_i^{*'} M y_i^*)^{2r} < \infty$,
- (iii) $X_i [(X_i X_i)^{-1} - (X_i X_i + \mu I)^{-1}] X_i'$ is positive semidefinite, and
- (iv) $E(Y_i^{*'} [X_i (X_i X_i)^{-1} X_i' - X_i (X_i X_i + \mu I)^{-1} X_i'] Y_i^* + \mu^*)^{-2r} \leq E\mu^{*-2r} < \infty$.

To establish the finiteness of the first factor on the right-hand side of the inequality (23), we find that

$$E(1-l)^{2r} \leq C_r(1+E l^{2r}) \leq C_r(1+E k^{*2r}) \dots (24)$$

where C_r is as defined in (16) and the second inequality is based on the result in (11). Consequently, it suffices to show that

$$E k^{*2r} = E \left[\frac{Y_i^{*'} M_i Y_i^*}{Y_i^{*'} M Y_i^*} \right]^{2r} < \infty \dots (25)$$

Noticing that M and M_i commute, we can write

$$E \left[\frac{Y_i^{*'} M Y_i^*}{Y_i^{*'} M Y_i^*} \right]^{2r} = E \left[1 + \frac{Y_i^{*'} p_2 p_2' Y_i^*}{Y_i^{*'} p_1 p_1' Y_i^*} \right]^{2r} \leq 2^{2r-1} \left[1 + \left(\frac{Y_i^{*'} p_2 p_2' Y_i^*}{Y_i^{*'} p_1 p_1' Y_i^*} \right)^{2r} \right] \dots (26)$$

where the columns of p_1 are the eigenvectors corresponding to $T-K$ unit roots of both M and M_i and the columns of p_2 are the eigenvectors corresponding to $K-K_i$ unit roots of M_i and $K-K_i$ zero roots of M . The variable $Y_i^{*'} p_2 p_2' Y_i^*$ is independent of $Y_i^{*'} p_1 p_1' Y_i^*$ because $p_2 p_2' p_1 p_1' = p_1 p_1' p_2 p_2' = 0$. Further, $Y_i^{*'} p_1 p_1' Y_i^*$ and $Y_i^{*'} p_2 p_2' Y_i^*$ are distributed as noncentral chi-square variables with $T-K$ and $K-K_i$ degrees of freedom respectively. These results show that the result in (25) is true if $T-K > 4r$. It remains now to establish the finiteness of $E |R_4|^r$ in (22). This result can be shown to hold for any $K-K_i \geq 1$ by the use of the same argument as given in the proof of Mehta and Swamy's (1978) theorem. Alternatively, we can show that for any $K-K_i \geq 1$

$$E |R_4|^r \leq \mu^{*-r} E |\rho Y_i^{*'} Q_{1\mu} Y_i^* + \xi^* Y_i^{*'} Q_{1\mu} y_i^*|^r < \infty \dots (27)$$

because $Y_i^{*'} [X_1'(X_i' X_i)^{-1} X_i' - X_i'(X_i' X_i + \mu I)^{-1} X_i'] Y_i^*$ is nonnegative almost surely and

$$E |\rho Y_i^{*'} Q_{1\mu} Y_i^* + \xi^* Y_i^{*'} Q_{1\mu} y_i^*|^r \leq C_r [\rho E(Y_i^{*'} Q_{1\mu} Y_i^*)^r + \xi^{*r} E |Y_i^{*'} Q_{1\mu} y_i^*|^r] < \infty. \text{ Q.E.D.}$$

The conditions of the above theorem also guarantee the existence of the $2r$ -th moments for the estimator (6). To see this, consider the r -th absolute moment of an arbitrary linear combination of $b_i(l, \mu)$, denoted $E |c' b_i(l, \mu)|^r$. Again using the inequality in Rao (1973, p. 149, Problem 8(a)), we obtain that

$$E |c' b_i(l, \mu)|^r \leq C_r (E |c'(X_i' X_i + \mu I)^{-1} X_i' y_i|^r + E |c'(X_i' X_i + \mu I)^{-1} X_i' Y_i g_i(l, \mu)|^r) \dots (28)$$

where it is assumed that $G_i = 1$. The first term on the right-hand side of this inequality is finite because y_i is a normal variable. The second term is also finite because by the Cauchy-Schwarz inequality

$$(E |c'(X_i' X_i + \mu I)^{-1} X_i' Y_i g_i(l, \mu)|^r)^2 \leq E [c'(X_i' X_i + \mu I)^{-1} X_i' Y_i]^{2r} \times E [g_i(l, \mu)]^{2r}. \dots (29)$$

The right-hand side of this inequality is finite under the conditions of our theorem.

To recapitulate, this work establishes that the modified LIML estimators in (5) and (6) possess finite moments for all T greater than some finite number. This is a decided improvement relative to the usual LIML estimators which do not, in general, possess finite means for any $T-K > 0$, see Mariano and Sawa (1972).

We remark that the estimators (5) and (6) do not exist when $T < K$ because they involve the regular inverse of $X'X$. Specifically, the root l is not defined when $T \leq K$ and $\text{rank}(X) = T$. In such cases we can use the Swamy and Holmes (1971) estimator. Further discussion of the Swamy-Holmes estimator is given in Swamy (1979). For general purposes it appears usually preferable to use the estimators (5) and (6) for larger values of $T-K > 0$ and the Swamy-Holmes estimator for smaller values of T . Anderson and Sawa's (1979) observations that (i) the distribution of the LIML estimator approaches normality much faster than that of the 2SLS estimator and (ii) the LIML estimator is generally median unbiased support this conclusion.

We may consider the use of the estimators (5) and (6) in economic forecasting. If we have a set of values available for the exogenous variables for the period in which the forecast is made, say x_f , (where x_f is a $K \times 1$ vector) we can then predict a value denoted \hat{y}_{if} , of the left-hand side endogenous variable in (1) by using the partially restricted reduced form estimators implied by (5) and (6). We then obtain

$$\hat{y}_{if} = x_f' \hat{\pi}_i : \hat{\pi}_i^* \tag{30}$$

where

$$\begin{aligned} \hat{\pi}_i &= \hat{\Pi}_{iu} g_l(l, \mu) + b_l(l, \mu), \\ \hat{\pi}_i^* &= \hat{\Pi}_{iu}^* g_l(l, \mu), \\ \hat{\Pi}_{iu} &= (X_i' M_i^* X_i)^{-1} X_i' M_i^* Y_i, \\ M_i^* &= 1 - X_i^* (X_i^* X_i)^{-1} X_i^*, \end{aligned}$$

and

$$\hat{\Pi}_{iu} (X_i^* X_i)^{-1} X_i^* (Y_i - X_i \hat{\Pi}_{iu}).$$

The argument given in the proof of Theorem 3 of Swamy and Mehta (1979) may be used here to show that under the conditions of our theorem, the forecast (30) possesses finite moments. The conditions of our theorem are sufficient to allow Goldberger, Nagar and Odeh (1961) type approximations to be developed for the first and second-order moments of $\hat{\pi}_i$ and $\hat{\pi}_i^*$.

REFERENCES

- ANDERSON, T. W. (1976): Estimation of linear functional relationships: approximate distributions and connections with simultaneous equations in econometrics. *J. Roy. Statist. Soc, Series B*, 38, 1-30.
- ANDERSON, T. W. and SAWA, T. (1979): Evaluation of the distribution function of the two-stage least squares estimate. *Econometrica*, 47, 163-182.
- GOLDBERGER, A. S., NAGAR, A. L., and ODEH, H. S. (1961): The covariance matrices of reduced-form coefficients and of forecasts for a structural econometric model. *Econometrica*, 19, 556-573.
- KADIYALA, K. R. (1970): An exact small sample property of the k -class estimators. *Econometrica*, 38, 930-932.
- LUKACS, E. (1975): *Stochastic Convergence*, second edition, New York: Academic Press.
- MARIAMO, R. S. and SAWA, T. (1972): The exact finite-sample distribution of the limited information maximum likelihood estimator in the case of two included endogenous variables. *Jour. Amer. Statist. Assoc.*, 67, 159-163.
- MEHTA, J. S. and SWAMY, P. A. V. B. (1978): The existence of moments of some simple Bayes estimators of coefficients in a simultaneous equation model. *Journal of Econometrics*, 7, 1-13.
- NAGAR, A. L. (1959): The bias and moment matrix of the general k -class estimators of the parameters in simultaneous equations. *Econometrica*, 27, 575-595.
- RAO, C. R. (1973): *Linear Statistical Inference and Its Applications*, second edition, New York: John Wiley and Sons.
- SWAMY, P. A. V. B. (1979): A comparison of estimators for undersized samples. *Special Studies Paper*, Federal Reserve Board, Washington, D. C.
- SWAMY, P. A. V. B. and HOLMES, J. (1971): The use of undersized samples in the estimation of simultaneous equation systems. *Econometrica*, 39, 455-459.
- SWAMY, P. A. V. B. and MEHTA, J. S. (1979): On the existence of moments of partially restricted reduced form coefficients. *Special Studies Paper*, Federal Reserve Board, Washington, D.C.
- THEIL, H. (1971): *Principles of Econometrics*, New York: John Wiley and Sons.
- ZELLNER, A. (1970): Estimation of regression relationships containing unobservable variables. *International Economic Review*, 11, 441-454.

Paper received: March, 1981.