DISCUSSION PAPER NO. 503

ASYMPTOTIC OPTIMALITY OF THE LIMITED INFORMATION MAXIMUM LIKELIHOOD ESTIMATOR IN LARGE ECONOMETRIC MODELS */

by

Naoto Kunitomo **/

May 1979

Revised October 1981

*/This research was partly supported by National Science Foundation Grant SOC 79-13976 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. I am grateful to T.W. Anderson, Takeshi Amemiya, Thomas MaCurdy, N.E. Savin, and James Powell for their comments on the earlier version of this article. I also wish to thank an anonymous referee and the editor of this journal for their encouragement to carry out the revision.

**/Assistant Professor, Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, Illinois 60201, U.S.A.

Abstract

A new asymptotic theory for large econometric models, called large-K_2 asymptotics, is introduced. The limited information maximum likelihood (LIML) estimator is shown to be consistent and attain the minimum asymptotic variance. The two stage least squares (TSLS) estimator is shown to be quite inferior to the LIML estimator when K_2, the number of excluded exogenous variables in the structural equation of interest is large.

1. Introduction

Several methods have been proposed for estimating the coefficients of a single equation in the complete system of simultaneous structural equations: limited information maximum likelihood (LIML), two stage least squares (TSLS), and ordinary least squares (OLS) estimation methods, for instance. Under certain appropriate conditions the first two methods yield consistent and asymptotically efficient estimators; the two sets of estimators normalized by the square root of the sample size have the same limiting normal distributions. Hence, for normative purposes, it has been difficult to choose one among these estimators on the ground of large sample aymptotic theory.

In this paper, we introduce a new asymptotic theory which may give a good approximation and expose differences between alternative estimators in contemporary economy-wide econometric models. In our asymptotic theory, called the large-K_2 asymptotics, the number of excluded exogenous varaibles, K_2, or the degrees of overidentification, L, increases along with the sample size.

In the early stage of macro-econometric model building, one prominent feature is that the size of model is small. For example, the Klein Model [1950] consists of three structural equations and several definitional equations. Another example is the Klein-Goldberger Model [1955] in which K_2's are much larger, but less than 40, and so it can be classified as a medium size model. After these pioneering attempts, econometric models became larger and larger in the 1960's. A possible explanation for this phenomenon in econometrics is the belief among economists that the more they disaggregated models, the more precisely they could predict the real economy. Basing his argument on some practical considerations in economics, Klein [1971] states:

> The typical structure of models is presently undergoing change in well defined directions, with the outcome that the models of the future are sure to be much larger.

As a result, it was not unusual by the late 1960's to find models which consisted of more than fifty equations. In some cases, for example, the Brookings Model [1969] and the Wharton Annual Model [1968], the number of equations was more than 100. It should be remarked that these recent large econometric models have a common property that K_2 (or L) is substantially large in each structural equation.

Returning to the theoretical econometrics, there have been several attempts to develop theories of statistical inference when the model is large. See Maddala [1980] for the details of these literatures. For example, Sargan [1975] examined some asymptotic properties of instrumental variables methods. Dhrymes [1971] and Theil [1971] proposed some modified estimators and proved that their estimators are consistent using regular large sample theory. However, since large sample theory implicitly assumes that K_2 (or L) is sufficiently small relative to the sample size, it is difficult to justify applying large sample theory to economy-wide econometric models. In fact, well-known good properties of the TSLS estimator in large sample theory are no longer valid when K_2 (or L) is very large.

The main purpose of this paper is to investigate the asymptotic properties of alternative estimators, particularly the LIML, TSLS, and OLS estimators, when K_2 is considerably large. We shall answer the question: what is the best estimation method for large-scale econometric models? However, except in Section 4, we will not explore the full information maximum likelihood estimation or the three stage least squares estimation for several reasons. First, it may not be practically feasible in computation when the complete system of simultaneous structural equations is very large. Second, it is difficult to formulate the problem in the statistical terminology since the system includes many nuisance parameters. In brief, the full information methods are interesting subjects but they are beyond the scope of this paper. Then, in the following analysis, our attention will be focusing on the limited information estimation approach.

The major result of this paper is that the LIML estimator has an asymptotically optimum property in the large-K_2 asymptotic sense. The LIML estimator is consistent and it minimizes the asymptotic covariance matrix among the R-class estimators, which includes the k-class estimators, under Assumptions 1-5 in Section 2. On the other hand, the TSLS estimator loses even consistency. This striking result, conjectured by Kunitomo [1980], means that the TSLS estimator is extremely inferior to the LIML estimator when K_2 is large. Therefore, the asymptotic theory in our sense throws a new light upon the choice of estimator of the parameters of a single structural equation in large econometric models.

Another result is that the LIML estimator is inadmissible in terms of the "higher order" asymptotic mean squared error (AMSE), which is defined to be the MSE obtained from the asymptotic expansion of the distribution of estimator when K_2 is large. We will examine the modifications of the LIML estimation method proposed by Morimune [1978] (the M estimator) and Fuller [1977] (the F estimator) and show that their estimators improve upon the LIML method in the large-K_2 asymptotics.

In Section 2 we define the model and estimators. The main results in the large-K_2 asymptotic theory are presented in Section 3. Then Section 4 discusses possible interpretations of our results in the complete simultaneous equation system, and gives some numerical evidence to justify our asymptotic theory.

Some concluding remarks are given in Section 5. The proof of theorems are in Section 6.

2. Model and Estimators

We consider a single structural equation represented by

$$y_1 = Y_2\beta + Z_1\gamma + u \quad , \tag{2.1}$$

where y_1 and Y_2 are $T \times 1$ and $T \times G_1$ matrices of T observations on the endogenous variables, Z_1 is a $T \times K_1$ matrix of T observations on the K_1 exogenous variables, β and γ are column vectors of G_1 and K_1 parameters, and u is a $T \times 1$ column vector of unobservable disturbances. The reduced form of the system of structural equations is defined as

$$Y = Z\Pi + V \quad , \tag{2.2}$$

where $Y = (y_1 Y_2)$, $Z = (Z_1 Z_2)$ is a $T \times K$ $(K = K_1 + K_2)$ matrix of exogenous variables, Π is a $K \times (1 + G_1)$ matrix of the reduced form coefficients, and $V = (v_1 V_2)$ is a $T \times (1 + G_1)$ matrix of unobservable disturbances. We make the following conventional assumptions.

Assumption 1: The rows of V are independently normally distributed, each row having mean 0 and nonsingular covariance matrix

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{pmatrix} . \tag{2.3}$$

Assumption 2: The matrix $\underset{\sim}{Z}$ is of rank K and $q = T - K > 0$. In order to relate (2.1) and (2.2), postmultiplying (2.2) by $(1 - \beta')'$, we obtain $\underset{\sim}{u} = \underset{\sim}{v}_1 - \underset{\sim}{V}_2\beta$, $\underset{\sim}{\gamma} = \underset{\sim}{\pi}_{11} - \Pi_{12}\underset{\sim}{\beta}$ and

$$(2.4) \qquad \underset{\sim}{\pi}_{21} = \underset{\sim}{\Pi}_{22}\underset{\sim}{\beta} \ ,$$

where

$$(2.5) \qquad \underset{\sim}{\Pi} = (\underset{\sim}{\pi}_1 \underset{\sim}{\Pi}_2) = \begin{pmatrix} \underset{\sim}{\pi}_{11} & \underset{\sim}{\Pi}_{12} \\ \underset{\sim}{\pi}_{21} & \underset{\sim}{\Pi}_{22} \end{pmatrix}$$

is partitioned into K_1 and K_2 rows, and into 1 and G_1 columns, respectively.

Assumption 3: The submatrix $(\underset{\sim}{\pi}_{21}\underset{\sim}{\Pi}_{22})$ is of rank G_1 and $\underset{\sim}{\Pi}_{22}$ is also of rank G_1.

This assumption implies that

$$(2.6) \qquad L = K_2 - G_1 \geq 0 \ ,$$

where L is the degree of overidentification of the structural equation. The components of $\underset{\sim}{u}$ are independently normally distributed with mean 0 and variance

$$(2.7) \qquad \sigma^2 = \omega_{11} - 2\underset{\sim}{\beta}'\underset{\sim}{\omega}_{21} + \underset{\sim}{\beta}'\underset{\sim}{\Omega}_{22}\underset{\sim}{\beta} \ .$$

Let $\underset{\sim}{p}_{21}$ and $\underset{\sim}{P}_{22}$ be the least squares estimator for $\underset{\sim}{\pi}_{21}$ and $\underset{\sim}{\Pi}_{22}$, and

$$(2.8) \qquad \underset{\sim}{G} = \underset{\sim}{Y}'(\bar{P}_{\underset{\sim}{Z}_1} - \bar{P}_{\underset{\sim}{Z}})\underset{\sim}{Y} = \begin{pmatrix} \underset{\sim}{p}'_{21} \\ \underset{\sim}{P}'_{22} \end{pmatrix} \underset{\sim}{A}_{22.1} (\underset{\sim}{p}_{21} \underset{\sim}{P}_{22})$$

$$= \begin{pmatrix} g_{11} & \underset{\sim}{g}_{12} \\ \underset{\sim}{g}_{21} & \underset{\sim}{G}_{22} \end{pmatrix} ,$$

where $\underset{\sim}{A}_{22.1} = \underset{\sim}{Z}_2'\underset{\sim}{Z}_2 - \underset{\sim}{Z}_2'\underset{\sim}{Z}_1(\underset{\sim}{Z}_1'\underset{\sim}{Z}_1)^{-1} \underset{\sim}{Z}_1'\underset{\sim}{Z}_2$. And also let

$$(2.9) \qquad \underset{\sim}{C} = \underset{\sim}{Y}'\bar{P}_{\underset{\sim}{Z}}\underset{\sim}{Y} = \begin{pmatrix} c_{11} & \underset{\sim}{c}_{12} \\ \underset{\sim}{c}_{21} & \underset{\sim}{C}_{22} \end{pmatrix} ,$$

where, for any matrix $\underset{\sim}{S}$, $P_{\underset{\sim}{S}} = \underset{\sim}{I} - \underset{\sim}{S}(\underset{\sim}{S}'\underset{\sim}{S})^{-1}\underset{\sim}{S}'$ is the projection onto the space orthogonal to the column vectors of $\underset{\sim}{S}$. The limited information maximum likelihood (LIML) estimator of $\underset{\sim}{\beta}$ is

$$(2.10) \qquad \hat{\underset{\sim}{\beta}}_{LI} = -\frac{\underset{\sim}{b}_2}{b_1} ,$$

where $\underset{\sim}{b}' = (b_1 \underset{\sim}{b}_2')$ is the characteristic vector corresponding to the smallest characteristic root, λ_{min}, of

$$(2.11) \qquad |\underset{\sim}{G} - \lambda \underset{\sim}{C}| = 0 .$$

For convenience we define the R-class estimator which includes the (fixed) k-class estimator and the LIML estimator:

$$\hat{\beta}_R = (G_{22} - \lambda^* C_{22})^{-1}(g_{21} - \lambda^* c_{21}) \quad , \tag{2.12}$$

and

$$\lambda^* = a\lambda_{min} + b \quad , \tag{2.13}$$

where a and b are constants. Then it is clear that this estimator is identical to the LIML estimator for $a = 1$ and $b = 0$, and is identical to the (fixed) k-class estimator for $a = 0$ and $b = k - 1$. In particular, it is equal to the two stage least squares (TSLS) estimator for $a = b = 0$, and is equal to the ordinary least squares (OLS) estimator for $a = 0$ and $b = -1$. It is also identical to a modification of the LIML estimator by Fuller [1977] for $a = 1$ and $b = -c/(T - K)$ where c is some constant. The R-class estimator includes a broad class of estimator proposed so far, so that we will confine ourselves to focusing on its asymptotic properties.

The estimator of the coefficients of included exogenous variables γ in any method discussed here is

$$\hat{\gamma} = (Z_1'Z_1)^{-1}Z_1'(y_1 - Y_2\hat{\beta}) \quad , \tag{2.14}$$

where $\hat{\beta}$ is an estimator of β.

The matrix C has a central Wishart distribution with q degrees of freedom and covariance matrix Ω. And the matrix G has a non-central Wishart distribution with K_2 degrees of freedom, covariance matrix $\underset{\sim}{\Omega}$, and the noncentrality matrix:

$$(2.15)\qquad \underset{\sim}{\Theta} = \begin{pmatrix} \underset{\sim}{\pi}'_{21} \\ \underset{\sim}{\Pi}'_{22} \end{pmatrix} \underset{\sim}{A}_{22.1}(\underset{\sim}{\pi}_{21}\underset{\sim}{\Pi}_{22}) = \begin{pmatrix} \underset{\sim}{\beta}' \\ \underset{\sim}{I}_{G_1} \end{pmatrix} \underset{\sim}{\Pi}'_{22}\underset{\sim}{A}_{22.1}\underset{\sim}{\Pi}_{22}(\underset{\sim}{\beta}\underset{\sim}{I}_{G_1}) .$$

The exact distributions and asymptotic expansions of distributions of estimators follow only from the distributions of two independent random matrices $\underset{\sim}{C}$ and $\underset{\sim}{G}$.

Let

$$(2.16)\qquad \underset{\sim}{A}(\Omega) = \frac{\underset{\sim}{\Omega}_{22}^{-\frac{1}{2}}\, \underset{\sim}{\Pi}'_{22}\underset{\sim}{A}_{22.1}\underset{\sim}{\Pi}_{22}\underset{\sim}{\Omega}_{22}^{-\frac{1}{2}}}{\mu^2} ,$$

where

$$(2.17)\qquad \mu^2 = \operatorname{tr}\ \underset{\sim}{\Omega}_{22}^{-\frac{1}{2}}\,(\underset{\sim}{I} + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')\ \underset{\sim}{\Pi}'_{22}\underset{\sim}{A}_{22.1}\underset{\sim}{\Pi}_{22}\underset{\sim}{\Omega}_{22}^{-\frac{1}{2}} ,$$

which is the sum of nonzero characteristic roots of the population equation $|\underset{\sim}{\theta} - \lambda\underset{\sim}{\Omega}| = 0$ and is called the generalized noncentrality parameter,

$$(2.18)\qquad \underset{\sim}{\alpha} = \frac{1}{\sqrt{\omega_{11.2}}}\,\underset{\sim}{\Omega}_{22}^{\frac{1}{2}}\,(\underset{\sim}{\beta} - \underset{\sim}{\Omega}_{22}^{-1}\underset{\sim}{\omega}_{21}) ,$$

which is the standardized coefficent and measures the difference between a structural parameter and a regression coefficient among disturbances, and

(2.19) $$\omega_{11.2} = \omega_{11} - \underset{\sim}{\omega}_{12}\underset{\sim}{\Omega}_{22}^{-1}\underset{\sim}{\omega}_{21} \quad .$$

Now in order to assure that the limiting distributions of estimators are proper distributions, we make the following assumption.

Assumption 4: There exists a nonsingular matrix $\underset{\sim}{A}$ such that

(2.20) $$\lim_{\mu\to\infty} \underset{\sim}{A}(\Omega) = \underset{\sim}{A} \quad .$$

This assumption implies that the ratio of the smallest to the largest characteristic root of $A(\Omega)$ is bounded uniformly in the noncentrality parameter μ^2. Since $\underset{\sim}{A}$ is positive definite, we can find a nonsingular matrix B such that $B'B = A$. Then we define the parameters which appear in the following arguments:

(2.21) $$\underset{\sim}{F}^* = \underset{\sim}{B}(I + \underset{\sim}{\alpha}\underset{\sim}{\alpha}')^{-\frac{1}{2}} ,$$

and

(2.22) $$\underset{\sim}{f}^* = - \underset{\sim}{B}(1 + \underset{\sim}{\alpha}'\underset{\sim}{\alpha})^{-\frac{1}{2}}\underset{\sim}{\alpha} \; .$$

Finally, it is convenient to derive some properties of estimators in terms of the standardized estimator:

$$\hat{e} = \begin{pmatrix} \hat{e}_\beta \\ \hat{e}_\gamma \end{pmatrix} = R \begin{pmatrix} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{pmatrix} , \tag{2.23}$$

where

$$R = \frac{1}{\sigma} \begin{pmatrix} (\Pi'_{22} A_{22.1} \Pi_{22})^{1/2} & 0' \\ (Z_1' Z_1)^{-1} Z_1' Z \Pi_2 & (Z_1' Z_1)^{\frac{1}{2}} \end{pmatrix} , \tag{2.24}$$

and $\hat{e}$ is divided into the first G_1 and the last K_1 elements. We shall denote $\hat{e}$ with the LIML estimator, for example, as $\hat{e}_{LI}$.

3.3 Statement of Main Results

In the regular large sample asymptotic theory (referred to as the large-T asymptotics) for simultaneous equations system, the sample size increases under the assumption that the noncentrality parameter increases also. For this parameter sequence, it is well known that the LIML and TSLS estimators are asymptotically equivalent: the two sets of estimators are consistent and the estimators normalized by the square root of the sample size have the same limiting joint normal distributions. Both are best asymptotically normal (BAN) estimators. However, in the large-K_2 asymptotics, this is not the case if we assume the following:

Assumption 5: There exist finite positive numbers δ and ν^2 such that

$$\lim_{n\to+\infty} \frac{\mu^2}{n} = \delta \quad , \quad \lim_{q\to\infty} \frac{\mu^2}{q} = \nu^2 \tag{3.1}$$

where $n = K_2$.

We call δ the reduced noncentrality parameter. We note that under Assumptions 4 and 5, K_2 and μ^2 increase at the same rate as the sample size.

Lemma 1: Under Assumptions 1-5, an estimator in the R-class estimator is consistent if and only if

$$(a - 1) \frac{\nu^2}{\delta} + b = 0. \tag{3.2}$$

Then

Theorem 1: The LIML estimator is consistent while the TSLS and OLS estimators are inconsistent in the large-K_2 asymptotics.

The theorem shows that the TSLS estimator is quite inferior to the LIML estimator in contemporary large econometric models. Kadane [1971] pointed out the inferiority of the TSLS estimator in large models using the asymptotic mean squared error (AMSE) in the small-σ asymptotics, in which the variance of the distrubances in the model is sufficienctly small.

Lemma 2: Under Assumptions 1-5, the asymptotic covariance matrix of the standardized R-class estimator is given by

$$AV(\hat{\underset{\sim}{e}}_R) = \underset{\sim}{Q} + \frac{2}{\delta}\left(1 + \frac{\nu^2}{\delta}\right)(1 - a)^2 \begin{pmatrix} \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} & \underset{\sim}{0}' \\ \underset{\sim}{0} & \underset{\sim}{0} \end{pmatrix}, \tag{3.3}$$

where

(3.4) $$\underset{\sim}{Q} = \begin{pmatrix} \underset{\sim}{I}_{G_1} + \frac{1}{\delta}(1 + \frac{\nu^2}{\delta}) \underset{\sim}{F}^* \underset{\sim}{F}^{*\prime} & \underset{\sim}{0}' \\ \underset{\sim}{0} & \underset{\sim}{I}_{K_1} \end{pmatrix} .$$

In order to compare matrices, we adopt a strong criterion: $A \geq B$ if and only if $A - B$ is non-negative definite. Then we can assert that the LIML estimator is asymptotically efficient.

Theorem 2: The LIML estimator attains the smallest asymptotic covariance matrix among the R-class estimators, and

(3.5) $$\hat{\underset{\sim}{e}}_{LI} \xrightarrow{d} N[\underset{\sim}{0}, \underset{\sim}{Q}] \quad .$$

The first part of the theorem follows directly from Lemma 2 and the proof of asymptotic normality is given in Lemma 4. We should emphasize that in the large-K_2 asymptotics the number of the unknown parameters increases at the same rate as the sample size. Therefore, the regular asymptotic properties of the maximum likelihood estimator, such as Cramér-Rao lower bound, do not hold in general. However, we proved that the LIML estimator is asymptotically efficient in terms of the asymptotic covariance. It is known that

(3.6) $$\hat{\underset{\sim}{e}}_{LI} \xrightarrow{d} N[0, \underset{\sim}{I}_{G_1} + K_1]$$

for the large-T asymptotics. Thus the asymptotic variance of the LIML estimator in the large-K_2 asymptotics is always larger than that in the large-T asymptotics.

In recent years, some econometricians have been trying to find more asymptotically efficient estimators than the LIML estimator. In this context, Theorem 2 implies that it is sufficient for us to focus on modifying the LIML estimation in order to construct a "higher order" asymptotically efficient estimator.

Morimune [1978] proposed a combined estimator: (the M estimator)

$$\hat{\beta}_M = \frac{L-1}{L}\hat{\beta}_{LI} + \frac{1}{L}\hat{\beta}_{TS} \tag{3.7}$$

and showed that the AMSE of (3.7) is always smaller than the AMSE of the LIML estimator. His arguments were based on the asymptotic moments in the small-σ asymptotics. Later, Morimune and Kunitomo [1980] proved that the ASME of the M estimator is always smaller than that of the LIML estimator in the large-K_2 asymptotics under the assumptions of known covariance and two endogenous variables.

On the other hand, Fuller [1977] modified the LIML estimation to the F estimator:

$$\hat{\beta}_F = (G_{22} - \lambda^* C_{22})^{-1}(g_{21} - \lambda^* c_{21}) \quad , \tag{3.8}$$

and

$$\lambda^* = \lambda_{min} - \frac{1}{q} \quad . \tag{3.9}$$

Both the M estimator and the F estimator are asymptotically mean-unbiased up to T^{-1} in the large-T asymptotics. A similar result can be obtained in the large-K_2 asymptotics.

Theorem 3: Under Assumptions 1-5,

$$\| AE(\hat{e}_M) \| < \| AE(\hat{e}_{LI}) \| \quad , \tag{3.10}$$

$$AE(\hat{e}_M \hat{e}_M') < AE(\hat{e}_{LI} \hat{e}_{LI}') \quad , \tag{3.11}$$

and

$$\| AE(\hat{e}_F) \| < \| AE(\hat{e}_{LI}) \| \quad , \tag{3.12}$$

$$AE(\hat{e}_F e_F') < AE(\hat{e}_{LI} e_{LI}') \quad , \tag{3.13}$$

where $\| \cdot \|$ is the Euclidean norm and AE(·) stands for the moments of approximate distribution of estimators.

The proof follows immediately from Lemmas 7 and 3. This theorem means that we can construct more asymptotically efficient estimators by modifying the LIML estimation method, so that the LIML estimator is inadmissible in the sense of AMSE. Here it should be remarked that both $\hat{\beta}_M$ and $\hat{\beta}_F$ are not asymptotically mean-unbiased in the large-K_2 asymptotics while they are so in the large-T asymptotics.

Unfortunately, we cannot compare the AMSE of these two estimators in the general case. Since the F estimator has exact finite moments while the M estimator does not, however, one may assert that the former is preferable to the latter.

Takeuchi [1978], and Takeuchi and Morimune [1979] proposed a third order asymptotically efficient estimator (the T estimator), which turned out to be asymptotically equivalent to the F estimator. Rothenberg [1978] argues that the F estimator minimizes the AMSE among the R-class estimators. Both of these results were shown within the framework of the large-T asymptotics. Kunitomo [1981] shows that the F estimator is asymptotically efficient in terms of not only the AMSE, but also the concentration of probability.

In the large-K_2 asymptotics, however, neither the F estimator nor the M estimator is third-order asymptotically efficient because they are asymptotically mean-biased in the term of $O(n^{-1})$. [See (6.68) and (6.83) below.] In this respect some ambiguity still remains on the optimality of modified estimation. Further investigations on the higher-order asymptotic efficiency of estimator in the large-K_2 asymptotics are undertaken in Kunitomo [1981].

4. Discussion

4.1 Interpretations of the large-K_2 asymptotic theory

Let the multivariate linear regression model including the structural equation of interest be

$$(4.1) \qquad \underset{\sim}{Y}^* = \underset{\sim}{Z}\underset{\sim}{\Pi}^*(T) + \underset{\sim}{V}^*$$

where $\underset{\sim}{Y}^*$ and $\underset{\sim}{Z}$ are T x G(T) and T x K(T) matrices of observations on regressands and regressors, respectively, $\underset{\sim}{\Pi}^*(T)$ is a K(T) x G(T) matrix of unknown parameters including $\underset{\sim}{\Pi}$, V* is a T x G(T) matrix of disturbances including $\underset{\sim}{V}$, and both $G(T)(>1+G_1)$ and $K(T)(> K)$ are monotone increasing functions of the sample size T. Let us assume (4.1) is the reduced form of the complete system of structural equations:

$$(4.2) \qquad \underset{\sim}{Y}^*\underset{\sim}{B} + \underset{\sim}{Z}\underset{\sim}{\Gamma} = \underset{\sim}{U} \quad ,$$

where $\underset{\sim}{B} = (b_{ij})$ and $\underset{\sim}{\Gamma} = (\gamma_{ij})$ are G(T) x G(T) and K(T) x G(T) matrices of unknown parameters, and $\underset{\sim}{U}$ is a T x G(T) matrix of disturbances.

There are several ways to interpret our large-K_2 asymptotic theory in the complete system of simultaneous equations. One simple way is to fix the number of equations and consider a parameter sequence such that $K_2(T)$ ($= K(T) - K_1$) goes to infinity. Since, in this case, the included exogenous variables in the structural equation of interest are fixed, the number of exogenous variables in other structural equations is growing along with the sample size. Another interpretation, which may be natural in many large systems, is based on the assumptions given by Sargan [1975] that the system of simultaneous equations is almost recursive and stable. Under his Assumptions 1-3, (5) in Sargan [1975] implies that the matrix $\underset{\sim}{\Pi}'_{22}(\frac{1}{T}\underset{\sim}{A}_{22.1})\underset{\sim}{\Pi}_{22}$ converges to a constant matrix. Then, his assumptions together with our Assumption 3 are sufficient conditions for Assumption 5 and Assumption 4. Hence our formulation of the model and assumption can be justified by considering the complete system of simultaneous equations.

However, there is a basic difference between our model and Sargan's model which leads to different results: Theorem 3 in Sargan [1975] implies that the TSLS estimator is consistent. The key fact is his Assumption 5 that the speed of increase in T is much faster than the order of K(T), that is, $\lim_{T\to\infty} K(T)/T = 0$ for fixed K_1. On the other hand, we assumed $0 < \lim_{q\to\infty} K_2/q = \nu^2/\delta < +\infty$ in Assumption 5. Hence, it is not surprising to find a different result using different sets of assumptions about the relative speed of key parameters. It should be also remarked that a similar theorem such as our Theorem 2 can be obtainable under the assumption that $\lim_{T\to+\infty} K(T)/T = 0$. It is also the case that the LIML estimator attains a possible lower bound, and it is efficient. (See Section 6.3 in Kunitomo [1981], for instance.)

4.2 Numerical Comparison of Densities

Anderson and Sawa [1979] compared the exact distribution of the TSLS estimator and the LIML estimator when the covariance is known (which is termed the LIMLK estimator). They concluded that the exact distribution of the TSLS estimator is extremely biased when K_2 is relatively large. Another study of the exact distribution of the LIML estimator by Anderson, Kunitomo, and Sawa [1981] led us to the same conclusion.

In the Appendix, we give nine figures of the estimated density functions of the OLS, TSLS, and LIML estimators with different values of the key parameters: T-K, K_2,α, and μ^2. The estimators are standardized by the asymptotic standard deviation in the large-T asymptotics; this makes it possible to compare the small sample properties of estimators in a systematic way. The estimation method is based on simulation. The procedure of our Monte Carlo experiments and the accuracy of estimation are discussed in Anderson, Kunitomo, and Sawa [1981] in detail.

Figures 1-3 are for the case T - K = 5. When the degrees of freedom T-K is relatively small, there is not much difference between the distributions of the OLS and TSLS estimators. It is reasonable because these estimators are identical in the extreme case of T - K = 0. The OLS and TSLS estimators are biased toward their negative values; the distribution of the LIML estimator is almost median-unbiased. As K_2 increases, the biases of the OLS and TSLS estimators increase, and the dispersion of the LIML estimator increases.

As for Figures 4-6, the TSLS estimator is substantially different from the OLS estimator because T-K is not small (=20). The OLS estimator is always extremely downward biased (because α is positive) and the density of the LIML estimator is very close to the standard normal. The distribution of the TSLS estimator is always between the distribution of the OLS and LIML estimators. As K_2 increases, the bias of the TSLS estimator increases; the TSLS and OLS estimators tend to share a similar bias.

As for Figure 7-9, the degrees of freedom T-K is fairly large (=40). Especially in Figure 7 the density of the TSLS estimator is similar to that of the LIML estimator: the bias of the TSLS is relatively small. This is reasonable since the large-T asymptotics may be appropriate in this case. Again, however, the bias of the TSLS estimator becomes serious as K_2 increases. The distribution of the OLS estimator is extremely biased; the distribution of the LIML estimator is always median-unbiased and is approximated well by the standard normal distribution.

5. Concluding Remarks

We initiated the large-K_2 asymptotics, which may explain the asymptotic properties of alternative estimators when the model is considerably large. Both the TSLS and OLS estimators lose even consistency when the model is large, while the LIML estimator is consistent as is seen in Table 1. In this sense, the LIML estimator has a robust property.

Furthermore, the LIML estimator attains a possible lower bound among a class of consistent estimators, and its limiting distribution is the joint normal distribution. Therefore, the LIML estimator is asymptotically efficient in the large-K_2 asymptotics.

According to our numerical analysis, the difference between consistent estimators (the LIML estimator and its modification) and inconsistent estimators (the TSLS, OLS estimators, for instance) is significant. On the other hand, the differences among the modifications of the LIML estimator are negligible in many cases for practical purposes. (See Tables in Kunitomo [1981].) On the whole, all of our results support the practical advantages of the LIML estimation in large-scale models under the assumption of normality. It can be conjectured that the comparisons of distributions are approximately valid if the distributions of the disturbances are not too far from normal.

Although it has been well known that the OLS estimator in simultaneous eqauations models is inconsistent in the large-T asymptotics, it has been us-usually used in practice by some applied econometricians. A common justification for this is the observation that the TSLS and OLS estimation methods give almost identical estimated values in large-scale models, and hence the OLS estimation is preferred because it is the least complicated in computation.

In the context of the large-K_2 asymptotics, this observation can be easily explained by the fact that both the TSLS and OLS estimators are inconsistent, and their probability limits are almost the same when the model is large. Therefore, it can be neither justifiable nor preferable to use the OLS estimator in large-scale models.

Table 1

Consistency Of Estimators Under Alternative Asymptotic Theories

	Small-σ	Large-T	Large-K_2
OLS	Consistent	Inconsistent	Inconsistent
TSLS	Consistent	Consistent	Inconsistent
LIML	Consistent	Consistent	Consistent

6. Proof of Theorems

In this section we shall give algebraic details. The method we shall use is an extension of Fujikoshi et al. [1979] and Kunitomo et al. [1980]. First we will give a reduction for the distribution of $\underset{\sim}{C}$ and $\underset{\sim}{G}$ in closed form. Let us define a $G_1 \times G_1$ matrix

$$(6.1) \qquad \underset{\sim}{\Gamma} = \underset{\sim}{I}_{G_1} + \underset{\sim}{\alpha}\underset{\sim}{\alpha}' \quad .$$

and a $(1 + G_1) \times (1 + G_1)$ matrix

$$(6.2) \qquad \underset{\sim}{Q}^* = \begin{bmatrix} 1/\sigma & -(1/\sigma)\underset{\sim}{\beta}' \\ (1/\sqrt{\omega_{11.2}})^{-1/2}\ \underset{\sim}{\alpha}\ \underset{\sim}{\Gamma}^{-1/2} & [\underset{\sim}{I}_{G_1} - (1/\sqrt{\omega_{11.2}})\underset{\sim}{\alpha}\underset{\sim}{\omega}_{12}\underset{\sim}{\Omega}_{22}^{-1/2}]\ \underset{\sim}{\Omega}_{22}^{-1/2} \end{bmatrix} .$$

Then

$$(6.3) \qquad \underset{\sim}{Q}^*\ \underset{\sim}{\Omega}\ \underset{\sim}{Q}^{*\prime} = \underset{\sim}{I}_{G_1 + 1} .$$

Let

$$(6.4) \qquad \underset{\sim}{\Theta}^* = \begin{bmatrix} 0 & \underset{\sim}{0}' \\ \underset{\sim}{0} & \mu_2(\underset{\sim}{F}^{*\prime}\underset{\sim}{F}^*)^{-1} \end{bmatrix} ,$$

$$(6.5) \qquad \underset{\sim}{C}^* = \underset{\sim}{Q}^*\ \underset{\sim}{C}\ \underset{\sim}{Q}^{*\prime}$$

$$(6.6) \qquad \underset{\sim}{G}^* = \underset{\sim}{Q}^*\ \underset{\sim}{G}\ \underset{\sim}{Q}^{*\prime}$$

Then $\underset{\sim}{C}^*$ and $\underset{\sim}{G}^*$ are independently distributed as a central Wishart distribution $W_{G_1 + 1}(T-K, \underset{\sim}{I}_{G_1 + 1})$ and a noncentral Wishart distribution $W_{G_1 + 1}(K_2, I_{G_1 + 1}; \underset{\sim}{\Theta}^*)$, respectively. We shall use the following expression for G*:

$$(6.7)\qquad G^* = \mu^2 \begin{pmatrix} 0 & 0' \\ 0 & (\underset{\sim}{F}^{*\prime}\underset{\sim}{F}^*)^{-1} \end{pmatrix}$$

$$+ \mu \begin{pmatrix} 0 & \underset{\sim}{x}_1' F^{*\prime -1} \\ \underset{\sim}{F}^{*-1}\underset{\sim}{x}_1 & \underset{\sim}{X}_1'\underset{\sim}{F}^{*\prime -1} + \underset{\sim}{F}^{*-1}\underset{\sim}{X}_1 \end{pmatrix}$$

$$+ \begin{pmatrix} \underset{\sim}{x}'\underset{\sim}{x} & \underset{\sim}{x}'\underset{\sim}{X} \\ \underset{\sim}{X}'\underset{\sim}{x} & \underset{\sim}{X}'\underset{\sim}{X} \end{pmatrix} ,$$

where the elements of $\underset{\sim}{x}' = (\underset{\sim}{x}_1' \ \underset{\sim}{x}_2') : 1 \times (G_1 + L)$ and $\underset{\sim}{X}' = (\underset{\sim}{X}_1' \ \underset{\sim}{X}_2') : G_1 \times (G_1 + L)$ are independent standard normal variables.

Next we consider a reduction for the distribution of the standardized estimator of $\underset{\sim}{\beta}$ and $\underset{\sim}{\gamma}$ defined by (2.23). After some calculation, the last K_1 elements of $\hat{\underset{\sim}{e}}$ are

(6.8) $$\hat{e}_{\gamma} = \frac{1}{\sigma} (Z_1'Z_1)^{-1/2} Z_1'u - \frac{1}{\mu} (Z_1'Z_1)^{-1/2} Z_1'V_2\Theta_{22}^{-1/2} \hat{e}_{\beta}$$

$$= u^* - \frac{1}{\mu} (MF^{*\prime} + u^*f^{*\prime}) \hat{e}_{\beta} \quad ,$$

where

(6.9) $$u^* = \sigma^{-1}(Z_1'Z_1)^{-1/2}Z_1'u \quad ,$$

(6.10) $$M = (Z_1'Z_1)^{-1/2}Z_1'(V_2\Omega_{22}^{-1/2} + (\sigma\sqrt{c})^{-1}u\alpha')\Gamma^{1/2} \quad ,$$

(6.11) $$f^* = -(1/\sqrt{c})\Theta_{22}^{-1/2}\Omega_{22}^{1/2}\alpha \quad ,$$

(6.12) $$\Theta_{22} = \Pi_{22}'A_{22.1}\Pi_{22}/\mu^2 \quad ,$$

(6.13) $$c = 1 + \alpha'\alpha = \sigma^2/\omega_{11.2} \quad .$$

We note that M is independent of u^*, and the elements of u^* and M are independent standard normal random variables. Furthermore, C^*, G^*, and M are mutually independent random matrices.

The LIML estimator (2.10) may be written as $\hat{\beta}$ satisfying

(6.14) $$(G^* - \lambda_{min} C^*) Q^{*\prime -1} \begin{pmatrix} -1 \\ \hat{\beta} \end{pmatrix} = 0$$

by using (6.5) and (6.6), where λ_{min} is the smallest characteristic root of G^*C^{*-1}. We may write (6.14) as

$$(6.15)\qquad (G^* - \lambda_{min} C^*) Q^{*\prime -1} (0 \quad \hat{\beta}' - \beta')' = (G^* - \lambda_{min} C^*) Q^{*\prime -1} (1 \quad -\beta')' \;.$$

Noting that

$$(6.16)\qquad Q^{*-1} = \begin{pmatrix} q^{11} & q^{12} \\ q^{21} & Q^{22} \end{pmatrix} = \begin{pmatrix} \beta' q^{21} + \sigma & \beta' Q^{22} \\ \Theta_{22}^{1/2} f^* & \Theta_{22}^{1/2} F^* \end{pmatrix} \;,$$

it follows that (6.15) becomes

$$(6.17)\qquad (G^* - \lambda_{min} C^*) \begin{pmatrix} f^* \\ F^* \end{pmatrix} \hat{e}_\beta = \mu (G^* - \lambda_{min} C^*) \begin{pmatrix} 1 \\ 0 \end{pmatrix} ,$$

where $\hat{e}_\beta$ is the first G_1 elements of $\hat{e}$ with the LIML estimator. Let

$$(6.18)\qquad (X_2' X_2)^{-1/2} X_2' x_2 = y$$

and

$$(6.19)\qquad C_{22}^{*-1/2} c_{21}^* = z \;,$$

then it is easily seen that

$$(6.20)\qquad y \sim N[0, I_{G_1}] \;,$$

$$(6.15)\qquad (G^* - \lambda_{min}C^*)Q^{*\prime -1}(0 \quad \hat{\beta}' - \beta')' = (G^* - \lambda_{min}C^*)Q^{*\prime -1}(1 \quad -\beta')' \quad .$$

Noting that

$$(6.16)\qquad Q^{*-1} = \begin{pmatrix} q^{11} & q^{12} \\ q^{21} & Q^{22} \end{pmatrix} = \begin{pmatrix} \beta' q^{21} + \sigma & \beta' Q^{22} \\ \Theta_{22}^{1/2} f^* & \Theta_{22}^{1/2} F^* \end{pmatrix} \quad ,$$

it follows that (6.15) becomes

$$(6.17)\qquad (G^* - \lambda_{min}C^*)\begin{pmatrix} f^* \\ F^* \end{pmatrix}\hat{e}_{\beta} = \mu(G^* - \lambda_{min}C^*)\begin{pmatrix} 1 \\ 0 \end{pmatrix} ,$$

where $\hat{e}_{\beta}$ is the first G_1 elements of $\hat{e}$ with the LIML estimator. Let

$$(6.18)\qquad (X_2'X_2)^{-1/2}X_2'x_2 = y$$

and

$$(6.19)\qquad C_{22}^{*-1/2}c_{21}^{*} = z \quad ,$$

then it is easily seen that

$$(6.20)\qquad y \sim N[0, I_{G_1}] \quad ,$$

$$(3.21) \qquad \underset{\sim}{z} \sim N[0, \underset{\sim}{I}_{G_1}] \quad ,$$

$$(3.22) \qquad \underset{\sim}{x}_2'\underset{\sim}{x}_2 - \underset{\sim}{x}_2'\underset{\sim}{X}_2(\underset{\sim}{X}_2'\underset{\sim}{X}_2)^{-1}\underset{\sim}{X}_2'\underset{\sim}{x}_2 \sim \chi^2(L - G_1) \quad ,$$

and

$$(3.23) \qquad c_{11}^* - \underset{\sim}{z}'\underset{\sim}{z} \sim \chi^2(q - G_1) \quad .$$

where we implicitly assumed $L - G_1 > 0$ and $q - G_1 > 0$.

We note that all of these random variables are independently distributed. Also let

$$(3.24) \qquad \text{plim} \frac{1}{\mu^2} \underset{\sim}{G}^* = \begin{pmatrix} g_{11}^{(0)} & \underset{\sim}{g}_{12}^{(0)} \\ \underset{\sim}{g}_{21}^{(0)} & \underset{\sim}{G}_{22}^{(0)} \end{pmatrix} \quad ,$$

and

$$(3.25) \qquad \text{plim} \frac{1}{\mu^2} \underset{\sim}{C}^* = \begin{pmatrix} \underset{\sim}{c}_{11}^{(0)} & \underset{\sim}{c}_{12}^{(0)} \\ \underset{\sim}{c}_{21}^{(0)} & \underset{\sim}{C}_{22}^{(0)} \end{pmatrix} \quad .$$

Then we obtain

$$(3.26) \qquad g_{11}^{(0)} = \frac{1}{\delta} \quad , \quad \underset{\sim}{g}_{12}^{(0)} = \underset{\sim}{g}_{21}^{(0)} = 0 \quad , \quad \underset{\sim}{G}_{22}^{(0)} = \frac{1}{\delta}\underset{\sim}{I}_{G_1} + (\underset{\sim}{F}^{*\prime}\underset{\sim}{F}^*)^{-1} \quad ,$$

and

$$(6.27) \qquad c_{11}^{(0)} = \frac{1}{\nu^2} \;,\quad \underset{\sim}{c}_{12}^{(0)} = \underset{\sim}{c}_{21}^{(0)} = 0 \;,\quad \underset{\sim}{C}_{22}^{(0)} = \frac{1}{\nu^2} \underset{\sim}{I}_{G_1} \;.$$

Lemma 3 :

$$(6.28) \qquad \lambda_{min} \overset{P}{\to} \lambda^{(0)} = \frac{\nu^2}{\delta} \;.$$

Proof:

Note that λ_{min} is the smallest root of the determinantal equation $|\underset{\sim}{G}^* - \lambda \underset{\sim}{C}^*| = 0$. Now, dividing by μ^2 gives

$$(6.29) \qquad \left|\frac{1}{\mu^2} \underset{\sim}{G}^* - \lambda \frac{1}{\mu^2} \underset{\sim}{C}^*\right| = 0 \;.$$

This equation converges in probability to

$$(6.30) \qquad \left| \begin{pmatrix} \frac{1}{\delta} & \underset{\sim}{0}' \\ \underset{\sim}{0} & \frac{1}{\delta} \underset{\sim}{I}_{G_1} + (F^{*\prime}F^*)^{-1} \end{pmatrix} - \lambda \begin{pmatrix} \frac{1}{\nu^2} & \underset{\sim}{0}' \\ \underset{\sim}{0} & \frac{1}{\nu^2} \underset{\sim}{I}_{G_1} \end{pmatrix} \right| = 0 \;,$$

so that λ_{min} converges to the smallest root of (6.29), which is ν^2/δ.

(QED)

Proof of Lemma 1 : Rewriting (6.17)

$$(6.31) \qquad (\underset{\sim}{f}^*\underset{\sim}{F}^*)\,(\underset{\sim}{G}^* - \lambda^*\underset{\sim}{C}^*)\,(\underset{\sim}{f}^*\underset{\sim}{F}^*)'\hat{\underset{\sim}{e}}_\beta = \mu(\underset{\sim}{f}^*\underset{\sim}{F}^*)\,(\underset{\sim}{G}^* - \lambda^*\underset{\sim}{C}^*)\begin{pmatrix} 1 \\ \underset{\sim}{0} \end{pmatrix} .$$

Then dividing by μ^3 gives

$$(6.32)\qquad [\underset{\sim}{f}^* \frac{1}{\mu^2}(g^*_{11} - \lambda^* c^*_{11}) + \underset{\sim}{F}^* \frac{1}{\mu^2}(\underset{\sim}{g}^*_{21} - \lambda^* \underset{\sim}{c}_{21})]\ (\underset{\sim}{f}^{*\prime} \frac{1}{\mu} \underset{\sim}{\hat{e}}_\beta - 1)$$

$$+ [\underset{\sim}{f}^* \cdot \frac{1}{\mu^2}(\underset{\sim}{g}^*_{21} - \lambda^* \underset{\sim}{c}^*_{21}) + \underset{\sim}{F}^* \frac{1}{\mu^2}(\underset{\sim}{G}^*_{22} - \lambda^* \underset{\sim}{C}^*_{22})]\underset{\sim}{F}^{*\prime} \frac{1}{\mu} \underset{\sim}{\hat{e}}_\beta$$

$$= 0 \quad .$$

Equation (6.32) converges in probability to

$$(6.33)\qquad \{(\frac{1}{\delta} - \frac{\lambda_*(0)}{\nu^2})f^*f^{*\prime} + [\underset{\sim}{I}_{G_1} + (\frac{1}{\delta} - \frac{\lambda_*(0)}{\nu^2})F^*F^{*\prime}]\}\underset{\sim}{e}_\beta^{(-1)}$$

$$= (\frac{1}{\delta} - \frac{\lambda_*(0)}{\nu^2})\underset{\sim}{f}^* \quad ,$$

where we denote

$$(6.34)\qquad \lambda_*^{(0)} = a\lambda^{(0)} + b \quad ,$$

and

$$(6.35)\qquad \text{plim}\ \frac{1}{\mu}\underset{\sim}{\hat{e}}_\beta = \underset{\sim}{e}_\beta^{(-1)} \quad .$$

Therefore, the estimator $\underset{\sim}{\hat{\beta}}_R$ is consistent iff

$$(6.36)\qquad \frac{1}{\delta} - \frac{\lambda_*(0)}{\nu^2} = 0 \quad .$$

Hence,

(6.37) $$(a - 1)\frac{\nu^2}{\delta} + b = 0 \quad .$$

This equation holds if $a = 1$ and $b = 0$, namely the LIML estimator. (QED)

Let v and w be independently normally distributed with mean 0 and variance 2 such that

(6.38) $$\mu[\frac{1}{\mu^2} \underset{\sim}{x}_2'\underset{\sim}{x}_2 - \frac{1}{\delta}] \overset{d}{\to} \frac{v}{\sqrt{\delta}} \quad ,$$

and

(6.39) $$\mu[\frac{1}{\mu^2} c^*_{11} - \frac{1}{\mu^2}] \overset{d}{\to} \frac{w}{\nu} \quad .$$

From (6.18) and (6.19)

(6.40) $$\frac{1}{\mu} X_2'x_2 \overset{d}{\to} \frac{1}{\sqrt{\delta}} \underset{\sim}{y} \quad ,$$

and

(6.41) $$\frac{1}{\mu} \underset{\sim}{c}^*_{21} \overset{d}{\to} \frac{1}{\nu} \underset{\sim}{z} \quad .$$

Lemma 4 :

(6.42) $$\hat{\underset{\sim}{e}}_{\beta} \overset{d}{\to} N[0, \underset{\sim}{Q}_1] \quad \text{for the LIML estimator} \quad ,$$

where

(6.43) $$\underset{\sim}{Q}_1 = \underset{\sim}{I}_{G_1} + \frac{1}{\delta}\left(1 + \frac{\nu^2}{\delta}\right)\underset{\sim}{F}^{*}\underset{\sim}{F}^{*\prime} \quad .$$

Proof: From (6.17), we obtain

(6.44) $$(\underset{\sim}{g}^{*}_{21} - \lambda_{\min}\, \underset{\sim}{c}^{*}_{21})\, (\underset{\sim}{f}^{*\prime}\hat{\underset{\sim}{e}}_{\beta} - \mu) + (\underset{\sim}{G}^{*}_{22} - \lambda_{\min}\, \underset{\sim}{C}^{*}_{22})\underset{\sim}{F}^{*\prime}\hat{\underset{\sim}{e}}_{\beta} = 0 \quad .$$

Dividing (6.44) by μ^2 gives

(6.45) $$[(\frac{1}{\mu^2}\, \underset{\sim}{g}^{*}_{21} - \lambda_{\min}\, \frac{1}{\mu^2}\, \underset{\sim}{c}^{*}_{21})\underset{\sim}{f}^{*\prime} + (\frac{1}{\mu^2}\, \underset{\sim}{G}^{*}_{22} - \lambda_{\min}\, \frac{1}{\mu^2}\, \underset{\sim}{C}^{*}_{22})\underset{\sim}{F}^{*\prime}]\hat{\underset{\sim}{e}}_{\beta}$$

$$= \tilde{\underset{\sim}{g}}^{(1)}_{21} - \lambda_{\min}\, \tilde{\underset{\sim}{c}}^{(1)}_{21} \quad .$$

By virtue of (6.40) and (6.41),

(6.46) $$\tilde{\underset{\sim}{g}}^{(1)}_{21} = \mu(\frac{1}{\mu^2}\, \underset{\sim}{g}^{*}_{21} - \underset{\sim}{g}^{(0)}_{21}) \overset{d}{\to} \underset{\sim}{g}^{(1)}_{21} = \frac{1}{\sqrt{\delta}}\, \underset{\sim}{y} + \underset{\sim}{F}^{*\prime -1}\underset{\sim}{x}_1 \quad ,$$

and

(6.47) $$\tilde{\underset{\sim}{c}}^{(1)}_{21} = \mu(\frac{1}{\mu^2}\, \underset{\sim}{c}^{*}_{21} - \underset{\sim}{c}^{(0)}_{21}) \overset{d}{\to} \underset{\sim}{c}^{(1)}_{21} = \frac{1}{\nu}\, \underset{\sim}{z} \quad .$$

Then using Lemma 3, equation (6.44) converges to

(6.48) $$[(\underset{\sim}{F}^{*\prime}\underset{\sim}{F}^{*})^{-1} + \frac{1}{\delta}\underset{\sim}{I}_{G_1} - \frac{\lambda^{(0)}}{\nu^2}\underset{\sim}{I}_{G_1}]\underset{\sim}{F}^{*\prime}\underset{\sim}{e}_\beta^{(0)}$$

$$= \frac{1}{\sqrt{\delta}}\underset{\sim}{y} + \underset{\sim}{F}^{*-1}\underset{\sim}{x}_1 - \frac{\lambda^{(0)}}{\nu}\underset{\sim}{z} \quad ,$$

where we denote

(6.49) $$\hat{\underset{\sim}{e}}_\beta \overset{d}{\to} \underset{\sim}{e}_\beta^{(0)} \quad .$$

Rearranging (6.46), we have

(6.50) $$\underset{\sim}{e}_\beta^{(0)} = \underset{\sim}{x}_1 + \underset{\sim}{F}^{*}(\frac{1}{\sqrt{\delta}}\underset{\sim}{y} - \frac{\lambda^{(0)}}{\nu}\underset{\sim}{z}) \quad .$$ (QED)

<u>Lemma 5</u>:

(6.51) $$\tilde{\lambda}^{(1)} = \mu(\lambda_{min} - \lambda^{(0)}) \overset{d}{\to} \lambda^{(1)} \quad ,$$

where

(6.52) $$\lambda^{(1)} \sim N[0, 2\frac{\nu^4}{\delta}(1 + \frac{\nu^2}{\delta})]$$

and is independent of $\underset{\sim}{x}_1$, $\underset{\sim}{y}$ and $\underset{\sim}{z}$.

<u>Proof</u>: From (6.17) we obtain

(6.53) $$(g_{11}^* - \lambda_{min} c_{11}^*)(\underset{\sim}{f}^{*\prime}\hat{\underset{\sim}{e}}_\beta - \mu) + (\underset{\sim}{g}_{12}^* - \lambda_{min}\underset{\sim}{c}_{12}^*)\underset{\sim}{F}^{*\prime}\hat{\underset{\sim}{e}}_\beta = 0 \quad .$$

Dividing (6.53) by μ^2 yields

$$(6.54) \qquad [(\frac{1}{\mu^2} g^*_{11} - \lambda_{\min} \frac{1}{\mu^2} c^*_{11}) \underset{\sim}{f}^{*\prime} + (\frac{1}{\mu^2} g^*_{12} - \lambda_{\min} \frac{1}{\mu^2} \underset{\sim}{c}^*_{12}) \underset{\sim}{F}^{*\prime}] \hat{\underset{\sim}{e}}_\beta$$

$$= \tilde{g}^{(1)}_{11} - \frac{c^*_{11}}{\mu^2} \tilde{\lambda}^{(1)} - \lambda^{(0)} \tilde{c}^{(1)}_{11} \quad ,$$

where

$$(6.55) \qquad \tilde{c}^{(1)}_{11} = \mu(\frac{1}{\mu^2} c^*_{11} - c^{(0)}_{11}) \overset{d}{\rightarrow} c^{(1)}_{11} = \frac{w}{\nu}$$

and

$$(6.56) \qquad \tilde{g}^{(1)}_{11} = \mu(\frac{1}{\mu^2} g^*_{11} - g^{(0)}_{11}) \overset{d}{\rightarrow} g^{(1)}_{11} = \frac{v}{\sqrt{\delta}} \quad .$$

Then in virtue of (6 26) and (6.27), equation (6.54) converges to

$$(6.57) \qquad 0 = \frac{v}{\sqrt{\delta}} - \frac{\lambda^{(1)}}{\nu^2} - \frac{\nu}{\delta} w \quad .$$

Hence, we get

$$(6.58) \qquad \lambda^{(1)} = \frac{\nu^2}{\sqrt{\delta}} v - \frac{\nu^3}{\delta} w \quad . \qquad \text{(QED)}$$

Proof of Lemma 2: Applying a similar argument in the proof of Lemma 5 to (6.32), we get

$$(6.59) \qquad \hat{\underset{\sim}{e}}_{\beta} \overset{d}{\to} \underset{\sim}{e}_{\beta}^{(0)} = \underset{\sim}{x}_1 + \underset{\sim}{F}^{*\prime}\left(\frac{1}{\sqrt{\delta}}\,\underset{\sim}{y} - \frac{\lambda^{*(0)}}{\nu}\,\underset{\sim}{z}\right) + (1-a)\,\frac{\lambda^{(1)}}{\nu^2}\,\underset{\sim}{f}^* \ .$$

Thus the asymptotic covariance matrix can be written

$$(6.60) \qquad AM(\underset{\sim}{e}_{\beta}^{(0)} \cdot \underset{\sim}{e}_{\beta}^{(0)\prime}) = \underset{\sim}{Q}_1 + \frac{2}{\delta}\left(1 + \frac{\nu^2}{\delta}\right)(1-a)^2 \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \ .$$

It is easily seen that this matrix can be minimized at $a = 1$. Then from (3.2) consistency requires $b = 0$, namely, the LIML estimator. (QED)

Lemma 6 (Asymptotic Bias):

$$(6.61) \qquad AE(\hat{e}_{LI}) = -\frac{1}{\mu}\begin{pmatrix} \underset{\sim}{Q}_1 \underset{\sim}{f}^* \\ 0 \end{pmatrix}$$

for the LIML estimator.

Proof: Let

$$(6.62) \qquad \begin{pmatrix} \hat{\underset{\sim}{e}}_{\beta} \\ \hat{\underset{\sim}{e}}_{\gamma} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{e}_{\beta}^{(0)} \\ \underset{\sim}{e}_{\gamma}^{(0)} \end{pmatrix} + \frac{1}{\mu}\begin{pmatrix} \underset{\sim}{e}_{\beta}^{(1)} \\ \underset{\sim}{e}_{\gamma}^{(1)} \end{pmatrix} + \frac{1}{\mu^2}\begin{pmatrix} \underset{\sim}{e}_{\beta}^{(2)} \\ \underset{\sim}{e}_{\gamma}^{(2)} \end{pmatrix} + O_p(\mu^{-3}) \ .$$

Recalling (6.44)

$$(6.63)\qquad -(\underset{\sim}{g}_{21}^{(2)} - \lambda^{(0)}\underset{\sim}{c}_{21}^{(2)} - \lambda^{(1)}\underset{\sim}{c}_{21}^{(1)} - \lambda^{(2)}\underset{\sim}{c}_{21}^{(0)})$$

$$+ (\underset{\sim}{g}_{21}^{(1)} - \lambda^{(0)}\underset{\sim}{c}_{21}^{(1)} - \lambda^{(1)}\underset{\sim}{c}_{21}^{(0)})\underset{\sim}{f}^{*\prime}\underset{\sim}{e}_{\beta}^{(0)} + (\underset{\sim}{g}_{21}^{(0)} - \lambda^{(0)}\underset{\sim}{c}_{21}^{(0)})\underset{\sim}{f}^{*\prime}\underset{\sim}{e}_{\beta}^{(1)}$$

$$+ (\underset{\sim}{G}_{22}^{(0)} - \lambda^{(0)}\underset{\sim}{C}_{22}^{(0)})\underset{\sim}{F}^{*\prime}\underset{\sim}{e}_{\beta}^{(1)} + (\underset{\sim}{G}_{22}^{(1)} - \lambda^{(0)}\underset{\sim}{C}_{22}^{(1)} - \lambda^{(1)}\underset{\sim}{C}_{22}^{(0)})\underset{\sim}{F}^{*\prime}\underset{\sim}{e}_{\beta}^{(0)}$$

$$= \underset{\sim}{0} \quad ,$$

where

$$(6.64)\qquad \underset{\sim}{g}_{21}^{(2)} = \underset{\sim}{g}_{21}^{*} - \mu^2\underset{\sim}{g}_{21}^{(0)} - \mu\underset{\sim}{g}_{21}^{(1)} \quad , \quad \underset{\sim}{G}_{22}^{(1)} = \mu(\frac{1}{\mu^2}\underset{\sim}{G}_{22}^{*} - \underset{\sim}{G}_{22}^{(0)}) \quad ,$$

$$(6.65)\qquad \lambda^{(2)} = \mu^2(\lambda_{min} - \lambda^{(0)} - \frac{1}{\mu}\lambda^{(1)}) \quad ,$$

and

$$(6.66)\qquad \underset{\sim}{c}_{21}^{(2)} = \underset{\sim}{c}_{21}^{*} - \mu^2\underset{\sim}{c}_{21}^{(0)} - \mu\underset{\sim}{c}_{21}^{(1)} \quad .$$

Using Lemmas 3 and 5, (6.26) and (6.27)

$$(6.67)\qquad \underset{\sim}{e}_{\beta}^{(1)} = -\underset{\sim}{f}^{*\prime}\underset{\sim}{e}_{\beta}^{(0)} \cdot \underset{\sim}{e}_{\beta}^{(0)} + \underset{\sim}{F}^{*}[(\underset{\sim}{g}_{21}^{(2)} - \lambda^{(0)}\underset{\sim}{c}_{21}^{(2)} - \lambda^{(1)}\underset{\sim}{c}_{21}^{(1)})$$

$$- (\underset{\sim}{G}_{22}^{(1)} - \lambda^{(0)}\underset{\sim}{C}_{22}^{(1)} - \lambda^{(1)}\underset{\sim}{C}_{22}^{(0)})\underset{\sim}{F}^{*\prime}\underset{\sim}{e}_{\beta}^{(0)}] \quad .$$

For $\underset{\sim}{e}_{\beta}^{(1)}$, since expectations of the terms in the parenthesis are zero, the only possible non-zero contribution comes from the first term. Hence,

(6.68) $$AE(\underset{\sim}{e}_{\beta}^{(1)}) = -\underset{\sim}{Q}_1\underset{\sim}{f}^* \quad .$$

Finally, from (3.8),

(6.69) $$\underset{\sim}{e}_{\gamma}^{(0)} = \underset{\sim}{u}^* \quad ,$$

(6.70) $$\underset{\sim}{e}_{\gamma}^{(1)} = -(\underset{\sim}{M}\underset{\sim}{F}^{*\prime} + \underset{\sim}{u}^*\underset{\sim}{f}^{*\prime})\underset{\sim}{e}_{\beta}^{(0)}$$

(6.71) $$\underset{\sim}{e}_{\gamma}^{(2)} = -(\underset{\sim}{M}\underset{\sim}{F}^{*\prime} + \underset{\sim}{u}^*\underset{\sim}{f}^{*\prime})\underset{\sim}{e}_{\beta}^{(1)} \quad ,$$

where $u^* \sim N(0,\underset{\sim}{I}_{K_1})$ and each row of $\underset{\sim}{M}$ is independently distributed as $N(0,\underset{\sim}{I}_{G_1})$, and they are mutually independent. Therefore, we get (6.61). (QED)

<u>Lemma 7 :</u> Let

(6.72) $$\hat{\underset{\sim}{e}}_M = \hat{\underset{\sim}{e}}_{LI} + \frac{c}{\ell}(\hat{\underset{\sim}{e}}_{TS} - \hat{\underset{\sim}{e}}_{LI}) \quad ,$$

where c is a positive constant. Then

(6.73) $$\| AM(\hat{\underset{\sim}{e}}_M) \| < \| AM(\hat{e}_{LI}) \| \qquad \text{if } 0 \leq c \leq 1 \quad ,$$

and

(6.74) $$AM(\hat{\mathbf{e}}_M \ \hat{\mathbf{e}}_M') < AM(\hat{\mathbf{e}}_{LI} \ \hat{\mathbf{e}}_{LI}') \qquad \text{if} \quad 0 \leq c \leq 8 \ .$$

Proof: For the TSLS estimator, write

(6.75) $$\frac{1}{\mu} \mathbf{e}_{\beta.TS} = \mathbf{e}_{\beta.TS}^{(-1)} + \frac{1}{\mu} \mathbf{e}_{\beta.TS}^{(0)} + 0_p(\mu^{-2}) \ .$$

Then, again using (6.17),

(6.76) $$(\mathbf{f}^*\mathbf{g}_{11}^* + \mathbf{F}^*\mathbf{g}_{21}^*)(\mathbf{f}^{*\prime}\hat{\mathbf{e}}_{\beta.TS} - \mu) + (\mathbf{f}^*\mathbf{g}_{12}^* + \mathbf{F}^*\mathbf{G}_{22}^*)\mathbf{F}^{*\prime}\hat{\mathbf{e}}_{\beta.TS} = \mathbf{0} \ .$$

Dividing by μ^3 and taking probability limits in each term yields

(6.77) $$\mathbf{e}_{\beta.TS}^{(-1)} = \frac{1}{\delta} \mathbf{D}^{-1}\mathbf{f}^* \ ,$$

where

(6.78) $$\mathbf{D} = \mathbf{I}_{G_1} + \frac{1}{\delta}(\mathbf{f}^*\mathbf{f}^{*\prime} + \mathbf{F}^*\mathbf{F}^{*\prime}) \ .$$

Similarly from (6.76),

(6.79) $$(\mathbf{f}^*\mathbf{g}_{11}^{(0)} + \mathbf{F}^*\mathbf{g}_{21}^{(0)})\mathbf{f}^{*\prime}\mathbf{e}_\beta^{(0)} + (\mathbf{f}^*\mathbf{g}_{11}^{(1)} + \mathbf{F}^*\mathbf{g}_{21}^{(1)})(\mathbf{f}^{*\prime}\mathbf{e}_\beta^{(-1)} - 1)$$
$$+ (\mathbf{f}^*\mathbf{g}_{12}^{(0)} + \mathbf{F}^*\mathbf{G}_{22}^{(0)})\mathbf{F}^{*\prime}\mathbf{e}_\beta^{(0)} + (\mathbf{f}^*\mathbf{g}_{12}^{(1)} + \mathbf{F}^*\mathbf{G}_{22}^{(1)})\mathbf{F}^{*\prime}\mathbf{e}_\beta^{(-1)} = \mathbf{0} \ .$$

(6.80) $$\underset{\sim}{e}^{(0)}_{\beta.TS} = -\underset{\sim}{D}^{-1} \{(\frac{v}{\sqrt{\delta}} \underset{\sim}{f}^* + \underset{\sim}{x}_1 + \frac{1}{\sqrt{\delta}} \underset{\sim}{F}^*\underset{\sim}{y}) (\underset{\sim}{f}^{*\prime}\underset{\sim}{e}^{(-1)}_{\beta} - 1)$$

$$+ [\underset{\sim}{f}^*(\underset{\sim}{F}^{*\prime -1}\underset{\sim}{x}_1' + \frac{1}{\sqrt{\delta}} \underset{\sim}{y}) + \underset{\sim}{F}^*(\underset{\sim}{X}_1'\underset{\sim}{F}^{*\prime -1} + \underset{\sim}{F}^{*-1}\underset{\sim}{X}_1 + \frac{1}{\sqrt{\delta}} \underset{\sim}{U}^{(1)})]\underset{\sim}{F}^{*\prime}\underset{\sim}{e}^{(-1)}_{\beta}\}$$

where

(6.81) $$(\frac{1}{\mu^2} \underset{\sim}{X}_2'\underset{\sim}{X}_2 - \frac{1}{\delta} \underset{\sim}{I}_{G_1}) \overset{d}{\rightarrow} \frac{1}{\sqrt{\delta}} \underset{\sim}{U}^{(1)} \quad ,$$

and the expectation of each element of $\underset{\sim}{U}^{(1)}$ is zero. Then the standardized combined estimator may be expanded

(6.82) $$\hat{\underset{\sim}{e}}_{\beta} = \underset{\sim}{e}^{(0)}_{\beta.LI} + \frac{1}{\mu} [\underset{\sim}{e}^{(1)}_{\beta.LI} + c\delta \underset{\sim}{e}^{(-1)}_{\beta.TS}]$$

$$+ \frac{1}{\mu^2} [\underset{\sim}{e}^{(2)}_{\beta.LI} + c\delta(\underset{\sim}{e}^{(0)}_{\beta.TS} - \underset{\sim}{e}^{(0)}_{\beta.LI})] + O_p(\mu^{-3}) \quad .$$

Taking expectations (with respect to the approximate distribution of the estimator) gives

(6.83) $$AM(\hat{\underset{\sim}{e}}_{\beta}) = \frac{1}{\mu} (-\underset{\sim}{Q}_1 + c\underset{\sim}{D}^{-1})\underset{\sim}{f}^* \quad ,$$

and

(6.84) $$AM(\hat{\underset{\sim}{e}}_{\beta} \quad \hat{\underset{\sim}{e}}_{\beta}') - AM(\hat{\underset{\sim}{e}}_{\beta.LI} \quad \hat{\underset{\sim}{e}}_{\beta.LI}')$$

$$= \frac{1}{\mu^2} AM[c^2\delta^2 \underset{\sim}{e}^{(-1)}_{\beta.TS} \quad \underset{\sim}{e}^{(-1)\prime}_{\beta.TS} + c\delta \underset{\sim}{e}^{(-1)}_{\beta.TS} \quad \underset{\sim}{e}^{(1)\prime}_{\beta.LI}$$

$$+ c\delta \underset{\sim}{e}^{(1)}_{\beta.LI} \quad \underset{\sim}{e}^{(-1)'}_{\beta.TS} + c\delta \underset{\sim}{e}^{(0)}_{\beta.LI} (\underset{\sim}{e}^{(0)}_{\beta.TS} - \underset{\sim}{e}^{(0)}_{\beta.LI})'$$

$$+ c\delta (\underset{\sim}{e}^{(0)}_{\beta.TS} - \underset{\sim}{e}^{(0)}_{\beta.LI}) \underset{\sim}{e}^{(0)'}_{\beta.LI}]$$

$$= \frac{1}{\mu^2} \{c^2 \underset{\sim}{D}^{-1} \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \underset{\sim}{D}^{-1} - c[\underset{\sim}{Q}_1 \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \underset{\sim}{D}^{-1} + \underset{\sim}{D}^{-1} \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \underset{\sim}{Q}_1]$$

$$+ \delta c [\underset{\sim}{D}^{-1} \underset{\sim}{Q}_2 + \underset{\sim}{Q}_2 \underset{\sim}{D}^{-1} - 2\underset{\sim}{Q}_1]$$

$$- c[\underset{\sim}{D}^{-1} \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \underset{\sim}{D}^{-1} \underset{\sim}{Q}_2 + \underset{\sim}{Q}_2 \underset{\sim}{D}^{-1} \underset{\sim}{f}^* \underset{\sim}{f}^{*\prime} \underset{\sim}{D}^{-1}$$

$$+ \underset{\sim}{f}^{*\prime} \underset{\sim}{D}^{-1} \underset{\sim}{f}^* (\underset{\sim}{D}^{-1} \underset{\sim}{Q}_2 + \underset{\sim}{Q}_2 \underset{\sim}{D}^{-1})]\} \quad ,$$

where

(6.85) $$\underset{\sim}{Q}_2 = \underset{\sim}{I}_{G_1} + \frac{1}{\delta} \underset{\sim}{F}^* \underset{\sim}{F}^{*\prime} \quad .$$

Now

(6.86) $$\delta[\underset{\sim}{D}^{-1} \underset{\sim}{Q}_2 + \underset{\sim}{Q}_2 \underset{\sim}{D}^{-1} - 2\underset{\sim}{Q}_1]$$

$$= -\underset{\sim}{D}^{-1}(\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime} + \underset{\sim}{f}^*\underset{\sim}{f}^{*\prime})\underset{\sim}{Q}_2 - \underset{\sim}{Q}_2(\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime} + \underset{\sim}{f}^*\underset{\sim}{f}^{*\prime})\underset{\sim}{D}^{-1} - 2\,\frac{\nu^2}{\delta}\,\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime}$$

$$< -2\underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1} \quad .$$

Because $\underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1} < \underset{\sim}{Q}_1\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1}$ and $\underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1} < \underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}\underset{\sim}{Q}_1$, (6.84) is non-positive definite if $0 \le c \le 8$. For $\hat{\underset{\sim}{e}}_\gamma$, from (6.69) (6.70).

(6.87).
$$AM(\hat{\underset{\sim}{e}}_\gamma\hat{\underset{\sim}{e}}_\gamma') = \underset{\sim}{I}_{K_1} + \frac{1}{\mu^2}\,AM[\underset{\sim}{e}_\gamma^{(1)}\underset{\sim}{e}_\gamma^{(1)\prime}$$

$$+ \underset{\sim}{e}_\gamma^{(2)}\underset{\sim}{e}_\gamma^{(0)\prime} + \underset{\sim}{e}_\gamma^{(0)}\underset{\sim}{e}_\gamma^{(2)\prime}]$$

$$= \underset{\sim}{I}_{K_1} + \frac{1}{\mu^2}\,[tr(\underset{\sim}{Q}_1\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime})\underset{\sim}{I}_{K_1} + 3\underset{\sim}{f}^{*\prime}\underset{\sim}{Q}_1\underset{\sim}{f}^*\underset{\sim}{I}_{K_1}$$

$$- 2c\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{I}_{K_1}]$$

$$= AM(\hat{\underset{\sim}{e}}_\gamma\hat{\underset{\sim}{e}}_\gamma')_{LI} - \frac{1}{\mu^2}\,[2c\underset{\sim}{f}^{*\prime}\underset{\sim}{D}^{-1}\underset{\sim}{f}^*\underset{\sim}{I}_{K_1}]$$

$$< AM(\hat{\underset{\sim}{e}}_\gamma\hat{\underset{\sim}{e}}_\gamma')_{LI} \qquad \text{if} \quad c \ge 0 \quad .$$

Then using the asymptotic independence of $\hat{\underset{\sim}{e}}_\beta$ and $\hat{\underset{\sim}{e}}_\gamma$, we get (6.74).

(QED)

Lemma 8 : Let

(6.88) $$\hat{e}_F = \hat{e}_R$$

and

(6.89) $$\lambda_F = \lambda_{min} - \frac{c}{q} \quad ,$$

where c is a positive constant. Then

(6.90) $$\|AM(\hat{e}_F)\| < \|AM(\hat{e}_{LI})\| \qquad \text{if} \quad 0 \leq c \leq 1 \quad ,$$

and

(6.91) $$AM(\hat{e}_F \hat{e}'_F) < AM(\hat{e}_{LI} \hat{e}'_{LI}) \qquad \text{if} \quad 0 \leq c \leq 8 \quad .$$

<u>Proof</u>: Let

(6.92) $$\hat{e}_{\beta.F} = e^{(0)}_{\beta.F} + \frac{1}{\mu} e^{(1)}_{\beta.F} + \frac{1}{\mu^2} e^{(1)}_{\beta.F} + O_p(\mu^{-3}) \quad ,$$

and

(6.93) $$\lambda_F = \lambda^{(0)} + \frac{1}{\mu} \lambda^{(1)} + \frac{1}{\mu^2} [\lambda^{(2)} - c\nu^2] + O_p(\mu^{-3}) \quad .$$

Recalling (6.17),

(6.94) $$(f^* F^*) \; (G^* - \lambda_F C^*) \begin{pmatrix} f^{*\prime}\hat{e}_\beta - \mu \\ F^{*\prime}\hat{e}_\beta \end{pmatrix} = 0$$

and the substitution of (6.91) and (6.92) gives

(6.95) $$\underset{\sim}{e}^{(0)}_{\beta.F} = \underset{\sim}{e}^{(0)}_{\beta.LI} \quad ,$$

(6.96) $$\underset{\sim}{e}^{(1)}_{\beta.F} = \underset{\sim}{e}^{(1)}_{\beta.LI} + c\underset{\sim}{f}^* \quad ,$$

(6.97) $$\underset{\sim}{e}^{(2)}_{\beta.F} = \underset{\sim}{e}^{(2)}_{\beta.LI} + c\nu^2 \{(c^{(1)}_{11} - 2c^{(0)}_{11}\underset{\sim}{f}^{*\prime}\underset{\sim}{e}^{(0)}_{\beta.LI})\underset{\sim}{f}^*$$

$$- c^{(0)}_{11}\underset{\sim}{f}^{*\prime}\underset{\sim}{f}^*\underset{\sim}{e}^{(0)}_{\beta.LI} + \underset{\sim}{F}^*\underset{\sim}{c}^{(1)}_{21} - \underset{\sim}{F}^*\underset{\sim}{C}^{(0)}_{22}\ \underset{\sim}{F}^{*\prime}\underset{\sim}{e}^{(0)}_{\beta.LI}$$

$$- c^{(0)}_{11}\underset{\sim}{F}^*(\underset{\sim}{G}^{(1)}_{22} - \lambda^{(0)}\underset{\sim}{C}^{(1)}_{22} - \lambda^{(1)}\underset{\sim}{C}^{(0)}_{22})\underset{\sim}{F}^{*\prime}\underset{\sim}{f}^*\} \quad .$$

Then after some calculation of expectations,

(6.98) $$AM(\hat{\underset{\sim}{e}}_{\beta.F}) = \frac{1}{\mu}(-\underset{\sim}{Q}_1 + c\underset{\sim}{I}_{G_1})\underset{\sim}{f}^* \quad ,$$

(6.99) $$AM(\hat{\underset{\sim}{e}}_{\beta.F}\hat{\underset{\sim}{e}}'_{\beta.F}) - AM(\underset{\sim}{e}_{\beta.LI}\underset{\sim}{e}'_{\beta.LI})$$

$$= \frac{1}{\mu^2}\{c^2\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime} - 3c\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime} - 3c\underset{\sim}{Q}_1\underset{\sim}{f}^*\underset{\sim}{f}^{*\prime}$$

$$- 2c\underset{\sim}{f}^{*\prime}\underset{\sim}{f}^*\underset{\sim}{Q}_1 - c\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime}\underset{\sim}{Q}_1 - c\underset{\sim}{Q}_1\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime} - 2c\frac{\nu^2}{\delta}\underset{\sim}{F}^*\underset{\sim}{F}^{*\prime}\} \quad .$$

The last term is non-positive definite if $0 \leq c \leq 8$. Similarly, for $\hat{\underset{\sim}{e}}_{\gamma}$,

(6.100)
$$AM(\hat{e}_{\gamma} \quad \hat{e}_{\gamma}') = I_{K_1} + \frac{1}{\mu^2}\{tr(Q_1 F^* F^{*\prime}) I_{K_1}$$

$$+ 3f^{*\prime} Q_1 f^* I_{K_1} - 2c f^{*\prime} f^* I_{K_1}\}$$

$$= AM(\hat{e}_{\gamma} \quad \hat{e}_{\gamma}')_{LI} + \frac{1}{\mu^2}[-2c f^{*\prime} f^* I_{K_1}] \quad .$$

Finally, using the asymptotic independence of $\hat{e}_{\beta}$ and $\hat{e}_{r}$, we complete the proof. (QED)

Appendix

Figure 1

Densities

T-K =5, K_2=15, α=1.0, μ^2=200.0

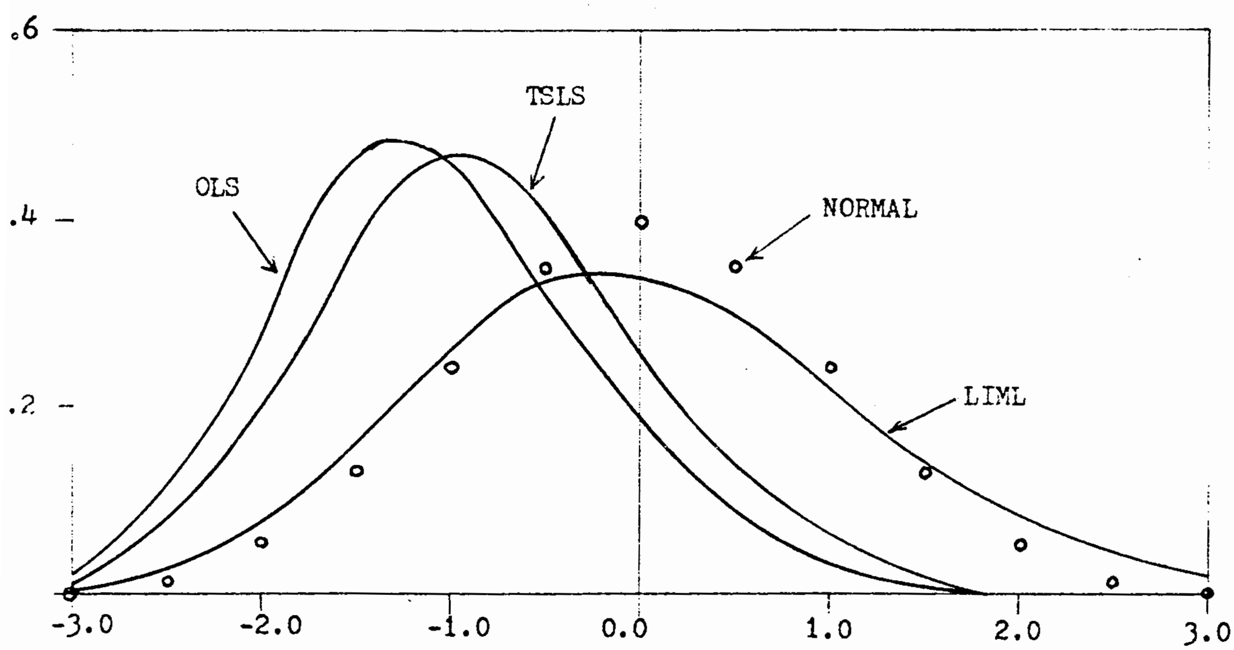


Figure 2

Densities

T-K =5, K_2=25, α=1.0, μ^2=300.0

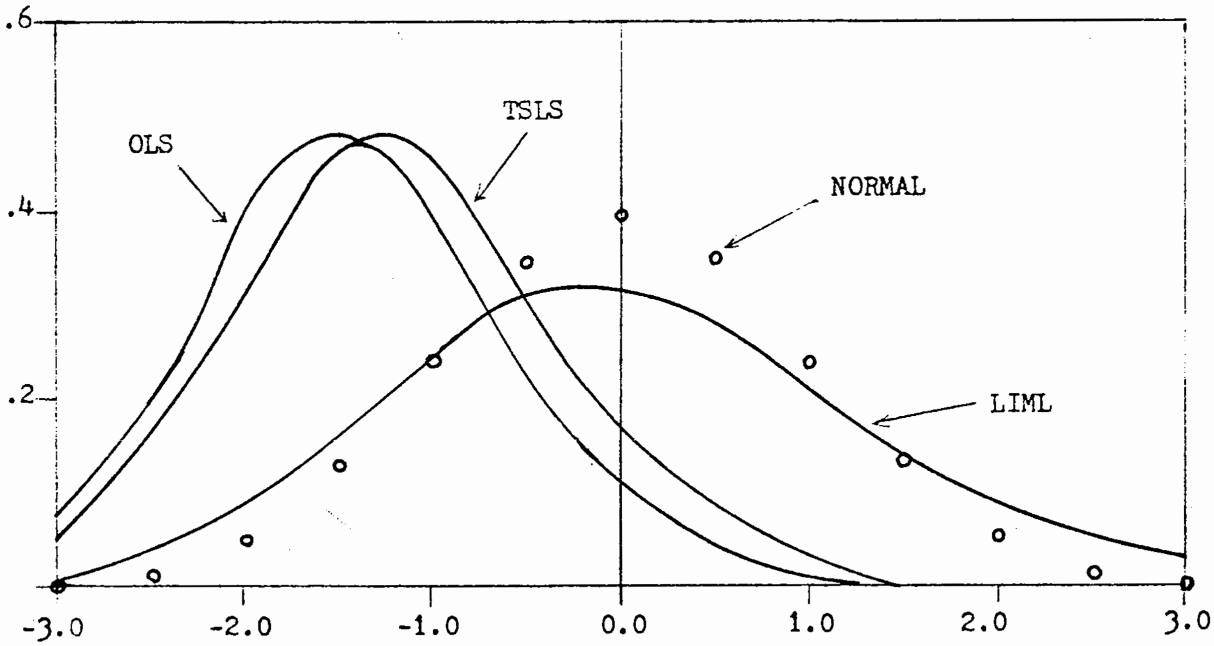


Figure 3

Densities

T-K=5, K_2=45, α=1.0, μ^2=500.0

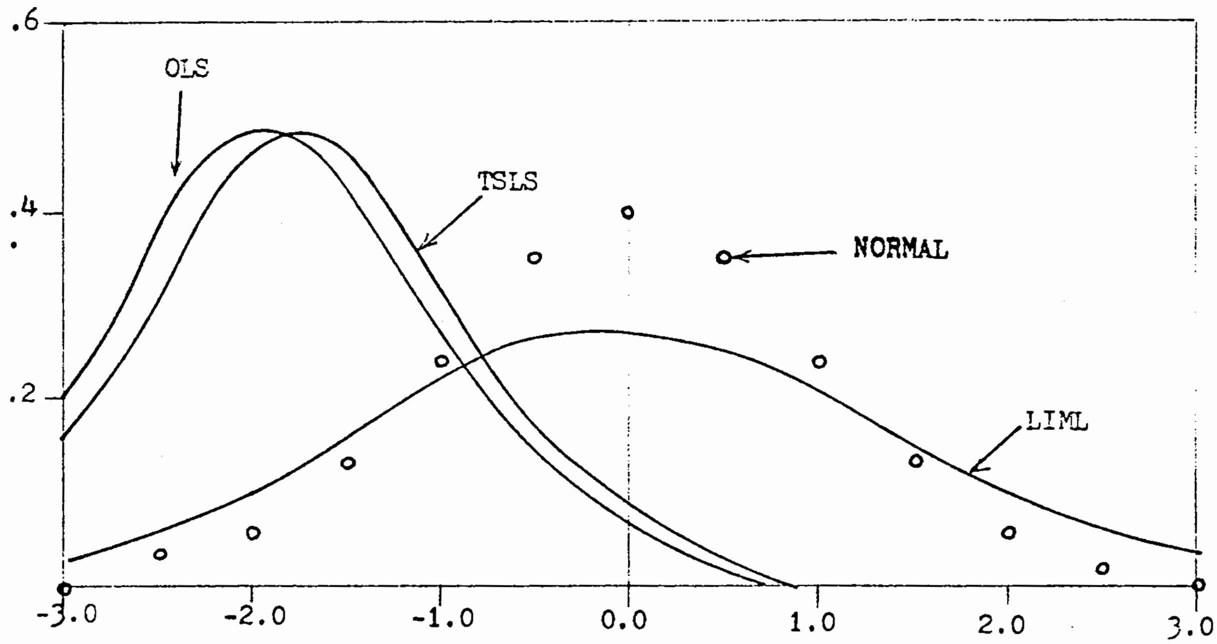


Figure 4

Densities

T-K=20 K_2=15, α=1.0, μ^2=350.0

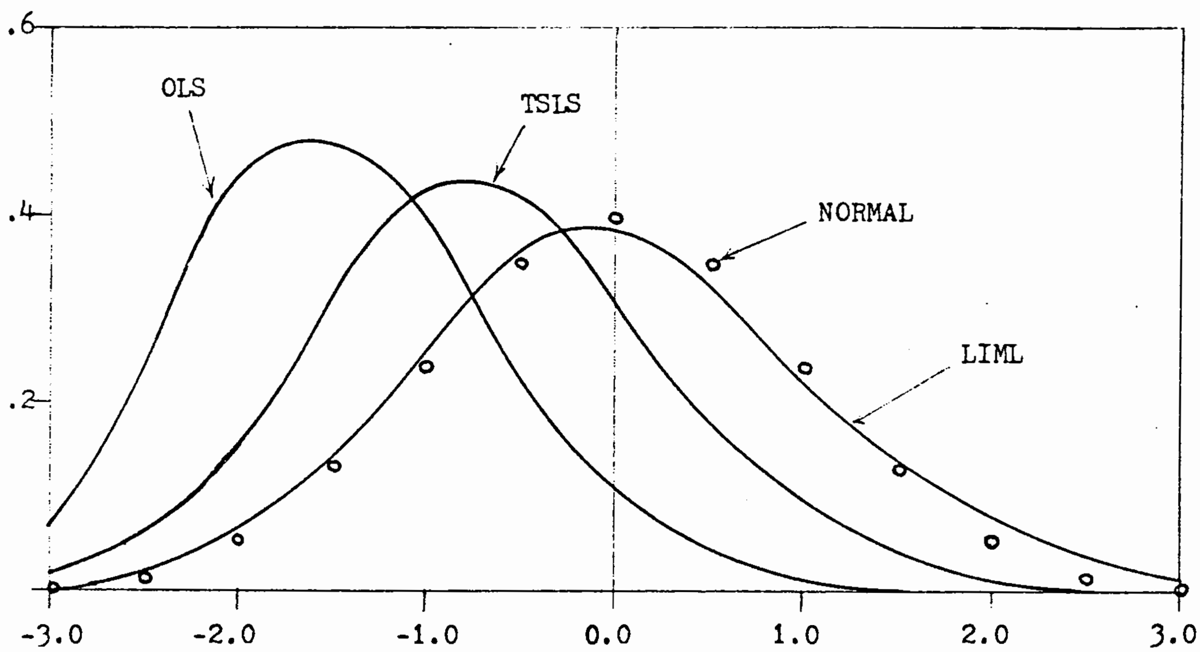
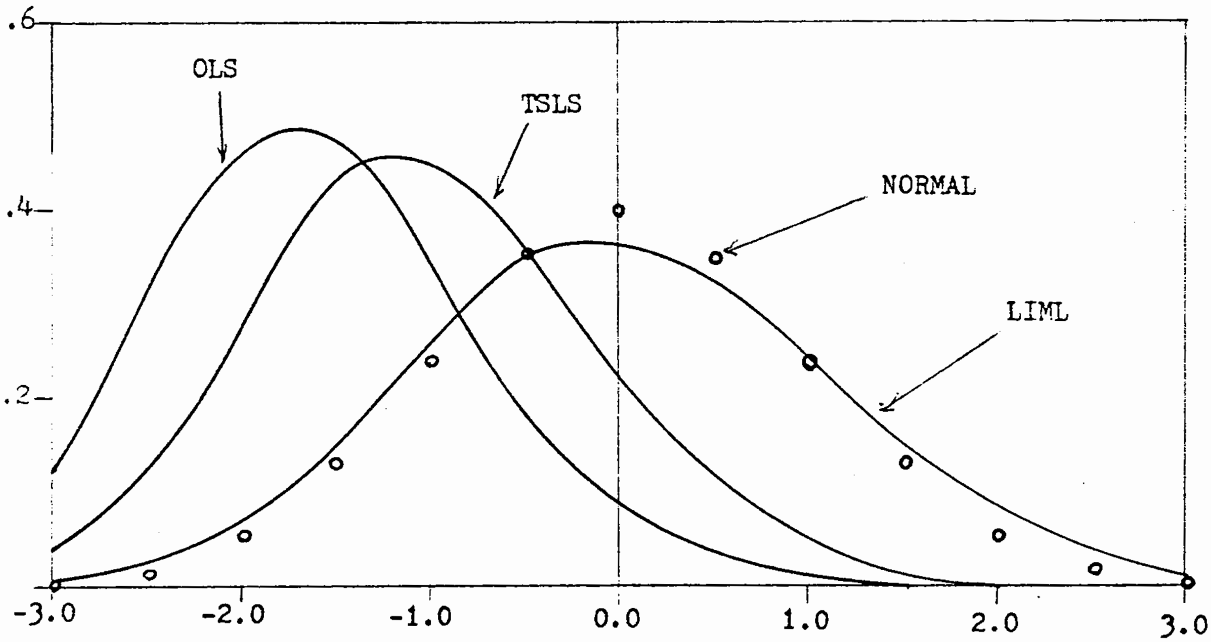


Figure 5

Densities

$T\text{-}K=20,\ K_2=25,\ \alpha=1.0,\ \mu^2=400.0$



A-6

Figure 6

Densities

T-K=20, K_2=45, α=1.0, μ^2= 650.0

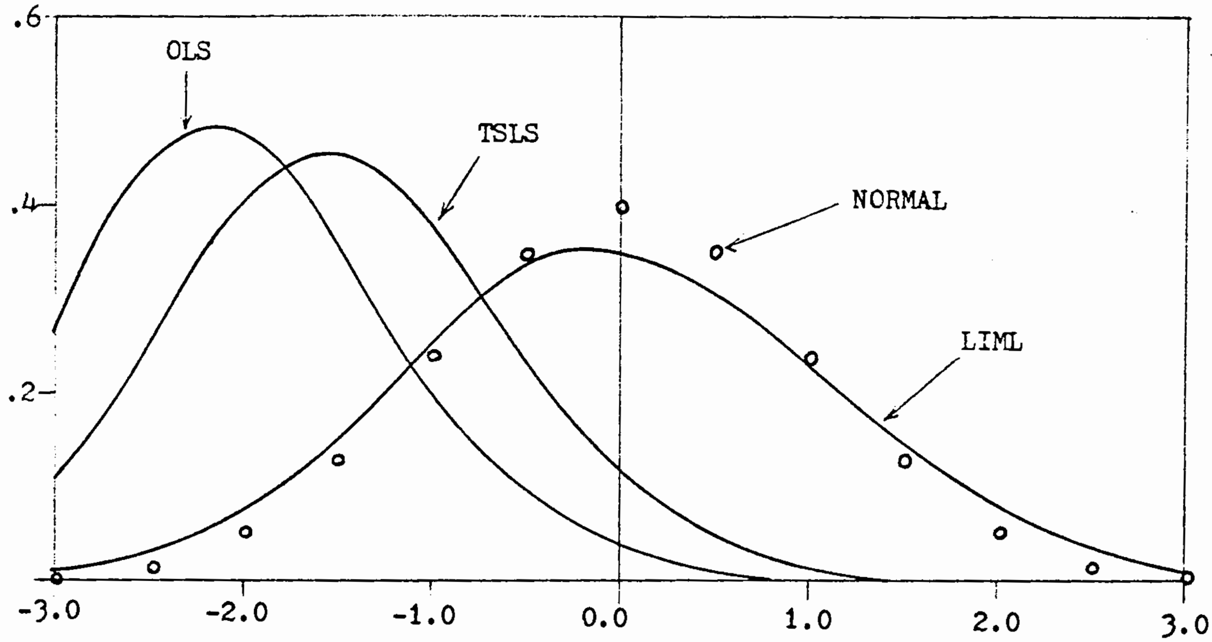


Figure 7

Densities

T-K=40, K_2=15, α=1.0, μ^2=550.0

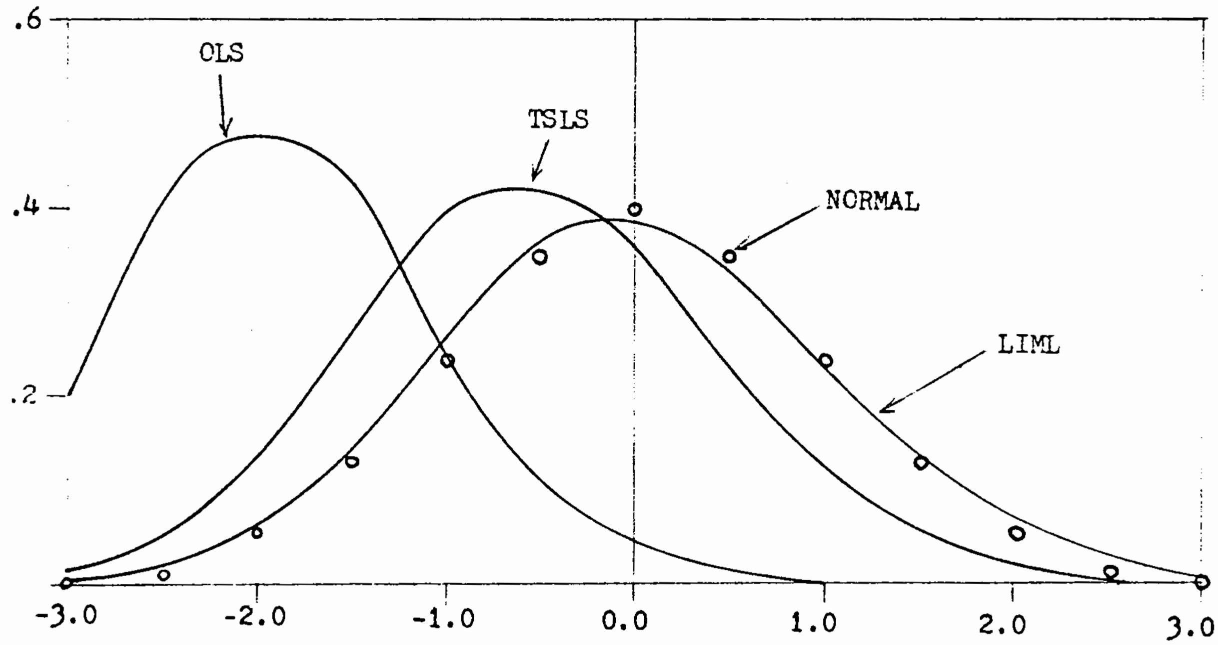


Figure 8

Densities

T-K=40, K_2=25, α=1.0, μ^2=650.0

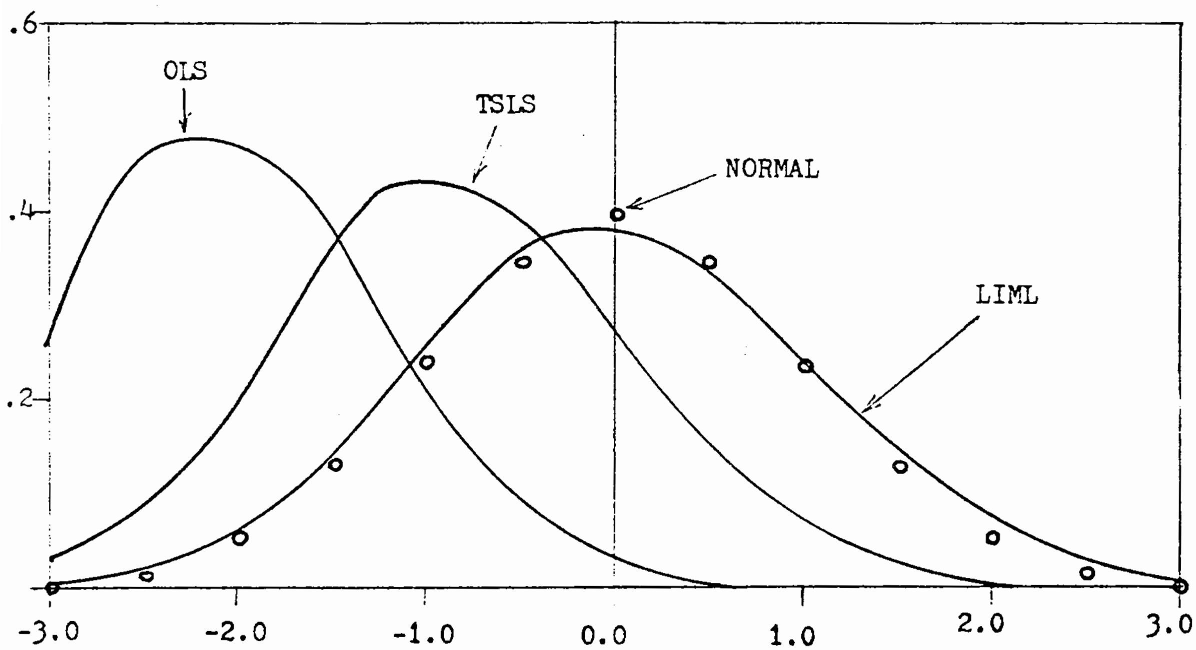
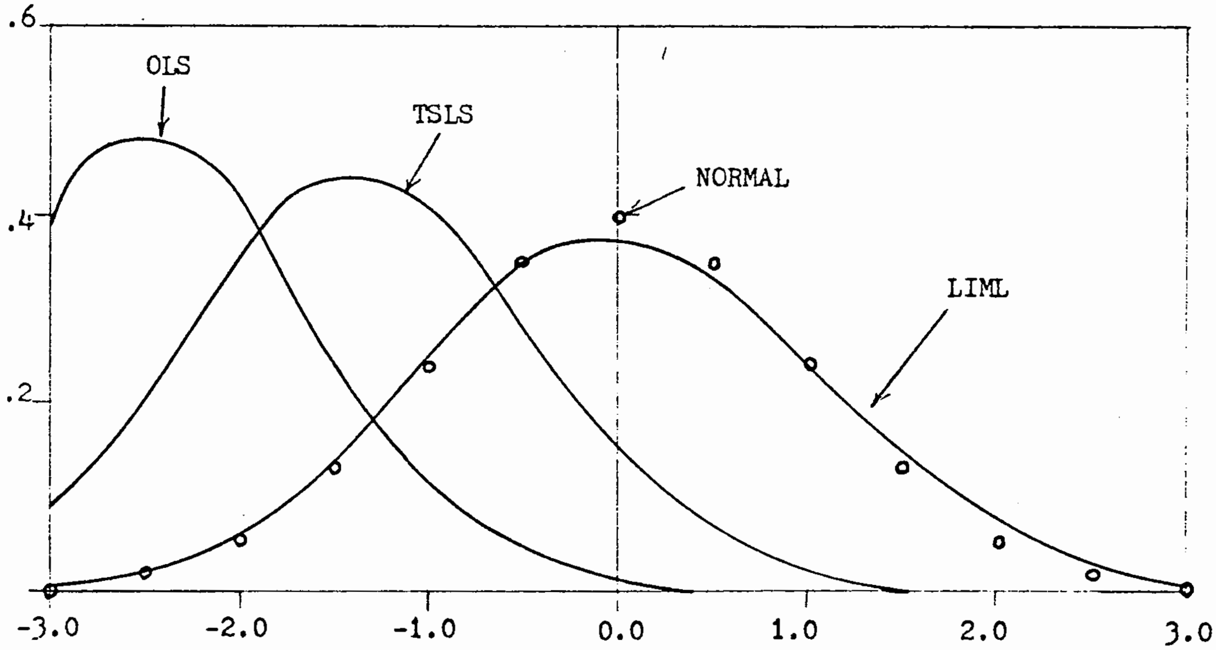


Figure 9

Densities

T-K=40, K_2=45, α=1.0, μ^2=850.0



REFERENCES

Anderson, T.W., Naoto Kunitomo and Takamitsu Sawa [1981]. "Evaluation of the Distribution Function of the Limited Information Maximum Likelihood Estimator," forthcoming in Econometrica.

Anderson, T.W. and Takamitsu Sawa. [1979]. "Evaluation of the Distribution Function of the Two-Stage Least Squares Estimate," Econometrica, 47, 163-183.

Dhrymes, P.J. [1971]. "A Simplified Structural Estimator for Large-Scale Econometric Models," The Australian Journal of Statistics, 13, 168-175.

Dusenberry, et al. [1969]. The Brookings Model: Some Further Results, Chicago: Rand McNally.

Evans, M. and R. Klein. [1968]. "The Wharton Econometric Forecasting Model," Economic Forecasting Unit, Department of Economics, University of Pennsylvania, Philadelphia.

Fujikoshi, Y., K. Morimune, N. Kunitomo and M. Taniguchi. [1979]. "Asymptotic Expansions of the Distributions of the Estimates of Coefficients in a Simultaneous Equation System," Technical Report No. 287, Economic Series, Institute for Mathematical Studies in the Social Sciences (IMSSS), Stanford University.

Fuller, W.A. [1977]. "Some Properties of a Modification of the Limited Information Estimator," Econometrica, 45, 939-953.

Kadane, Joseph B. [1971]. "Comparison of k-class Estimators When the Disturbances Are Small," Econometrica, 65, 723-737.

Klein, L. R. [1950]. Economic Fluctuations in the United States, New York: John Wiley and Sons.

Klein, L. R. [1971]. "Forecasting and Policy Evaluation Using Large Scale Econometric Models: The State of the Art," in Intrigator, M.D., ed., Frontiers of Quantitative Economics, Amsterdam: North-Holland.

Klein, L. R. and A. Goldberger. [1955]. An Econometric Model of the United States, 1929-1952, Amsterdam: North Holland.

Kunitomo, Naoto. [1980]. "Asymptotic Expansions of the Distributions of Estimators in a Linear Functional Relationship and Simultaneous Equations," Journal of the American Statistical Association, 75, 693-700.

Kunitomo, Naoto [1981]. "Asymptotic Efficiency and Higher Order Efficiency of the Limited Information Maximum Likelihood Estimator in Large Econometric Models," Unpublished Ph.D. Dissertation, Stanford University.

Kunitomo, Naoto, K. Morimune and Y. Tsukuda. [1980]. "Asymptotic Expansions of the Distributions of k-class Estimators When the Disturbances Are Small," Technical Report No. 307, Economics Series, Institute for Mathematical Studies in the Social Sciences (IMSSS), Stanford University.

Maddala, G.S. [1980]. "Statistical Inference in Relation to the Size of the Model," unpublished manuscript.

Morimune, Kimio [1978]. "Improving the Limited Information Maximum Likelihood Estimator When the Disturbances Are Small," Journal of the American Statistical Association, 73, 867-871.

Morimune, Kimio and Naoto Kunitomo. [1980]. "Improving the Maximum Likelihood Estimate in Linear Functional Relationships for Alternative Parameter Sequences," Journal of the American Statistical Association, 75, 230-237.

Rothenberg, T. [1978]. "Second Order Efficiency of Estimators and Tests in Simultaneous Equations Models," Notes for Joint Berkeley/Stanford Seminar, November 30.

Sargan, J.D. [1975]. "Asymptotic Theory and Large Models," International Economic Review, 16, 75-91.

Takeuchi, Kei. [1978]. "Asymptotic Higher Order Efficiency of the ML Estimators of Parameters in Linear Simultaneous Equations," Paper presented at the Kyoto University.

Takeuchi, Kei and K. Morimune. [1979]. "Asymptotic Completeness of the Extended Maximum Likelihood Estimators in a Simultaneous Equation System," Unpublished manuscript.

Theil, Henri. [1971]. "A Simple Modification of the Two-Stage Least Squares Procedure for Undersized Sample," Report 7107, Center for Mathematical Studies in Business and Economics, University of Chicago.

Maddala, G.S. [1980]. "Statistical Inference in Relation to the Size of the Model," unpublished manuscript.

Morimune, Kimio [1978]. "Improving the Limited Information Maximum Likelihood Estimator When the Disturbances Are Small," Journal of the American Statistical Association, 73, 867-871.

Morimune, Kimio and Naoto Kunitomo. [1980]. "Improving the Maximum Likelihood Estimate in Linear Functional Relationships for Alternative Parameter Sequences," Journal of the American Statistical Association, 75, 230-237.

Rothenberg, T. [1978]. "Second Order Efficiency of Estimators and Tests in Simultaneous Equations Models," Notes for Joint Berkeley/Stanford Seminar, November 30.

Sargan, J.D. [1975]. "Asymptotic Theory and Large Models," International Economic Review, 16, 75-91.

Takeuchi, Kei. [1978]. "Asymptotic Higher Order Efficiency of the ML Estimators of Parameters in Linear Simultaneous Equations," Paper presented at the Kyoto University.

Takeuchi, Kei and K. Morimune. [1979]. "Asymptotic Completeness of the Extended Maximum Likelihood Estimators in a Simultaneous Equation System," Unpublished manuscript.

Theil, Henri. [1971]. "A Simple Modification of the Two-Stage Least Squares Procedure for Undersized Sample," Report 7107, Center for Mathematical Studies in Business and Economics, University of Chicago.