RESEARCH PAPER | Economics | Volume : 3 | Issue : 6 | June 2013 | ISSN - 2249-555X

Maximum Likelihood Estimation in the Presence of Errors in Variables: A Modified Approach

KEYWORDS Parameter, Errors in Variables, Maximum Likelihood estimator (M.L.E), Mean Square Error (MSE).

Prof. V. Balakrishnama Naidu	Dr. P. Surya Kumar	Dr. A. Siva Sankar
Professor & Head, Department of Econometrics, S.V.University,Tirupati-517502	Academic Consultant, Department of Econometrics, S. V. University, Tirupati-517 502	Academic Consultant, Department of Econometrics, S.V. University, Tirupati- 517502

ABSTRACT *In general, most of the measurements employed in economic analysis contain sizeable errors of measurement. Any realistic model must take this fact in to consideration. In the presence of measuremental errors, the OLS Method of estimation of parameters of the general linear model breaks down. Maximum Likelihood (M.L) method assumes error variances to be known. Even if it is taken that the error variances are known, the variance covariance matrix of true variables thus obtained using covariance matrix of error variables, need not necessarily be positive definite. This paper discusses how to get the M.L estimator of covariance matrix of true variables which is positive definite matrix and hence obtain the modified M.L estimator of the parameter vector. It is also found that the modified M.L Estimator is better than the OLE from the point of view of bias and Mean Square Error (MSE).*

INTRODUCTION

Let there be a general model with stochastic repressors

$$Y = X \ \beta + \in \qquad 1.1$$

n.1 n.k k.1 n.1

when there are errors in variables, then, y, x_j (j=1, … k) are known as true variables and the observed variables are y_j^* x_j^*, with additive errors v and z.

$$y^* = y + v$$

$$x^* = x + z \qquad 1.2$$

Under the usual assumptions of errors in variables model, there exists an observed model

$$Y^* = X^* \beta^* + \in^* \qquad 1.3$$

Corresponding to (1.1) [(Refer: Lindley (1947)]

For all practical purposes of estimation and testing one has to depend on the observed model, (1.3), since true variables are not known.

It is well known that OLS method breakdown in the sense that it cannot provide even consistent estimators of parameter vector β [Johnston (1963)].

There is voluminous literature on the topic suggesting different methods of estimation of parameters of true model. However, there is no satisfactory solution.

Making use of the correspondence between true model (1.1) and observed model (1.3) a new estimator is provided, which is a linear transform of the OLS, β^*, based on the observed model.

This new estimator is

$$b = \widetilde{M}_x^{-1} \ \hat{M}_{x^*y^*} \qquad 1.4$$

Where, M_{x^*}, M_x, M and $M_{x^*y^*}$ are the variance covariance matrices of x_j^*, x_j, (x_j, z_j) and (x_j^*, y^*) respectively, whose sample estimators are given by

$$\hat{M}_{x^*} = \left(\frac{x^{*1} x^*}{n} \right) \qquad 1.5$$

$$\hat{M}_x = \left(\frac{x^1 \, x}{n} \right)$$

$$\hat{M} = \frac{x^1 z + z^1 x + z^1 z}{n}$$

$$\hat{M}_{x^*y^*} = \left(\frac{x^* \, y^*}{n} \right)$$

and the asymptotic variance covariance matrix of b is obtained as

$$\sigma^2 \ M_x^{-1} \ M_{x^*} \ M_x^{-1} \qquad 1.6$$

[Refer: V. B. Naidu et all 1992].

However, the same estimator can also be obtained without making use of the correspondence between true model and observed model. But there is some definite advantage in assuming the correspondence between the models.

It is clear from the discussion, that $\hat{M}_{x^*}$ and $\hat{M}_{x^*y^*}$ alone can be computed in practice, directly and they are M.L. estimators. If M is known, b is nothing but M.L. estimator of β provided M_x is positive matrix [Refer, Judge, C.J. et al. (1980)].

However, since, $x_j^* = x_j + z_j$, $M_{x^*} = M_x + M$ and hence $\widetilde{M}_x = \widetilde{M}_{x^*} - M$ and it is an unbiased estimator of M_x. But it is not certain that M_x thus obtained is positive definite and much less whether it is M.L. estimator of M_x. More often, in economic situations, M has to be choosen from extraneous information in which case it is not admissible in general as a covariance matrix. And hence, b, can no longer be M.L. estimator. Now the question is can we have M.L. estimator of M_x, which is positive definite? In this article an attempt is made to answer this question.

There are certain procedures developed to provide general method of analysis of covariance structures. In cases analytical solutions are not possible to get maximum likelihood estimators of covariance matrix numerical maximization of likelihood functions are proposed to provide maximum likelihood estimators of covariance matrix, under multivariate normality assumption of explanatory variables, along with imposed conditions on the nature of parameter matrix [Ref: Bock, R.D. and Bergman, R.D. (1966), Fletcher, R and Powell, N.J.D. (1963), Wiley D.W. et al., (1975)].

Volume : 3 | Issue : 6 | June 2013 | ISSN - 2249-555X

The basic problem is, it is given that, $M_{x*} = M_x + M$, where M_{x*}, M_x are positive definite and M is symmetric matrix and further M.L. estimator of M_{x*} is $\hat{M}_{x*}$. Using this information, if M is known, can we arrive at M.L. estimator of M_x, which is positive definite?

RESULT

There exists a matrix P such that $P^1 M_x P = D_2$, a diagonal matrix with positive elements.

Proof: We have $\hat{M}_{x*} = \widetilde{M}_x + M$ (M known)

It is given that $\hat{M}_{x*}$ is positive definite and M is symmetric. Then there is a standard matrix algebra result, which says that, there exists a matrix P such that $P^1 \hat{M}_{x*} P = I$, and

$P^1 M P = D_1$ (diagonal matrix) with $D_{1i} > 0$.

$$\therefore P' \hat{M}_{x*} P = P' \widetilde{M}_x P + P' M P \quad 1.7$$

$$P' \widetilde{M}_x P = P' \hat{M}_x P - P' M P \quad 1.8$$

$$\text{i.e. } D_2 = (I - D_1) \quad 1.9$$

where D_2 is a diagonal matrix. If D_2 is to be positive definite, D_2 must be a diagonal matrix with positive elements i.e.

$D_{2i} = (I - D_{1i}) > 0$ 1.10

which implies that $D_{1i} < 1$, and it is possible to have such D_{1i}. Hence the result.

From (1.8)

Consider, $P' \widetilde{M}_x P = P' \hat{M}_{x*} P - P' M P$

i.e., $\widetilde{M}_x = Q' \ (I - D_1) Q$, 1.11

where $Q = P^{-1}$

From (1.11) we have $\hat{M}_x = Q' \ (I - D_1^*) \ Q$ 1.12

Where $D_1^* = diag \ (D_{1i} \ ; O < D_{1i} < 1)$ is a positive definite and it is M.L. estimator of M_x.

[Ref: Bock R.D. and Vandenberg, S.G. (1968)].

Thus the modified maximum likelihood estimator of β can be obtained by

$$\hat{\beta} = [Q' \ (I - D_1^*) \ Q]^{-1} \ \hat{M}_{x*y*} \quad 1.13$$

In classical M.L. estimation of β vector the variance covariance of the estimator is not given. However, following the correspondence between the true model and observed model, we can have from (1.6) the asymptotic variance covariance matrix of β as

Asy $v \ (\hat{\beta}) = \sigma^2 \ M_x^{-1} \ M_{x*} \ M_x^{-1}$

and an estimator of it is given by

Estimated Asy $v \ (\hat{\beta}) = \hat{\sigma}^2 \ \hat{M}_x^{-1} \ \hat{M}_{x*} \ \hat{M}_x^{-1}$ 1.14

MONTE CARLO STUDY

To study the behaviour of the modified approach of M L procedure in the presence of errors in variables linear model, the following experiment is designed

The model considered as $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \in$ 1.15

Where $y* = y + v$

$x_j^* = x_j^* + z_j$, j = 1, 2, 3...

Where y and x_j and true variables, y*, x_j^* and observed variables with errors v and Z_j. The efficiency of the modified M L estimator in examined for the following case where all the variables follow normal distribution.

$X_j \sim N(0, 1)$

$Z_j \sim N(0, 0.1)$, j = 1, 2, 3 . . . 1.16

and $V \sim N(0, 0.1)$

Further in the above case the following two parameters structures are considered

(i) $\beta_1 = 0.01$, $\beta_2 = 0.3$ $\beta_3 = 0.6$ 1.17

(ii) $\beta_1 = -0.3$, $\beta_2 = 3.0$ $\beta_3 = 6.0$ 1.18

with the parameter structures (1.17), (1.18) and the observation generated on X_j (j=1, 2, 3. . .), the observations on the dependent variables Y are obtained by using the relation (1.15). Errors Z_j and V are introduced into the variables X_j and Y according to the relation (1.16).

50 sets of samples of size 10, 20, 30, 50 and 100 considered in computing the estimator in each of the two different structures.

The modified M L estimator is $\hat{\beta} = [Q' \ (I - D_1^*) \ Q]^{-1} \ \hat{M}_{x*y*}$ 1.19

The measure of efficiency of the new estimator to that of OLE with errors viz,.

$e = \left| \frac{MSE(\hat{\beta}^*)}{MSE(\hat{\beta})} \right|$ is provided.

Table-1, represents measure of efficiency of structure-I and structure-II

TABLE-1
MEASURE OF EFFICIENCY

Sample size	Structure-I	Structure-II
10	10^2 (0.301459)	10^3 (0.532870)
20	10^4 (0.837204)	10^2 (0.247035)
30	10^5 (0.143926)	10^6 (0.159920)
50	10^6 (0.492883)	10^4 (0.648397)
100	10^9 (0.275319)	10^7 (0.394682)

SUMMARY OF THE RESULTS

From the Monte Carlo study the following observations are made.

1. The bias is the modified M L estimation is less than the bias in OLE
2. For the two structures the value of e>1, implying that the modified M L estimator is more efficient than OLE

REFERENCE 1. Bock, R.D. & Bargmann, R.E. (1966), "Analysis of Covariance Structures", Psychometreka, 9, pp.507-534. | 2. Bock, R.D. &Vandenberg, S.G. (1968), "Components of Heritable Variation in Mental Scores, in (Ed) Vandenberg, S.G. Progress in Human Behaviour Genetics", John Hopkins Press, Baltimore, pp.233-260. | 3. Fletcher, R & Powell, M.J.D. (1963), "A Rapidly Converging Descent Method of Minimization", Computer Journal, Vol.6, pp.163-168. | 4. Intriligator, M.D (1980), "Econometric Models, Techniques and Applications", Prentice, Hall of India, New Delhi. | 5. James, A.T. (1973), "The Variance Information Manifold in (Ed) P.R. Krishnaiah, Multivariate Analysis-III, Academic Press, New York, pp.166-167. | 6. Johnston, J. (1963), "Econometric Methods", Mc Graw Hill Book Company, New York, First Edition. | 7. Judge, G.J. et al (1980), "The theory and Practice of Econometrics", John Wiley and Sons, New York. | 8. Lankipalle, K.N,et al (1984), "Adjusted Estimator", Proceedings of the twenty second annual conference of the Indian Econometric Society, held at Bangalore, 1984. | 9. Lankipalle, K.N (1984a), "Econometrics Methods and Data Uncertainties: A Monograph", Submitted to Indian Council of Social Science Research. | 10. Lankipalle, K.N. (1975), "Errors in variables linear models: A Survey", Proceedings of the 15th conference of Indian Econometric society, vol.1, pp.21-72. | 11. Lindley, D.V. (1947), "Regression Lines and Linear Functional Relationships", Journal of Royal Statistical Society B, 9, pp.219-244. | 12. V. B. Naidu and V. K. Reddy, (2006), "Estimation of Parameters in the Presence of Measurement Errors and Missing Data: A New Estimation", Asian-African Journal of Economics and Econometrics, Vol.6, No.2, pp.219-225. | 13. V. B. Naidu et al (1992), "Adjusted Linear Estimation: An application to farm management data", Artha vijnana, Vol.34, No.1, pp.41-56. | 14. V. K. Reddy and V. B. Naidu (2006), "Errors in variables and Auto-Correlation: A Monte Carlo Study", Asian-African Journal of Economics and Econometrics, Vol.6, No.2, pp.227-233. | 15. Wiley, D.W. et al (1975), "Studies of a Class of Covariance Structure Models", Journal of the American Statistical Association, 50, pp.317-323. |