Sankhyā : The Indian Journal of Statistics
1984, Volume 46, Series B, Pt. 1, pp., 83–89.

A NOTE ON MODIFIED LIML ESTIMATORS

By S. HARISH
University of Mysore, India
and
N. S. IYENGAR*
Indian Statistical Institute

SUMMARY. In a recent paper Swamy, Mehta and Iyengar (1983) have modified the classical LIML estimator and proved the finiteness of the moments of their estimator. In this note, a further modification is introduced through a simple adjustment. This adjusted LIML estimator not only possesses finite moments but also overcomes certain limitations of the Swamy-Mehta-Iyengar estimator. A simple numerical example is provided to illustrate the various steps involved in the estimation process.

1. Introduction

Mehta-Swamy (1978) and Swamy-Mehta-Iyengar (1983) modified the classical LIML estimators and established the finiteness of the moments of their modified estimators. In this note, we propose further modifications to the Mehta-Swamy (1978) estimator, which may be referred to as the 'adjusted' LIML estimator. These estimators preserve the effect of the adjustment factor (μ) on both the endogenous and the exogenous coefficient estimators. It is shown that the adjusted LIML estimators not only possess finite moments but also appear simpler than the Swamy-Mehta-Iyengar estimator. In addition, the former estimators overcome certain inherent limitations of the latter.

In Section 2, we define our adjusted LIML estimator using the same model as described in Swamy-Mehta-Iyengar (1983) and state its main properties in Section 3. A simple numerical illustration is provided in Section 4 taking a 3-equation macro economic model from Koutsoyiannis (1979, p. 389).

* This note is based on a paper presented at the 21st Indian Econometric conference held at Annamalai Nagar, 10–12 January 1983. The authors are grateful to the anonymous referee for making valuable suggestions, as also to the Computer Centre, University of Mysore for the computer facilities.

AMS (1980) *subject classification :* 62J05, 62F10.

Key words and phrases : Limited information maximum likelihood (LIML) ; Modified LIML : Adjusted LIML ; Adjustment factor ; Ridge regression ; Marginal propensity to consume ; Finite sample moments.

2. The adjusted LIML estimator

We propose the following estimators for the unknown coefficients γ_i and β_i in the Swamy-Mehta-Iyengar (1983) model.

$$g_i(\lambda, \mu) = (Y_i'Q_\lambda Y_i + \mu I)^{-1} Y_i'Q_\lambda y_i \qquad \ldots \quad (1)$$

where
$$Q_\lambda = M_i - \lambda M$$
$$M_i = I - X_i(X_i'X_i)^{-1}X_i'$$
$$M = I - X(X'X)^{-1}X'$$
$\mu \neq 0$ is an arbitrary constant[1]

and
$$b_i(\lambda, \mu) = (X_i'X_i)^{-1}X_i'(y_i - Y_i g_i(\lambda, \mu)). \qquad \ldots \quad (2)$$

In Section 3, we show that these adjusted estimators possess finite sample moments under mild assumptions about the sample size.

It may be noted that these estimators and the results based on them are more general than those obtained for the "modified" LIML estimators by Mehta-Swamy (1978) and Swamy-Mehta-Iyengar (1983), in the sense that the Mehta-Swamy (1978) estimators reduce to our 'adjusted' LIML estimators when $\mu_1 = \mu$ and $\mu_2 = 0$ where μ_1 and μ_2 are respectively the adjustment factors in $g_i(\lambda, \mu)$ and $b_i(\lambda, \mu)$, but they are both assumed equal to μ in their study, as well as in Swamy-Mehta-Iyengar (1983). As can be seen from our analysis, the adjustment factor μ, appearing in $g_i(\lambda, \mu)$, also affects $b_i(\lambda, \mu)$, since the latter is a function of the former. Moreover, this definition of the LIML estimators enables us to prove the existence of finite moments for their sampling distributions even for moderately small sample sizes.

3. Properties of the adjusted LIML estimator

The estimator (1) in reparametrized form can be written as[2]

$$g_i^*(\lambda, \mu) = \frac{\rho Y_i^{*\prime} Q_\lambda Y_i^* + \xi^* Y_i^{*\prime} Q_\lambda y_i^*}{Y_i^{*\prime} Q_\lambda Y_i^* + \mu^*} \qquad \ldots \quad (3)$$

Re-arranging the Q_λ matrix as

$$Q_\lambda = Q_1 + (1-\lambda) M,$$

[1]In Swamy-Mehta-Iyengar (1983) and Mehta-Swamy (1978) $\mu > 0$. In our case it need not be positive. Since the marginal propensity to consume (γ) cannot exceed unity, it is easy to show that μ has the lower limit equal to $Y_i'Q_\lambda(y_i - Y_i)$, which can be positive or negative. Also since β cannot be more than unity, one can show that the upper limit for μ is

$$\frac{X_i' Y_i Y_i' Q_\lambda y_i - Y_i' Q_\lambda Y_i X_i'(y_i - X_i)}{X_i' y_i - X_i' X}.$$

In our case, as already stated, $G_i = 1$ and $K_i = 1$.

[2]The notation $g_i^*(\lambda, \mu)$ has been used to indicate that this makes use of the transformed data y_i^* and Y_i^*.

where

$$Q_1 = I - X_i(X_i' X_i)^{-1} X_i' - M$$

and writing this in equation (3) we get

$$g_i^*(\lambda, \mu) = \frac{(1-\lambda)(\rho Y_i^{*\prime} M Y_i^* + \xi^* Y_i^{*\prime} M y_i^*) + (\rho Y_i^{*\prime} Q_1 Y_i^* + \xi^* Y_i^{*\prime} Q_1 y_i^*)}{(1-\lambda) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* + \mu^*}, \quad \ldots \quad (4)$$

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

$$E|g_i^*(\lambda, \mu)| \leqslant C_r|(E|R_1|^r + E|R_2|^r)$$

where

$$C_r = \begin{cases} 1 & \text{for } r \leqslant 1 \\ 2^{r-1} & \text{for } r \geqslant 1. \end{cases} \quad \ldots \quad (5)$$

Let us write

$$R_1 = \frac{(1-\lambda)(\rho Y_i^{*\prime} M Y_i^* + \xi^* Y_i^{*\prime} M y_i^*)}{(1-\lambda) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* + \mu^*} \quad \ldots \quad (6)$$

and

$$R_2 = \frac{\rho Y_i^{*\prime} Q_1 Y_i^* + \xi^* Y_i^{*\prime} Q_1 y_i^*}{(1-\lambda) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* + \mu^*}. \quad \ldots \quad (7)$$

Consequently, the r-th absolute moment of $g_i^*(\lambda, \mu)$ is finite if the r-th absolute moments of R_1 and R_2 are finite. We know that

$$k^* > \lambda \geqslant 1 \text{ (Kadiyala, 1970)}$$

where

$$k^* = \min_{\mathbf{c} \neq 0} \left\{ \frac{\mathbf{c}' Y_i' M_i Y_i \mathbf{c}}{\mathbf{c}' Y_i' M Y_i \mathbf{c}} \right\}. \quad \ldots \quad (8)$$

From this, it follows that

$$E|R_1|^r \leqslant E\left| \frac{(1-\lambda)(\rho Y_i^{*\prime} M Y_i^* + \xi^* Y_i^{*\prime} M y_i^*)}{(1-k^*) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* + \mu^*} \right|^r \quad \ldots \quad (9)$$

and

$$E|R_2|^r \leqslant E\left| \frac{\rho Y_i^{*\prime} Q_1 Y_i^* + \xi^* Y_i^{*\prime} Q_1 y_i^*}{(1-k^*) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* + \mu^*} \right|^r \quad \ldots \quad (10)$$

Since

$$k^* = (Y_i' M Y_i)^{-1} Y_i' M_i Y_i = (Y_i^{*\prime} M Y_i^*)^{-1} Y_i^{*\prime} M_i Y_i^* \quad \ldots \quad (11)$$

for $G_i = 1$, the denominators in the right-hand side of (9) and (10) get reduced to μ^*,[3] so that we may write

$$E|R_1|^r \leqslant E\left| \frac{(1-\lambda)(\rho Y^{*\prime} M Y_i^* + \xi^* Y_i^{*\prime} M y_i^*)}{\mu^*} \right| \quad \ldots \quad (12)$$

[3]This is because $(1-k^*) Y_i^{*\prime} M Y_i^* + Y_i^{*\prime} Q_1 Y_i^* = 0$. Also, for $G_i = 1$ since Y_i is a $T \times 1$ vector, c in (8) is a scalar. Hence (8) reduces to an equality in (11).

and

$$E|R_2|^r < E\left|\frac{\rho Y_i^{*\prime}Q_1Y_i^*+\xi^*Y_i^{*\prime}Q_1y_i^*}{\mu^*}\right|^r \quad \ldots \quad (13)$$

Now, by Cauchy-Schwarz inequality

$$(E|R_1|^r)^2 < E(1-\lambda)^{2r}.E\left|\frac{\rho Y_i^{*\prime}MY_i^*+\xi^*Y_i^{*\prime}My_i^*}{\mu^*}\right|^{2r}. \quad \ldots \quad (14)$$

Since $MX_i = 0$ and y_i^* and Y_i^* are independent normal variables, the second factor on the right-hand side of (14) is finite because

(i) The numerator is independent of the denominator, since μ^* is a constant; and

(ii) E (numerator) is finite.

Similarly, the finiteness of the first factor on the right-hand side of (14), can be arrived at using the same argument as in Swamy-Mehta-Iyengar (1983).

It remains now to establish the finiteness of $E|R_2|^r$ in (13). This result can be shown to hold good for any $K-K_i > 1$ by using the same argument as given in the proof of Mehta-Swamy's (1978) theorem. Alternatively, we can use the Swamy-Mehta-Iyengar (1983) argument to show that for any $K-K_i > 1$

$$E|R_2|^r \leqslant \mu^{*-r}E|\rho Y_i^{*\prime}Q_1Y_i^*+\xi^*Y_i^{*\prime}Q_1y_i^*|^r < \infty. \quad \ldots \quad (15)$$

This readily follows from the following inequality

$$E|\rho Y_i^{*\prime}Q_1Y_i^*+\xi^*Y_i^{*\prime}Q_1y_i^*|^r < C_r[\rho E|Y_i^{*\prime}Q_1Y_i^*|^r+\xi^*E|Y_i^{*\prime}Q_1y_i^*|^r] < \infty. \quad \ldots \quad (16)$$

Hence, the $2r$-th moment of the estimator $g_i^*(\lambda, \mu)$ exists finitely.

Similarly, it can be shown that the $2r$-th moment of the estimator $b_i^*(\lambda, \mu)$ also exists. To see this, consider the r-th absolute moment of an arbitrary linear combination of $b_i^*(\lambda, \mu)$, denoted by $E|c'b_i^*(\lambda, \mu)|^r$.

Again, using the inequality in Rao (1973, p. 149, Problem 8a) we obtain,

$$E|c'b_i^*(\lambda,\mu)|^r < C_r[E|c'(X_i'X_i)^{-1}X_i'y_i|^r+E|c'(X_i'X_i)^{-1}X_i'Y_ig_i^*(\lambda,\mu)|^r] \quad \ldots \quad (17)$$

where it is assumed that $G_i = 1$.

The first term on the right-hand side of the inequality is finite because y_i is a normal variable. The second term is also finite because, by the Cauchy-Schwarz inequality,

$$[E|c'(X_i'X_i)^{-1}X_i'Y_ig_i^*(\lambda,\mu)|^r]^2 < E|c'(X_i'X_i)^{-1}X_i'Y_i|^{2r}.E|g_i^*(\lambda,\mu)|^{2r}. \quad \ldots \quad (18)$$

The first factor on the right-hand side of (18) is finite because Y_t is a normal variable and the finiteness of the second factor has already been established.

Hence, the $2r$-th moment of the estimator $b_t^*(\lambda, \mu)$ is also finite.

Thus, we have established that the "adjusted" LIML estimators in (1) and (2) possess finite moments for all T (sample size) greater than some finite number. This is decidedly an improvement over the usual LIML estimators which do not, in general, possess any finite moments for $T-K > 0$. See Mariano and Sawa (1972).

4. An illustration

To illustrate the application of our method we use it for the estimation of an overidentified relationship in a typical Keynesian model of income determination given in Koutsoyiannis (1979, p. 389).

$$\begin{aligned} y_{1t} &= \gamma_{11} y_{2t} + \beta_{11} x_{1t} + u_{1t} \\ y_{3t} &= \gamma_{31} y_{2t} + \beta_{21} x_{2t} + u_{2t} \\ y_{2t} &= y_{1t} + y_{3t} + x_{3t} \end{aligned} \quad \ldots \quad (19)$$

where y_{1t} = consumption, x_{1t} = lagged consumption, $(= y_{1,t-1})$, y_{2t} = Income, x_{2t} = lagged income $(= y_{2,t-1})$, y_{3t} = Investment, x_{3t} = Government expenditure.[4] Variables y_{1t}, y_{2t}, and y_{3t} are assumed to be endogenous and x_{3t} exogenous; the lagged variable $y_{1,t-1}$ and $y_{2,t-1}$ are also treated as exogenous and denoted by x_{1t} and x_{2t} respectively. The consumption equation according to the order and rank conditions is overidentified and permits application of our method. We may estimate the parameters of the consumption function by using the adjusted LIML method. For this, we have used British data from 1961 to 1968 as given in the Appendix. These figures have been rounded off to the nearest billion, and have not been adjusted for population changes. The major computational steps are indicated in the Appendix. Using the values of these computations the estimated consumption function which corresponds to the admissible optimal value of μ is obtained as

$$\hat{y}_{1t} = 2{\cdot}639145 + 0{\cdot}54514895 y_{2t} + 0{\cdot}027915776 x_{1t}. \quad \ldots \quad (20)$$

The corresponding minimum $D^2(\mu)$ turns out to be 0·090574362.

This was obtained from the graph showing the squared distance $D^2(\mu)$ against μ. The estimated coefficients have relevant signs and appear plausible in terms of economic theory. The long-term marginal propensity to consume (MPC) turns out to be < 1, as it should be.

---

[4] In terms of the original formulation these variables are the same as

y_t, X_t, Y_t and $(x_{2t}, x_{3t}) = X_t^*$.

5. Concluding remarks

The modifications proposed by Mehta-Swamy (1978) are on similar lines as those proposed earlier by Hoerl and Kennard (1970) for single-equation linear models. In the latter case the optimal value for the adjustment (μ) factor was found by minimizing the MSE of the OLS estimator. However, in the case of the simultaneous equations model that procedure ends up in intractable form and does not seem to be feasible. Hence the choice is entirely arbitrary. However, one way of choosing μ would be to assume that it is proportional to $Y_i'Q_1Y_i$ which is scalar when $G_i = 1$. There could be several other ways of choosing μ, but we have not tried them. Our numerical illustration is perhaps restrictive in the sense that we cannot compare our results with those of Koutsoyiannis, as the number of observations and the units of measurement in the two cases are different.

Appendix I

NATIONAL INCOME ACCOUNTS OF THE U.K.
(£ billion at 1963 prices)

year	consumption (y_{1t})	income (y_{2t})	government expenditure (x_{3t})	lagged income (x_{2t})	lagged consumption (x_{1t})
1961	18.846	29.091	4.945	28.134	18.418
1962	19.258	29.450	5.100	20.091	18.846
1963	20.125	30.705	5.184	29.450	19.258
1964	20.819	32.372	5.272	30.705	20.125
1965	21.169	33.152	5.420	32.372	20.819
1966	21.617	33.764	3.561	33.152	21.169
1967	22.039	34.411	5.825	33.764	21.617
1968	22.562	35.429	5.851	34.411	22.039

Source : Koutsoyiannis, *op. cit.*, p. 390.

Appendix II

Important intermediate results :

(1) $(X_1'X_1)^{-1} = 0{\cdot}079655.$

(2) $(X'X)^{-1}$

$$= \begin{bmatrix} 11{\cdot}036214 & -5{\cdot}973645 & -1{\cdot}956928 \\ & 3{\cdot}803477 & -2{\cdot}902246 \\ & & 29{\cdot}175600 \end{bmatrix}$$

(3) The best root λ of $|Y_0'M_tY_0-\lambda Y_0'MY_0| = 0$ is 1·511421.

(4) Estimators at the admissible optimum $\mu^* = 0{\cdot}05\%$ of $Y_t'Q_\lambda Y_t$ are :

$$\hat{\gamma} = 0{\cdot}545149$$
$$\hat{\beta} = 0{\cdot}027916$$

(5) The minimum distance $D^2(\mu^*) = 0{\cdot}090574$.

References

Hoerl Arthur, E. and Robert, W. Kennard (1970a) : Ridge regression—based estimation for non orthogonal problems. *Technometrics*, 12, No. 1, 55–67.

——— (1970b) : Ridge regression : Applications to non orthogonal problems. *Technometrics*, 12, No. 1, 69–82.

Kadiyala, K. R. (1970) : An exact small sample property of k-class estimators. *Econometrica*, 38, 930–932.

Koutsoyiannis, A. (1979) : *Theory of Econometrics*, 2nd Ed., MacMillan (India).

Lukacs, E. (1975) : *Stochastic Convergence*, 2nd Edition, New York, Academic Press.

Mariano, R. S. and Sawa, T. (1972) : The exact finite sample distribution of the limited information maximum likelihood estimator in the case of two included endogeneous 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.

Rao, C. R. (1973) : *Linear Statistical Inference and its Applications*, 2nd edition, New York, John Wiley & Sons.

Swamy, P. A. V. B., Mehta, J. S. and Iyengar, N. S. (1983) : Finite sample properties of a modification of the limited information maximum likelihood estimator. *Sankhyā*, B, 45, No. 3, pp. 389-397.

Theil, H. (1971) : *Principles of Econometrics*, New York, John Wiley and Sons.

Paper received : July, 1982.
Revised : July, 1983.