COMPARISON OF SINGLE EQUATION METHODS OF PREDICTION IN A SIMULTANEOUS EQUATION SYSTEM*

By YOSHIHIKO TSUKUDA

1. Introduction

Prediction of endogenous variables in a simultaneous equation model is usually based on estimates of the reduced form coefficients which have been derived from consistently estimated structural parameters. In an important paper by Klein (1960), he argued in a heuristic manner that as long as there exists no specification error the derived estimates of the reduced form coefficients will be more efficient than those obtained with no restriction procedure, consequently the prediction based on the former will be more efficient than that based on the latter. Goldberger, Nagar and Odeh (1961) gave the asymptotic covariance matrix of the two stage least squares (TSLS) induced restricted reduced form estimator. Later, Dhrymes (1973) showed that the TSLS induced restricted reduced form estimators are not necessarily asymptotically efficient relative to the unrestricted reduced form estimators. It should be also noticed that the induced restricted reduced form estimators are highly sensitive to specification errors since the entire system must be specified prior to carrying out the estimation procedure.

Nager and Sahay (1978) derived the exact and large noncentrality parameter asymptotic moments of the partially restricted reduced form predictor. This method was originally proposed by Amemiya (1966) and later analyzed by Kakwani and Court (1972).

On the other hand, Takeuchi (1972) proposed a predictor which is obtained by the use of the maximum likelihood estimates of the reduced form coefficients taking account of the overidentifying restrictions imposed on a relevant structural equation. This method is called the single equation method of prediction (SEMP) in analogy with the single equation method of estimation for the structural parameters. This predictor is known to be asymptotically efficient relative to the unrestricted reduced form predictor. Moreover it is not affected by specification errors in equations other than the relevant equation.

Employing Kadane's (1970, 1971) small-sigma asymptotics, Hasegawa (1976) derived the bias (to the order $\sigma^2$) and the mean square error (to the order $\sigma^4$) of the SEMP associated with the general $k$-class estimators. The resultant expressions of moments in Hasegawa (1976) are,
however, so complicated that it is not easy to compare the efficiencies among the SEMP.

In this paper, we derive the asymptotic expansions of the bias and the mean square error of the SEMP associated with the limited information maximum likelihood (LIML) and the TSLS estimators when the sample size increases. We refer to the former predictor as the LIML predictor and the latter as the TSLS predictor. Our approach is summarized as follows. Firstly the random variables of the predictor are approximated by other random variables close to the original variables in probability. The approximating random variables are derived from a Taylor series expansion in terms of $T^{-1/2}$. Then, the properties of the approximating random variables are examined in order to investigate the properties of the predictor. No claim is made that the moments of the approximating random variables are close to the moments of the original random variables. See Brown, Kadane and Ramage (1974).

The results are much simpler than those of Hasegawa (1976) and we are able to compare the efficiencies of predictors. Both the LIML and the TSLS predictors asymptotically dominate the partially restricted reduced form predictor as well as the unrestricted reduced form predictor for any values of parameters. The LIML predictor is unbiased to the order $T^{-1}$ while the TSLS predictor is biased in the order $T^{-1/2}$. The SEMP shares the common properties with the single equation method of estimation for structural parameters. In other words, large degrees of overidentifiability favor the LIML predictor, and the small standardized parameters of structural parameters $\beta$ in absolute values favor the TSLS predictor (Anderson (1974), and Fujikoshi, et al. (1982)).

Section 2 defines the model and the predictors, and states the assumptions. In Section 3, we describe the resulting expressions of the bias and the mean square error of the predictors in theorems, and make comparisons among predictors. Proofs of theorems are given in Appendix.

2. Model and Predictors
Let a single structural equation be

\[ y_1 = y_2^\beta + Z_1^\gamma + y, \]

where $y_1$ and $y_2$ are $T \times 1$ and $T \times G_1$ matrices, respectively, of observations on the endogenous variables, $Z_1$ is a $T \times K_1$ matrix of observations of the $K_1$ exogenous variables, $\beta$ and $\gamma$ are column vectors with $G_1$ and $K_1$ unknown parameters, and $u$ is a column vector of $T$ disturbances. We assume that (2.1) is the first equation in a simultaneous system of $G_1 + 1$ linear stochastic equations relating $G_1 + 1$ endogenous variables and $K (K = K_1 + K_2)$ exogenous variables. The reduced form of $y = (y_1, y_2)$ is defined as

\[ y = Z_1^\Pi + v = (Z_1, Z_2)(\pi_1^1, \pi_2^1) + (v_1, v_2), \]

where $Z_1$ is a $T \times K$ matrix of exogenous variables (full rank), $\pi_1^1 = (\pi_1^1, \pi_2^1)$ and $\Pi_2^1 = (\Pi_1^2, \Pi_2^2)$ are, respectively, $1 \times (K_1 + K_2)$ and $G_1 \times (K_1 + K_2)$ matrices of the reduced form coefficients, and $(v_1, v_2)$ is a $T \times (1 + G_1)$ matrix of disturbances. We make the following two assumptions about the model:

Assumption 1. The rows of $v$ are independently normally distributed, each row having mean $0$ and (non-singular) covariance matrix,
Assumption 2. The matrix \((\pi_{21}, \pi_{22})\) is of rank \(G_1\) and \(\pi_{22}\) is also of rank \(G_1\).

In order to relate (2.1) and (2.2) postmultiply (2.2) by \((1, -\beta')'\), then \(u = v_1 - V_2\beta, \gamma = \pi_{11} - \pi_{12}\beta,\) and

\[
(2.4) \quad \pi_{21} = \pi_{22}\beta.
\]

The components of \(u\) are independently normally distributed with mean 0 and variance

\[
(2.5) \quad \sigma^2 = \omega_{11} - 2\beta'\omega_{21} + \beta'\Omega_{22}\beta.
\]

Further, for convenience, we define

\[
(2.6) \quad L = K_2 - G_1,
\]

which is the degree of overidentification,

\[
(2.7) \quad \rho = \frac{1}{\sigma^2}(\omega_{11} - \omega_{12}\beta),
\]

which represents the covariance between the structural disturbance \(u\) and the reduced form disturbance \(v_1\) divided by the structural variance \(\sigma^2\),

\[
(2.8) \quad g_1 = \frac{1}{\sigma^2}(\omega_{21} - \Omega_{22}\beta) : G_1 \times 1,
\]

\[
(2.9) \quad C_1 = g_1 g_1' : G_1 \times G_1,
\]

\[
(2.10) \quad C_2 = \frac{1}{\sigma^2} \Omega_{22} - C_1 : G_1 \times G_1,
\]

\[
(2.11) \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \frac{1}{T}(Z_1Z_2)'(Z_1Z_2) : (K_1 + K_2) \times (K_1 + K_2),
\]

\[
(2.12) \quad Q_{11} = (\pi_{21}\pi_{22}, I)\pi_{22}^{-1} : G_1 \times G_1,
\]

where \(A_{22} = A_{2} - A_{21}A_{11}^{-1}A_{12}\). We note that the above notations are essentially conformable with those of Fujikoshi, et al. (1982). Additionally we define, for any matrix \(F\) with independent columns,

\[
(2.13) \quad P_F = F(F'F)^{-1}F', \quad \bar{P}_F = I - F(F'F)^{-1}F'.
\]

We further assume the following:

Assumption 3. \(A = 0(1)\).

We are interested in predicting the expected value of the endogenous variable appearing in the lefthand side of (2.1), conditionally on the values of a set of exogenous variables, say a \((K_1 + K_2)\)-dimensional vector \(z^0 = (z_1^0, z_2^0)'\). In other words, we shall predict the expected value of \(y^0_i, i.e. z^0_i\). where \(y^0_i\) has the form

\[
(2.14) \quad y^0_i = z^0_i \pi_1 + v^0_i = z^0_i \pi_1 + z^0_i \pi_21 + v^0_i,
\]

where \(v^0_i\) is assumed to be distributed with mean zero and variance \(\omega_{11}\), and independent of \((v_1, Y_2)\).

The *unrestricted reduced form estimator* of \(\pi_1\) in (2.2) yields a predictor of \(z^0_i\pi_1\) as

\[
(2.15) \quad \hat{y}_{UR} = z^0_i \hat{\pi}_1,
\]

where

\[
(2.16) \quad \hat{\pi}_1 = (Z'Z)^{-1}Z'y_1.
\]
Note that $\tilde{\pi}_1$ has a $K$-dimensional multivariate normal distribution with a mean vector $\pi_1$ and a covariance matrix $\omega_1(Z'Z)^{-1}$.

Amemiya (1966) and Kakwani and Court (1972) proposed the partially restricted reduced form estimator of $\tilde{\pi}_1$ as

(2.17) $\tilde{\pi}_{1,PR} = \left[ (Z'Z)^{-1}Z'Y_2 \right] \left[ \begin{array}{c} I_k \\ \hat{\beta} \end{array} \right],$

where $\hat{\beta}$, $\tilde{\pi}$ denote either the LIML or the TSLS estimators of $\beta$ and $\gamma$ respectively. This estimator brings the partially restricted reduced form predictor

(2.18) $y_{PR} = Z_0' \tilde{\pi}_{1,PR}.$

Nagar and Sahay (1978) derived the exact and large noncentrality asymptotic moments of (2.18). The equation (2.18) is also written as

(2.19) $\tilde{y}_{PR} = \tilde{\beta} + z_1' \tilde{\pi},$

where

(2.20) $\tilde{\beta}_2 = z_1' (Z'Z)^{-1}Z'Y_2.$

The single equation method of prediction (SEMP) proposed by Takeuchi (1972) is obtained by the use of the maximum likelihood estimator of $\pi_1$ maximizing the likelihood function of (2.2) subject to the overidentifying restriction (2.4):

(2.21) $\tilde{y}_{SE} = Z_0' \tilde{\pi}_{1,SE},$

where

(2.22) $\tilde{\pi}_{1,SE} = \tilde{\pi}_1 - \hat{\beta} \left[ \begin{array}{c} -(Z'Z)^{-1}Z'Y_2 \\ I_k \end{array} \right] \left[ Z_0' \tilde{P} Z_1 Z_2 \right]^{-1} Z_0' \tilde{P} Z_1 Y \left( \frac{1}{1 - \hat{\beta}} \right),$

and

(2.23) $\hat{\beta} = \frac{y_1' \tilde{P} Z Y (1 - \hat{\beta})}{y_1' \tilde{P} Z Y (1 - \hat{\beta})}. $

Note that $\hat{\rho}$ is a consistent estimator of $\rho$. We refer to $\tilde{y}_{LI}$ as $\tilde{y}_{SE}$ with the LIML estimator of $\beta$, and to $\tilde{y}_{TS}$ with the TSLS estimator. The SEMP is a weighted sum of two other predictors

(2.24) $\tilde{y}_{SE} = (1 - \hat{\rho}) \tilde{y}_{UR} + \hat{\rho} \tilde{y}_{PR}.$

This relation indicates that $\tilde{y}_{PR}$ and $\tilde{y}_{UR}$ are formally obtained by (2.21) with replacement of $\hat{\rho} = 1$ and $\hat{\rho} = 0$, respectively.

The reduced form of (2.2) satisfies the assumptions of Zellner's (1962) seemingly unrelated regressions (SUR). However, the SUR estimator of $\pi_1$ collapses to yield the unrestricted reduced form estimator of (2.16) since (2.2) has common regressors to all columns of $\Pi$.

3. Statement of Results

In this section, we first give the approximate formulae of the first and second moments of the partially restricted reduced form predictor and the SEMP when the sample size increases. Our results reveal that the SEMP asymptotically dominates both the unrestricted reduced form predictor and the partially restricted reduced form predictor for any values of parameters. Next, we compare the LIML predictor with the TSLS predictor based on the approximate mean square error. The proofs of Theorems 1 through 3 in this section are given in Appendix.
We define the prediction error associated with each predictor normalized by the square root of the sample size, for instance, as

\[ (3.1) \quad \tilde{e}_{UR} = \sqrt{T} (\hat{y}_{UR} - z_0\beta_1). \]

We define \( \tilde{e}_{SE} \) and \( \tilde{e}_{PR} \) by the righthand side of (3.1) with \( \hat{y}_{SE} \) and \( \hat{y}_{PR} \) in the place of \( \hat{y}_{UR} \) respectively. It is well known that \( \hat{y}_{UR} \) is an unbiased predictor of \( z_0\beta_1 \) and its mean square error is given by

\[ (3.2) \quad E\{e^2_{UR}\} = \omega_{11}z_0'\beta_1A^{-1}z_0. \]

It must be noticed that all the results in this section are calculated from the first three terms of the stochastic expansion of (3.1) in terms of \( T^{-1/2} \). The resulting moments do not necessarily mean the approximation to the exact finite sample moments of (3.1).

**Theorem 1:** As \( T \to \infty \), the mean square error of the partially restricted reduced form predictor defined by (2.18), to the order \( O(1) \), is given by

\[ (3.3) \quad AM_T\{e^2_{PR}\} = E\{\tilde{e}^2_{UR}\} - \sigma^2(2\rho - 1)z_0'\beta_1(A_{22,1}^{-1} - \Pi_{22Q_{11}}\Pi_{22Q_{11}})z_0^0, \]

where \( z_0^0 = z_0^0 - Z_0^0Z_1^0(Z_1^0Z_1^0)^{-1}z_0^0 \), and the first term of the righthand side is defined by (3.2), and \( AM_T(\cdot) \) stands for the mean in terms of the stochastic expansion of the prediction error up to \( O_p(T^{-1}) \).

The second term of (3.3) is positive if and only if \( \rho < 1/2 \). The partially restricted reduced form predictor is asymptotically efficient relative to the unrestricted reduced form predictor if and only if \( \rho > 1/2 \).

**Theorem 2:** As \( T \to \infty \), the biases of the single equation methods of prediction associated with the LIML and the TSLS estimators defined by (2.21), to the order \( T^{-1} \), are respectively given by

\[ (3.4) \quad AM_T\{e_{LI}\} = 0, \]

\[ (3.5) \quad AM_T\{e_{TS}\} = \frac{1}{\sqrt{T}} \sigma^2\rho Lz_0^0\Pi_{22Q_{11}}Q_{11}g_1. \]

It is remarkable that the LIML predictor is unbiased to the order \( T^{-1} \).

**Theorem 3:** As \( T \to \infty \), the mean square errors of the single equation methods of prediction associated with the LIML and the TSLS estimators defined by (2.21), to the order \( T^{-1} \), are respectively given by

\[ (3.6) \quad AM_T\{e^2_{LI}\} = \omega_{11}z_0'\beta_1A^{-1}z_0 - \sigma^2\rho^2z_0'\beta_1(A_{22,1}^{-1} - \Pi_{22Q_{11}}\Pi_{22Q_{11}})z_0^0 \]

\[ + \frac{1}{T} \sigma^4z_0'\beta_1(\rho^2L\Pi_{22Q_{11}}^2C_2Q_{11}\Pi_{22Q_{11}} + \xi_{LI}(A_{22,1}^{-1} - \Pi_{22Q_{11}}\Pi_{22Q_{11}})z_0^0, \]

\[ (3.7) \quad AM_T\{e^2_{TS}\} = \omega_{11}z_0'\beta_1A^{-1}z_0 - \sigma^2\rho^2z_0'\beta_1(A_{22,1}^{-1} - \Pi_{22Q_{11}}\Pi_{22Q_{11}})z_0^0 \]

\[ + \frac{1}{T} \sigma^4z_0'\beta_1(\Pi_{22Q_{11}}[\rho^2L^2C_1 - \rho^2LC_2 - 2\rho L(g_1\beta' \beta + C_2\beta g_1)]Q_{11}\Pi_{22Q_{11}} \]

\[ + \xi_{TS}(A_{22,1}^{-1} - \Pi_{22Q_{11}}\Pi_{22Q_{11}})z_0^0, \]

where

\[ (3.8) \quad \xi_{LI} = \rho^2tr(C_2Q_{11}) + \beta'(C_2 + C_2Q_{11}C_2)\beta, \]

\[ (3.9) \quad \xi_{TS} = \xi_{LI} + 2\rho L\beta'C_2Q_{11}g_1. \]

The first two terms in (3.6) and (3.7) were derived by Takeuchi (1972). The difference of the mean square errors of the LIML and the TSLS predictors appears in the terms of the order \( T^{-1} \).
Corollary 1: The single equation method of prediction dominates both the unrestricted reduced form predictor and the partially restricted reduced form predictor in the following sense:

\[
(3.10) \lim_{T \to \infty} \{AM_T[\hat{e}_{UR}^2] - AM_T[\hat{e}_{SE}^2]\} \geq 0,
\]

\[
(3.11) \lim_{T \to \infty} \{AM_T[\hat{e}_{PR}^2] - AM_T[\hat{e}_{SE}^2]\} \geq 0.
\]

The equality holds if and only if \( \rho = 0 \) in (3.10), and if and only if \( \rho = 1/2 \) in (3.11).

Proof: It is obvious from (3.2), (3.3), (3.6), and (3.7). (Q.E.D.)

Hasegawa (1976) obtained

\[
(3.12) \lim_{\sigma \to 0} \{AM_P[\hat{e}_{UR}^2] - AM_P[\hat{e}_{SE}^2]\} \geq 0,
\]

if and only if

\[
(3.13) \frac{\sigma^2}{\omega_{11}} \rho^2 \geq \frac{1}{T - K + 1}.
\]

The righthand side of (3.13) is positive because the sample size is fixed in small-sigma approach.

We further compare the mean square errors of two single equation methods of prediction. The LIML predictor is preferred to the TSLS predictor for large \( L \) because a term of \( L^2 \) is included in (3.7) but not in (3.6). However, it is not easy to extract more concrete properties of the mean square errors of the LIML and the TSLS predictors. In order to make more precise comparison we rewrite Theorem 3 for the case of two including endogenous variables \((G_1 = 1)\) in the structural equation (2.1). Anderson (1974) first derived the asymptotic expansion of the distribution of the LIML estimator of \( \beta \) for \( G_1 = 1 \), and compared the mean square error of the LIML estimator with that of the TSLS estimator in terms of \( L \) and the so-called standardized structural parameter \( \alpha \), which is defined as

\[
(3.14) \quad \alpha = \frac{1}{\sqrt{\Omega}} (\omega_{22} \beta - \omega_{21}).
\]

His results reveal that large \( L \) and large \( \alpha^2 \) favor the LIML estimator. We compare two predictors in terms of \( L \) and \( \alpha \). Let us define

\[
(3.15) \quad \theta = \frac{z_0' \Pi_{22} Q_{11} \Pi_{22} z_0}{z_{21}' A_{21}^{-1} z_{21}}.
\]

We note \( 0 \leq \theta \leq 1 \) since a \( K_2 \times K_2 \) matrix \( A_{21}^{-1} = \Pi_{22} Q_{11} \Pi_{22} \) is positive semi-definite.

When \( G_1 = 1 \), Theorem 3 is written as follows.

Corollary 2: When \( G_1 = 1 \), as \( T \to \infty \), the mean square errors of the single equation methods of prediction associated with the LIML and the TSLS estimators, to the order \( T^{-1} \), are respectively given by

\[
(3.16) \quad AM_T[\hat{e}_{LI}^2] = \omega_{11} \sigma^2 \theta^2 \sigma^{-1} A_0^{-1} z_0 - \sigma^2 \rho^2 z_{02} A_{21} z_{21} (1 - \theta)
\]

\[
+ \frac{1}{T} \sigma^2 z_{02} A_{21}^{-1} \{ \rho^2 L Q_{11} C_2 \theta + x_{LI} (1 - \theta) \},
\]

\[
(3.17) \quad AM_T[\hat{e}_{TS}^2] = AM_T[\hat{e}_{LI}^2] + \frac{1}{T} \sigma^4 \rho^2 L C_2 Q_{11} z_{02} A_{21} z_{21}.
\]

\[
\left\{ \left[ L a^2 - \left( 6 - \frac{4}{\rho} \right) \right] \theta + 2 \left( 1 - \frac{1}{\rho} \right) (1 - \theta) \right\}.
\]
Theorem 4: When $G_1 = 1$, the following inequality holds
\[ \lim_{T \to \infty} T \{ AM_T \{ \hat{\varepsilon}_T^2 \} - AM_T \{ \tilde{\varepsilon}_T^2 \} \} \geq 0, \]
if
\[ La^2 \geq 6 - \frac{4}{\rho} \quad \text{and} \quad \rho < 0 \quad \text{or} \quad \rho \geq 1, \]
the inequality of (3.18) is reversed if
\[ La^2 \leq 6 - \frac{4}{\rho} \quad \text{and} \quad \frac{2}{3} \leq \rho \leq 1. \]

We can say that small $\alpha^2$ favors the TSLS predictor and large $L$ favors the LIML predictor. This is compatible with the properties of the LIML and the TSLS estimators of $\beta$. Our conclusion, however, depends on the unknown parameter $\rho$. When $0 < \rho < 2/3$, it is impossible to deduce the values of the parameters $L$, $\alpha$ and $\rho$ which satisfy (3.18) for any values of $\theta (0 \leq \theta \leq 1)$.

4. Concluding Remarks

We derived the asymptotic expansions of the moments of the single equation methods of prediction (SEMP) associated with the LIML and the TSLS estimators of a structural equation when the sample size increases. The SEMP asymptotically dominates both the unrestricted reduced form predictor and the partially restricted reduced form predictor, the latter was first proposed by Amemiya (1966), in the sense of the mean square error (m.s.e.). The SEMP utilizes the LIML estimator of the reduced form coefficients taking account of the overidentifying restrictions. The asymptotic m.s.e. of $\hat{\pi}_{1, SE}$ is given by
\[ AM_T \{ \sqrt{T} (\hat{\pi}_{1, SE} - \pi_1) \sqrt{T} (\hat{\pi}_{1, SE} - \pi_1)' \} = \omega_{11} \{(A^2 - r^2(-A_{21}A_{11}^{-1}, I_{21})'(A_{21}A_{11}^{-1}, I_{21}')(A_{22}A_{12}^{-1}, I_{22})'(-A_{21}A_{11}^{-1}, I_{21})\}, \]
where $r = \text{cov} (u_1, v_1)/(\sigma \sqrt{\omega_{11}})$ stands for the correlation coefficient between $u_1$ and $v_1$. Since the first term represents the m.s.e. of the ordinary least squares (OLS) estimator of $\pi_1$, (4.1) shows that the m.s.e. of the LIML estimator is not greater than that of the OLS estimator. This is why the SEMP dominates the unrestricted reduced form predictor.

The LIML predictor is unbiased to the order $T^{-1}$ but the TSLS predictor is biased in the order $T^{-1/2}$. The m.s.e.s of the LIML and the TSLS predictors are more complicated than those of the LIML and the TSLS estimators of the structural parameters. It is not easy to compare the LIML predictor with the TSLS predictor in general for the case $G_1 > 1$. Instead, in two including endogenous variables case ($G_1 = 1$) we are able to compare the LIML and the TSLS predictors in terms of $L$ and $\alpha$ as Anderson (1974) did for comparing the LIML and the TSLS estimators of $\beta$. Our result implies that large $L$ and large $\alpha^2$ favor the LIML predictor. The above properties are compatible with those of the LIML and the TSLS estimators.

The problem of predicting the lefthand side endogenous variable of a structural equation is essentially the same as the problem of estimating the reduced form coefficients. Asymptotic expansions of the distributions of $\hat{\pi}_{1, SE}$ may shed more illuminating insights into the distributional properties of $\hat{\pi}_{1, SE}$. Since we explicitly derived the stochastic expansions of $\hat{\pi}_{1, SE}$, it may
not be difficult to obtain the approximate distributions of \( \hat{\pi}_{1,SE} \). We will discuss these problems in another article.

The conclusions of this paper crucially depend on the assumption of permitting no lagged endogenous variables in the model. The above-mentioned assumption is restrictive from the practical point of view. Even if there exist lagged endogenous variables in the model, Colorally 1 may still hold under certain regularity conditions. Because (3.10) and (3.11) are valid in the limit as the sample size goes to infinity, the usual asymptotic theory for the time series analysis is applicable. See, for instance, Chapter 5 of Anderson (1971). However, it is very hard to develop the higher order asymptotic expansions of the moments of the SEMP for the model having lagged endogenous variables. Even the higher order approximations to the moments of neither LIML nor TSLS estimators have been derived for these models in spite of their importance. We anticipate research works on the dynamic model including lagged endogenous variables.

Appendix  Proofs of Theorems 1, 2, and 3

This appendix gives proofs of Theorems 1, 2, and 3 in Section 3. The distributional properties of the predictors considered in this paper depends solely on the estimators of the reduced form coefficients. To make clear the role of the estimators of the reduced form coefficients, we explicitly derive the moments of the estimators of \( \pi_1 \) in the reduced form. The basic idea for proving theorems is, first to obtain the stochastic expansions of the sampling errors to appropriate order of \( T^{-1/2} \) by the full use of the results by Fujikoshi, et al. (1982), next to evaluate the expectations of the approximating random variables. Since the calculations are rather cumbersome, we only show the outline of the proofs.

The sampling error of \( \hat{\pi}_{1,SE} \) defined by (2.22) is written as

\[
(A.1) \quad \hat{\pi}_{1,SE} = \sqrt{T} (\hat{\pi}_{1,SE} - \pi_1)
\]

\[
= (A_{11}^{1/2}, 0)'u_{21} + D' \tilde{A}_{21}^{-1/2}u_{11} - \hat{\rho}D' \tilde{A}_{21}^{-1/2} \left\{ \xi_1 - \left( F + \frac{1}{\sqrt{T}} U_{12} \right) \tilde{\beta} \right\},
\]

where

\[
(A.2) \quad D = (-A_{21}A_{11}^{-1}, I_{K_2}) \quad : \quad K_2 \times (K_1 + K_2),
\]

\[
(A.3) \quad F = \tilde{A}_{21}^{1/2}U_{22} \quad : \quad K_2 \times G_1,
\]

\[
(A.4) \quad \xi_1 = u_{11} - U_{12}\hat{\beta} \quad : \quad K_2 \times 1,
\]

\[
(A.5) \quad \xi_{21} = (Z'Z)^{-1/2}Z' \xi_{11} \quad : \quad K_1 \times 1,
\]

\[
(A.6) \quad \xi_{11}, U_{12} = (Z' \tilde{P} z_1 Z' z_2)^{-1/2}Z' \tilde{P} z_1 (u_1, V_2) \quad : \quad K_2 \times (1 + G_1),
\]

\[
(A.7) \quad \xi_{21} = \sqrt{T} (\hat{\beta} - \hat{\beta}),
\]

and \( \hat{\beta} \) is either the LIML or the TSLS estimator of \( \beta \). The sampling errors of \( \hat{\pi}_1 \) and \( \hat{\pi}_{1,PR} \) are obtained by (A.1) with the replacement of \( \hat{\rho} \) by zero and unity respectively. Fujikoshi, et al. (1982) gave the stochastic expansions of \( \hat{\pi}_\beta \) as

\[
(A.8) \quad \hat{\pi}_\beta = \varepsilon^{(0)} + \frac{1}{\sqrt{T}} \varepsilon^{(1)} + \frac{1}{T} \varepsilon^{(2)} + O_p(T^{-3/2}),
\]

for both the LIML and the TSLS estimators. We do not reproduce each term of \( \varepsilon^{(i)}(i = 0, 1, 2) \) for the sake of space.
Lemma 1: As $T \to \infty$, we have the following Taylor series expansion of (A.1) in terms of $T^{-1/2}$.

\begin{align}
\frac{1}{T - K} Y' \bar{P} Y &= \Omega + \frac{1}{\sqrt{T}} \psi' \begin{pmatrix} w_{11} & w_{12} \\ w_{12}' & W_{22} \end{pmatrix} \psi \\
+ \frac{1}{T} \psi' \begin{pmatrix}
\frac{1}{3} (w_{11}^2 - 2) & \frac{1}{2} w_{11} w_{12} \\
\frac{1}{2} w_{11} w_{12}' & O_p(1)
\end{pmatrix} \psi,
\end{align}

where

\begin{align}
\psi &= \sigma \begin{pmatrix}
\rho & g_1' \\
\beta & \xi_1
\end{pmatrix},
\end{align}

and $w_{11}, w_{12}, W_{22}$ are all independently distributed with $w_{11} \sim N(0, 2), w_{12} \sim N(0, \sigma^2 C_2)$, and $E(W_{22}) = Q$.

**Proof:** Let us define two random matrices $S$ and $A$ as

\begin{align}
S &= \sqrt{\Psi} \Psi^{-1} : T \times (1 + G_1), \\
A &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = S' \bar{P} S : (1 + G_1) \times (1 + G_1).
\end{align}

The matrix $A$ of (A.12) is used only for proving Lemma 1, and is different from the definition in (2.11). There should be no confusion. Then, each row of $S$ is identically independently distributed as $N(Q, \Sigma)$, where

\begin{align}
\Sigma &= \begin{pmatrix} 1 & 0 \\ O & C_2 \end{pmatrix},
\end{align}

and $A$ has a Wishart distribution $W_{1+G_1}(T - K, \Sigma)$. Moreover, a random matrix $A_{22,1} = A_{22} - A_{21} A_{11}^* A_{12}$ has $W_{G_1}(T - K - 1, C_2)$ and is independent of $A_{11}$ and $A_{12}$. The conditional distribution of $A_{12}$, given $A_{11}$, is $N(Q, C_2 A_{11})$, and $A_{11}$ is distributed as $\chi^2(T - K)$. (See Muirhead (1982) p.93, for instance.) Hence, $A_{12}^* = A_{11}^{-1/2} A_{12}$ is independent of $A_{11}$ and has $N(Q, C_2)$. Finally, we have the following stochastic expansions to the order $T^{-1}$:

\begin{align}
\frac{1}{T - K} A_{11} &= 1 + \frac{1}{\sqrt{T}} w_{11} + \frac{1}{T} \frac{1}{3} (w_{11}^2 - 2), \\
\frac{1}{T - K} A_{12} &= \frac{1}{T - K} A_{12}^* A_{12}^* = \frac{1}{\sqrt{T}} w_{12} + \frac{1}{T} \frac{1}{2} w_{11} w_{12}, \\
\frac{1}{T - K} A_{22} &= \frac{1}{T - K} \{A_{22,1} + A_{12}^* A_{12}^* \} = C_2 + \frac{1}{\sqrt{T}} W_{22} + O_p(T^{-1}).
\end{align}

Noting $Y' \bar{P} Y = \psi' \Gamma \psi$ , we obtain the required result. (Q.E.D.)

Lemma 2: As $T \to \infty$, we have the stochastic expansion of $\hat{\rho}$ to the order $T^{-1}$,

\begin{align}
\hat{\rho} = \rho^{(0)} + \frac{1}{\sqrt{T}} \rho^{(1)} + \frac{1}{T} \rho^{(2)},
\end{align}

where

\begin{align}
\rho^{(0)} = \rho,
\end{align}
The Economic Studies Quarterly

(A.19) \( \rho^{(1)} = w_{12} \beta + \left[ 2 \rho g' - \left( \frac{1}{\sigma^2} w_{12} \right) \xi^{(0)}_\beta, \right. \)

(A.20) \( \rho^{(2)} = -\frac{1}{2} w_{11} w_{12} \beta + \left( w_{11} \left( \frac{1}{\sigma^2} w_{12} \right) - \rho (w_{11} g'_1 - w_{12}) + \beta' (w_{12} \beta - W_{22}) \right) \xi^{(0)}_\beta + \left[ 2 \rho g' - \left( \frac{1}{\sigma^2} w_{12} \right) \xi^{(1)}_\beta + \xi^{(0)}_\beta \right] \rho (3 \zeta_1 - \zeta_2) - 2 g_1 \left( \frac{1}{\sigma^2} w_{12} \right) \xi^{(0)}_\beta. \)

Proof: From (A.10) of Kunitomo, et al. (1983), the denominator of (2.23), divided by \( \sigma^2 (T - K) \), is written as

(A.21) \( 1 + \frac{1}{\sqrt{T}} (w_{11} - 2 g'_1 \xi^{(0)}_\beta) + \frac{1}{T} \left( \frac{1}{3} (w_{11}^2 - 2) \right) - 2 \left( (w_{11} g'_1 + w_{12}) \xi^{(0)}_\beta + g'_1 \xi^{(1)}_\beta + \xi^{(0)}_\beta (\zeta_1, \zeta_2) \xi^{(0)}_\beta \right) + O_p(T^{-3/2}). \)

Similarly the numerator, divided by \( \sigma^2 (T - K) \), is written as

(A.22) \( \rho + \frac{1}{\sqrt{T}} \left( w_{11} \rho + w_{12} \beta - \left( \frac{1}{\sigma^2} w_{12} \right) \xi^{(0)}_\beta \right) + \frac{1}{T} \left( \frac{1}{3} (w_{11}^2 - 2) \rho + \frac{1}{2} w_{11} w_{12} \beta \right) - \left[ \rho (w_{11} g'_1 + w_{12}) + \beta' (w_{12} g'_1 + W_{22}) \right) \xi^{(0)}_\beta - \left( \frac{1}{\sigma^2} w_{12} \right) \xi^{(1)}_\beta \right] + O_p(T^{-3/2}). \)

We used Lemma 1 to derive (A.21) and (A.22). Then, a Taylor series expansion of the inverse of (A.21) in terms of \( T^{-1/2} \), multiplied by (A.22), yields (A.17). (Q.E.D.)

Substituting (A.8) and (A.17) into (A.1), we have the following lemma.

Lemma 3: As \( T \rightarrow \infty \), we have the stochastic expansion of (A.1) to the order \( T^{-1} \),

(A.23) \( \tilde{e}_{n,SE} = \xi^{(0)}_{\tilde{e}_{n,SE}} + \frac{1}{\sqrt{T}} \xi^{(1)}_{\tilde{e}_{n,SE}} + \frac{1}{T} \xi^{(2)}_{\tilde{e}_{n,SE}}, \)

where

(A.24) \( \xi^{(0)}_{\tilde{e}_{n,SE}} = (A_{11}^{-1/2}, Q) u_{11} + D' A_{22}^{-1/2} (u_{11} - \tilde{F}_2 \xi_1^{(0)}), \)

(A.25) \( \xi^{(1)}_{\tilde{e}_{n,SE}} = D' A_{22}^{-1/2} (\xi_1^{(1)} + U_{12} \xi_2^{(0)}) \rho^{(0)} - \tilde{F}_2 \xi_1^{(1)} \rho^{(1)}), \)

(A.26) \( \xi^{(2)}_{\tilde{e}_{n,SE}} = D' A_{22}^{-1/2} (\xi_1^{(2)} + U_{12} \xi_2^{(1)}) \rho^{(0)} + (\xi_1^{(1)} + U_{12} \xi_2^{(0)}) \rho^{(1)} - \tilde{F}_2 \xi_1^{(2)} \rho^{(2)}. \)

The stochastic expansion of \( \tilde{e}_{n,PR} \) is obtained by \( \tilde{e}_{n,SE} \) with \( \rho^{(0)} = 1, \rho^{(1)} = 0, \) and \( \rho^{(2)} = 0 \) in (A.23), and \( \tilde{e}_{n,UR} \) is identical to the first term of (A.23) with \( \rho^{(0)} = 0. \)

Lemma 3 plays the fundamental role for deriving the moments in theorems. The notations \([\cdot \cdot \cdot]^*\) is used to denote a single expression, valid for both the LIML and the TSLS estimators, the upper part referring to the LIML and the lower part referring to the TSLS estimators.

Proof of Theorem 1: Noting that

(A.27) \( \tilde{e}_{n,PR} = (A_{11}^{-1/2}, Q) u_{21} + D' A_{22}^{-1/2} (u_{11} - \tilde{F}_2 \xi_1^{(0)}) + O_p(T^{-1/2}), \)

we have

(A.28) \( AM_T \{ \tilde{e}_{n,PR}, \tilde{e}_{n,PR}' \} = \omega_{11} A^{-1} + \sigma^2 (1 - 2 \rho) D' (A_{22}^{-1} - \Pi_{22} Q_{11} \Pi_{12}^') D. \) (Q.E.D.)

Proof of Theorem 2: Since the terms of the order of \( T^0 \) and \( T^{-1} \) in the righthand side of (A.23) include only odd orders of polynomials of normal random variables with means zero, the possible
nonzero expectations appear in the terms of the order $T^{-1/2}$, which is

\begin{equation}
E\left\{ e^{(1)\text{SE}}_x \right\} = \sigma^2 \rho \begin{pmatrix} 0 & \star \\ L & \end{pmatrix} D^T \Pi_{22} Q_{11} q_1.
\end{equation}

This leads to Theorem 2. (Q.E.D.)

**Proof of Theorem 3:** Taking the expectations of each term of the square of the righthand side in (A.23) and making cumbersome calculations, we obtain the followings:

\begin{equation}
E\left\{ e^{(0)\text{SE}}_x e^{(0)\text{SE}}_x^T \right\} = \sigma^2 \rho^2 D^T (A^{-1}_{22,1} - \Pi_{22} Q_{11} \Pi_{22}) D,
\end{equation}

\begin{equation}
E\left\{ e^{(1)\text{SE}}_x e^{(0)\text{SE}}_x^T \right\} = 0,
\end{equation}

\begin{equation}
E\left\{ e^{(1)\text{SE}}_x e^{(1)\text{SE}}_x^T \right\} = \sigma^4 D^T \left( \rho^2 L \Pi_{22} Q_{11} C_2 Q_{11} \Pi_{22} + \rho L (A^{-1}_{22,1} - \Pi_{22} Q_{11} \Pi_{22}) \right) D
\end{equation}

\begin{equation}
+ \sigma^4 \rho^2 \begin{pmatrix} 0 & \star \\ L(L+2) & \end{pmatrix} D^T \Pi_{22} Q_{11} C_1 Q_{11} \Pi_{22} D,
\end{equation}

\begin{equation}
E\left\{ e^{(2)\text{SE}}_x e^{(0)\text{SE}}_x^T \right\} = -\sigma^4 \rho \begin{pmatrix} 0 & \star \\ L & \end{pmatrix}^* D^T \left( \rho \Pi_{22} Q_{11} (C_1 + C_2) Q_{11} \Pi_{22} \right.
\end{equation}

\begin{equation}
+ \left. \Pi_{22} Q_{11} (g_1 D C_2 + C_2 D g_1^T) Q_{11} \Pi_{22} - \beta^T C_2 Q_{11} q_1 (A^{-1}_{22,1} - \Pi_{22} Q_{11} \Pi_{22}) D \right).
\end{equation}

For deriving the above expectations we use the fact that a random matrix $U_{12} - 2 \tilde{q}_1^T \tilde{q}_1$ is independently distributed from $\tilde{q}_1$. Theorem 3 directly follows from (A.30), (A.31), (A.32), and (A.33). (Q.E.D.)

First draft received March 20, 1987; final draft accepted November 9, 1987.

**REFERENCES**


