1978

Generalized Two Stage Least Squares Estimators for a Structural Equation with Both Fixed and Random Coefficients

Baldev Raj

V. K. Srivastava

Aman Ullah

Follow this and additional works at: <https://ir.lib.uwo.ca/economicsresrpt>

Part of the [Economics Commons](#)

Citation of this paper:

Raj, Baldev, V. K. Srivastava, Aman Ullah. "Generalized Two Stage Least Squares Estimators for a Structural Equation with Both Fixed and Random Coefficients." Department of Economics Research Reports, 7835. London, ON: Department of Economics, University of Western Ontario (1978).

RESEARCH REPORT 7835
GENERALIZED TWO STAGE LEAST SQUARES
ESTIMATORS FOR A STRUCTURAL EQUATION
WITH BOTH FIXED AND RANDOM COEFFICIENTS

by

Baldev Raj
V. K. Srivastava
Aman Ullah

November 1978

GENERALIZED TWO STAGE LEAST SQUARES ESTIMATORS FOR A
STRUCTURAL EQUATION WITH BOTH FIXED AND RANDOM COEFFICIENTS

by

Baldev Raj^{*}

V. K. Srivastava^{**}

Aman Ullah^{***}

^{*}Universities of Wilfrid Laurier and Western Ontario

^{**}University of Lucknow, India

^{***}University of Western Ontario

GENERALIZED TWO STAGE LEAST SQUARES ESTIMATORS FOR A STRUCTURAL EQUATION WITH BOTH FIXED AND RANDOM COEFFICIENTS

by

Baldev Raj, V. K. Srivastava and Aman Ullah¹

1. INTRODUCTION

The problem of statistical inference in the context of classical regression models with random coefficients have been analyzed by Hildreth and Houck [1968], Rubin [1950], Theil and Mennes [1959], Swamy [1970], Zellner [1966] and others. However, attempts to solve the identification and estimation problems connected with systems of simultaneous linear structural equations with random parameters are few. For example, in his study of the Cobb-Douglas production functions, Nerlove [1965] developed a simultaneous equation model with random coefficients but did not solve completely the problems of identification and estimation of such models. Zellner [1969], gave conditions under which the usual macro two stage least squares estimator is a consistent estimator for the mean of individual's coefficient vectors but did not consider the case in which the coefficient vector differs from observation

¹The authors are grateful to an anonymous referee for his helpful comments and for suggesting the proofs of the Lemma and the Theorem of this paper. These proofs are more rigorous than those presented by the authors in an earlier version of this paper. The authors are also thankful to P. A. V. B. Swamy with whose help the paper is brought to its present form. They have also benefited from discussions on the subject matter of this paper with A. Zellner, R. D. Terrell and D. Héndry. The usual caveat with respect to responsibility holds.

to observation when one set of observations is available. Kelejian [1974] showed how the identifiability rules for simultaneous equations models with fixed coefficients carry over to random coefficients simultaneous equations models but his conditions are very restrictive and hold only in some special cases.

In this paper, we show how it is possible to estimate consistently a single structural equation in which the coefficients of exogenous variables are random and the coefficients of endogenous variables are fixed. Our specifying assumptions may be justified on the ground that the policy makers have control over some exogenous variables and the coefficients of exogenous variables are not invariant with respect to changes in policy, see Sargent [1976]. In the case where the coefficients of endogenous variables are fixed and the coefficients of exogenous variables are random, the problem of estimation is considerably simplified but remains nontrivial. The estimators we develop in this paper take the form of a modified two stage least squares (TSLS) estimator. It would be possible to pursue this line of thinking further and to develop an analogue of three stage least squares estimators for our model. But, we do not develop it here.

In Section 2 of this paper the model and the estimators of its parameters are described. Section 3 presents the asymptotic properties of the estimators developed in Section 2. Then a numerical example is given in Section 4. Finally, conclusions and suggestions for further research are presented in Section 5.

2. THE MODEL AND ESTIMATORS

2.1 Model Specification

Consider the possibilities for estimation of any single equation in a complete system of M structural equations in M jointly dependent and K exogenous variables². Without loss of generality, the first equation may be selected out of this system and written as

$$(2.1) \quad y_{11r} = \sum_{i=2}^{m+1} y_{1ir} \beta_{1i} + \sum_{j=1}^k x_{1jr} \gamma_{1j}(r) \quad \bar{r} = 1, 2, \dots, N$$

where y_{11r} is the r-th observation on the first jointly dependent variable, y_{1ir} and x_{1jr} are the r-th observations on the right hand side i-th endogenous and j-th exogenous variables respectively, and β_{1i} and $\gamma_{1j}(r)$ are the coefficients in this equation. We subsume the usual disturbance term in the equation into the varying intercept term by assuming that the elements of $x_{11r} = 1$ for every r.

Equation (2.1) can be compactly written as:

$$(2.2) \quad y_1 = Y_1 \beta_1 + \text{Diag}[x'_{11}, x'_{12}, \dots, x'_{1N}] [\gamma_1(1), \gamma_1(2), \dots, \gamma_1(N)]'$$

$$= Y_1 \beta_1 + D_{x_1} \gamma_1$$

where y_1 is a $N \times 1$ vector of observations on a left hand endogenous variable, Y_1 is a $N \times m$ matrix of observations on the other endogenous variables which appear in the equation, x'_{1r} is the r-th row of a $N \times k$ matrix of observations

²For simplicity of exposition, we assume that lagged endogenous variables do not appear in equation (2.1). The results established in this paper are true even when this assumption is not true, provided some additional conditions hold. It is easy to formulate these additional conditions.

on k exogenous variables that appear in the equation, and β_1 and $\gamma_1(r)$ are coefficient vectors of dimensions $m \times 1$ and $k \times 1$ respectively. Further, D_{x_1} and γ_1 are of order $N \times Nk$ and $Nk \times 1$ respectively.

We assume that the elements of β_1 are fixed while the elements of $\gamma_1(r)$ are random such that

$$(2.3) \quad \gamma_1(r) = \bar{\gamma}_1 + \epsilon_1(r) \quad r = 1, 2, \dots, N$$

where $\bar{\gamma}_1$ is a $k \times 1$ vector with nonstochastic elements and $\epsilon_1(r)$ is a $k \times 1$ discrepancy vector.

The assumptions concerning $\epsilon_1(r)$ are:

Assumption 1:

$$(i) \quad E \epsilon_1(r) = 0 \quad r = 1, 2, \dots, N$$

$$(2.4) \quad (ii) \quad E [\epsilon_{1j}(r) \epsilon_{1j'}(r')] = \begin{cases} \sigma_j & \text{for } j = j' \text{ and } r = r' \\ 0 & \text{otherwise,}^3 \end{cases}$$

where $\epsilon_{1j}(r)$ is the j -th element of $\epsilon_1(r)$, $j=1, 2, \dots, k$.

(iii) The elements of D_{x_1} are independent of the elements of $\epsilon_1(r)$.

Substituting (2.3) in (2.2), gives

$$(2.5) \quad y_1 = Y_1 \beta_1 + X_1 \bar{\gamma}_1 + w_1$$

where X_1 is a $N \times k$ matrix of observations on k exogenous variables and

$$(2.6) \quad w_1 = D_{x_1} [\epsilon'_1(1), \epsilon'_1(2), \dots, \epsilon'_1(N)]' = D_{x_1} \epsilon_1$$

Clearly, under Assumption 1

$$(2.7) \quad E w_1 = 0$$

$$E w_1 w_1' = \Omega = \text{Diag}(\Omega_1, \Omega_2, \dots, \Omega_N)$$

³Assumption 1(ii) is maintained throughout this paper for simplicity. The analysis of the paper could be extended to the case where $E(\epsilon_1(r) \epsilon_1'(r))$ is nondiagonal (see Swamy and Mehta [1975]).

where

$$(2.8) \quad \Omega_r = \sum_{j=1}^k x_{1jr}^2 \theta_j > 0 \quad r = 1, 2, \dots, N$$

Finally, we can write equation (2.5) as

$$(2.9) \quad y_1 = A_1 \delta_1 + w_1$$

where $A_1 = (Y_1 : X_1)$ and $\delta_1 = \begin{pmatrix} \beta_1 \\ \vdots \\ \gamma_1 \end{pmatrix}$.

A necessary and sufficient condition for the identifiability of β_1 and γ_1 in equation (2.9) is exactly as it would be in the fixed coefficients version of (2.1) (see Kelejian [1974]).⁴

Assumption 2: Equation (2.9) is identifiable.

2.2 Estimators

The TSLS estimator of δ_1 which does not take account the heteroscedasticity⁵ of w_1 is

⁴Kelejian [1974] suggested an approach to determining identifiability in simultaneous structural models if the coefficients of right hand endogenous variables are also random.

⁵It may be noted that the model (2.9) is not the usual kind of regression model with heteroscedastic errors. The latter leads to the problem of 'incidental' parameters if the number of distinct error variances varies with the sample size.

$$(2.10) \quad \hat{\delta}_1 = [A_1' P A_1]^{-1} A_1' P y_1$$

where $P = X(X'X)^{-1}X'$, X being a $N \times K$ matrix of observations on all the K exogenous variables in the system. In what follows, we treat X as a fixed matrix.

Assumption 3: The rank of matrix X is $K < N$.

It is worth mentioning here that if the rank of X is $q \leq N < K$, then one can use the generalized inverse solution suggested by Swamy and Holmes [1971].

It is well known that under general conditions the TSLS estimator $\hat{\delta}_1$ is consistent and the second order moment matrix of the asymptotic distribution of $\sqrt{N}(\hat{\delta}_1 - \delta_1)$ is given by

$$(2.11) \quad \text{Plim}_{N \rightarrow \infty} N [A_1' P A_1]^{-1} A_1' P \Omega P A_1 [A_1' P A_1]^{-1}$$

It is easy to verify that (2.11) reduces to the conventional result

$$(2.12) \quad \sigma^2 \text{Plim}_{N \rightarrow \infty} N [A_1' P A_1]^{-1}$$

when $\Omega = \sigma^2 I_N$.

Now since $E w_1 w_1' = \Omega \neq \sigma^2 I_N$, an efficient TSLS estimator of δ_1 in (2.9)

can be obtained as follows. Given that Ω is positive definite, we can write

$$(2.13) \quad X' \Omega^{-1} y_1 = X' \Omega^{-1} A_1 \delta_1 + X' \Omega^{-1} w_1.$$

Straight-forward application of Aitken's generalized least squares (see Theil et al. [1961], sec. 6.c.4., pp. 344-355) to (2.13) would then lead to estimation of δ_1 as

$$(2.14) \quad \tilde{\delta}_1 = (A_1' R A_1)^{-1} A_1' R y_1$$

where $R = \Omega^{-1} X(X' \Omega^{-1} X)^{-1} X' \Omega^{-1}$. 6,7

The estimator $\tilde{\delta}_1$ is, however, of limited use because it involves the θ_j 's which are unknown. A solution to the problem is therefore to replace these unknowns by their sample estimates. For this purpose we propose a few consistent methods of estimating these unknown parameters. First, we define $\theta = (\theta_1, \dots, \theta_k)$, $M_1 = I - X_1(X_1' X_1)^{-1} X_1'$, $\dot{M}_1 = M_1 * M_1$ and $\dot{X}_1 = X_1 * X_1$ where $*$ represents the Hadamard products, see Rao [1973, p. 30].

⁶Note that if $\theta_j = 1$ for every j in Ω then $\tilde{\delta}_1$ becomes

$$[A_1' \Omega^{*-1} X(X' \Omega^{*-1} X)^{-1} X' \Omega^{*-1} A_1]^{-1} \\ \cdot A_1' \Omega^{*-1} X(X' \Omega^{*-1} X)^{-1} X' \Omega^{*-1} y_1$$

where $\Omega^* = \text{diag}[\sum_{j=1}^k x_{1j1}^2, \sum_{j=1}^k x_{1j2}^2, \dots, \sum_{j=1}^k x_{1jN}^2]$. This estimator can be

computed and it will be relatively more efficient than (2.10) in the cases in which the true values of θ_j 's lie in a close neighborhood of 1.

⁷If we had applied Aitken's procedure to equation (2.13) with $\Omega^{-1} = I$ we would have obtained the following as an estimator of δ_1 :

$$\hat{\delta}_1 = [A_1' S A_1]^{-1} A_1' S y_1$$

where $S = X(X' \Omega X)^{-1} X'$. The estimator $\hat{\delta}_1$ can be shown to be consistent but it may be asymptotically less efficient than the estimator (2.14).

Next let

$$(2.15) \quad y_1^* = y_1 - y_1 \beta_1 = X_1 \bar{y}_1 + w_1.$$

Assume for a moment that β_1 is known. Then the Hildreth and Houck [1968] estimator of θ is ⁸

$$(2.16) \quad \hat{\theta} = [(\dot{M}_1 \dot{X}_1)' (\dot{M}_1 \dot{X}_1)]^{-1} (\dot{M}_1 \dot{X}_1)' \hat{w}_1$$

when the rank of $\dot{M}_1 \dot{X}_1$ is k . The $\hat{w}_1$ in (2.16) is a $N \times 1$ vector obtained by squaring each element of $\hat{w}_1 = M_1 y_1^*$.

Now relaxing the assumption that β_1 is known, we may replace β_1 in y_1^* by its consistent and approximately unbiased TSLS estimator ⁹

⁸It is well known that $\hat{\theta}$ defined in (2.16) is an unbiased, consistent but inefficient estimator of θ . Further, if $\xi = \hat{w}_1 - \dot{Z}_1 \theta$, where $\dot{Z}_1 = \dot{M}_1 \dot{X}_1$, is normally distributed, then it can be shown that $E(\xi \xi') = 2 \dot{\psi}_1 = 2 M_1 \Omega M_1^* M_1 \Omega M_1$.

Thus, we can obtain the Stepwise Aitken Least Squares Estimator of θ as (see Raj [1975]) $\tilde{\theta} = [\dot{Z}_1' \hat{\psi}_1^{-1} \dot{Z}_1]^{-1} \dot{Z}_1' \hat{\psi}_1^{-1} \hat{w}_1$. It is not clear whether we should use $\tilde{\theta}$ or $\hat{\theta}$ in place of the true value of θ used in (2.14) to obtain a good approximation to (2.14).

⁹We have corrected the conventional TSLS estimator for its bias, utilizing the approximation obtained by Nagar [1959] and subsequently by Kadane [1971]; one can use the exact expression derived by Sawa [1969] if $m = 1$. It may, however, be mentioned that asymptotically Kadane's and Sawa's results are the same.

$$(2.17) \quad \hat{\beta}_1^* = \hat{\beta}_1 - (K-k-m-1)\hat{Q}_{11} \hat{q}_1$$

where $\hat{\beta}_1$ consists of the first m elements of $\hat{\delta}_1$ in (2.10),

$$(2.18) \quad \hat{Q} = \begin{bmatrix} Y_1' & P & Y_1 & Y_1' X_1 \\ X_1' & Y_1 & X_1' X_1 \end{bmatrix}^{-1} = \begin{bmatrix} \hat{Q}_{11} & \hat{Q}_{12} \\ \hat{Q}_{21} & \hat{Q}_{22} \end{bmatrix};$$

$$\hat{Q}_{11} = (Y_1' [P - X_1 (X_1' X_1)^{-1} X_1'] Y_1)^{-1}, \text{ and}$$

$$(2.19) \quad \hat{q}_1 = \frac{1}{N} [Y_1' (I - P) (y_1 - A_1 \hat{\delta}_1)].$$

The estimator $\hat{\beta}_1^*$ is in fact constructed by subtracting a consistent estimator of the asymptotic bias in the TSLS estimator of β_1 from $\hat{\beta}_1$.

It should be noted that the estimator (2.17) provides, to the order N^{-1} , an unbiased estimator of β_1 if $E \hat{\beta}_1 < \infty$. This condition is satisfied if the degree of overidentification of equation (2.9) is at least one.¹⁰

¹⁰Alternatively, one could use the Mehta and Swamy [1975] estimators

$$(Y_1' [P - X_1 (X_1' X_1 + \mu I)^{-1} X_1'] Y_1 + \mu I)^{-1}$$

$$\cdot Y_1' [P - X_1 (X_1' X_1 + \mu I)^{-1} X_1'] y_1,$$

$$(Y_1' [P - X_1 (X_1' X_1 + \mu I)^{-1} X_1'] Y_1 + \mu I)^{-1} \text{ and } (A_1' P A_1 + \mu I)^{-1} A_1' P y_1 \text{ with any } \mu > 0$$

in place of $\hat{\beta}_1$, $\hat{Q}_{11}$ and $\hat{\delta}_1$ respectively used in (2.17). The Mehta-Swamy

estimators possess finite moments of all integer order regardless of the degree of over-identification of the equation being estimated. Also,

asymptotically the estimator $(A_1' P A_1 + \mu I)^{-1} A_1' P y_1$ will have the same

properties as the TSLS estimator (2.10).

Put

$$(2.20) \quad \hat{y}_1^* = y_1 - Y_1 \hat{\beta}_1^*.$$

Using $\hat{y}_1^*$ in place of y_1^* used in $\hat{w}_1 = M_1 y_1^*$, we obtain $\hat{w}_1^* = M_1 \hat{y}_1^*$. An operational version of the estimator (2.16) is

$$(2.21) \quad \hat{\theta}^* = [(\dot{M}_1 \dot{X}_1)' (\dot{M}_1 \dot{X}_1)]^{-1} (\dot{M}_1 \dot{X}_1)' \hat{w}_1^*.$$

Now an estimator $\hat{\Omega}^*$ of Ω can be constructed by replacing the unknown θ in (2.7) by its estimates $\hat{\theta}^*$ in (2.21). It can be shown that the estimator $\hat{\Omega}^*$ is a consistent estimator of Ω .

For this first we observe that

$$\begin{aligned} (2.22) \quad \hat{w}_1^* &= M_1 \hat{y}_1^* \\ &= M_1 [y_1 - Y_1 \hat{\beta}_1^*] \\ &= M_1 [y_1 - Y_1 \hat{\beta}_1 + (K-k-m-1) Y_1 \hat{Q}_{11} \hat{q}_1] \\ &= M_1 [A_1 \delta_1 + w_1 - A_1 \hat{\delta}_1 + (K-k-m-1) Y_1 \hat{Q}_{11} \hat{q}_1] \\ &= M_1 [w_1 - A_1 (\hat{\delta}_1 - \delta_1) + (K-k-m-1) Y_1 \hat{Q}_{11} \hat{q}_1] \\ &= M_1 [w_1 - A_1 (\hat{\delta}_1 - \delta_1) + (K-k-m-1) A_1 \hat{Q} \begin{pmatrix} \hat{q}_1 \\ 0 \end{pmatrix}] \end{aligned}$$

Now using (2.22) to form $\hat{w}_1^*$ and substituting $\hat{w}_1^*$ in (2.21), we can

verify that

$$(2.23) \quad \hat{\theta}^* - \theta = \eta$$

where the elements of η are of order $O_p(N^{-g})$; $g \cong \frac{1}{2}$.

Finally using (2.23), we find that

$$(2.24) \quad \hat{\Omega}^* - \Omega = O_p(N^{-g}), \quad g \cong \frac{1}{2},$$

whence $\hat{\Omega}^*$ is a consistent estimator of Ω .

Now, the operational estimator of δ_1 may be written as

$$(2.25) \quad \tilde{\delta}_1^* = (A_1' \hat{R}^* A_1)^{-1} A_1' \hat{R}^* y_1,$$

where

$$(2.26) \quad \hat{R}^* = \hat{\Omega}^{*-1} X(X' \hat{\Omega}^{*-1} X)^{-1} X' \hat{\Omega}^{*-1}, \text{ provided } \hat{\Omega}^{*-1} \text{ exists.}$$

The estimator (2.25) will be termed the GTSLSR estimator (the first and the last letter in this mnemonic stand for generalized and random respectively). But, any element of $\hat{\theta}^*$ can be negative with positive probability. We cannot adopt a negative estimate for a positive parameter. Hildreth and Houck [1968] have suggested that one may replace negative elements of $\hat{\theta}^*$ by zeros.¹¹ Following this suggestion, we obtain another estimator of Ω which, when used in place of Ω used in (2.14), gives another estimator of δ_1 . This new estimator may be termed the GTSLSRM estimator (the last letter in this mnemonic stands for modified). Notice that the GTSLSRM estimator will be the same as the GTSLSR estimator if none of the elements of $\hat{\theta}^*$ is negative.

¹¹Alternatively, one may apply a quadratic programming method to obtain non-negative estimates for the θ 's.

3. ASYMPTOTIC PROPERTIES

Following Fuller and Battese [1973, p. 629] and Swamy and Mehta [1977]

we assume that

Assumption 4:

(i) The elements of $R = \Omega^{-1} X (X' \Omega^{-1} X)^{-1} X' \Omega^{-1}$ are functions of the parameter vector θ such that the elements of the matrices $(\frac{\partial}{\partial \theta_i} R)$, $i = 1, 2, \dots, k$, where θ_i is the i -th element of θ , are continuous functions of θ in an open sphere S of θ^0 , the true value of θ ;

(ii) the matrices A_1 and R are such that

$$(3.1) \quad \text{Plim}_{N \rightarrow \infty} \frac{1}{N} A_1' R A_1 = M(\theta)$$

is a finite matrix, $M^{-1}(\theta)$ exists for all θ in S , and

$$(3.2) \quad \text{Plim}_{N \rightarrow \infty} \frac{1}{N} A_1' \left(\frac{\partial}{\partial \theta_i} R \right) A_1 = H_i(\theta)$$

is a finite matrix whose elements are continuous functions of θ ,

$i = 1, 2, \dots, k$; and

(iii) an estimator $\hat{R}^* \equiv R(\hat{\theta}^*)$ for $R \equiv R(\theta^0)$ is available and $\hat{\theta}^*$ satisfies the condition

$$(3.3) \quad \hat{\theta}^* = \theta^0 + O_p(N^{-g}), \quad g > 0$$

Assumption 5:

(i) Let f_r' be the r -th row of $\Omega^{-\frac{1}{2}}$. Then the elements $f_r' w_1 = f_r' D_{x_1} \epsilon_1$, $r = 1, 2, \dots, N$, are independent and $f_r' w_1$ has distribution function $F_r(f_r' w_1)$, $r = 1, 2, \dots, N$, such that

$$(3.4) \quad \sup_{r=1, 2, \dots, N} \int_{|f_r' w_1| > c} (f_r' w_1)^2 d F_r(f_r' w_1) \rightarrow 0 \quad \text{as } c \rightarrow \infty$$

(ii) $\text{Plim}_{N \rightarrow \infty} \max_{1 \leq r \leq N} \left(\frac{x_{1jr}^{*2}}{N} \right) = 0$, $j = 1, 2, \dots, K$, where x_{1jr}^* is the

(r, j) -th element of $\Omega^{-\frac{1}{2}} X$;

(iii) $\text{Plim}_{N \rightarrow \infty} \left(\frac{X' \Omega^{-1} X}{N} \right)$ is finite and nonsingular, and $\text{Plim}_{N \rightarrow \infty} \left(\frac{A_1' \Omega^{-1} X}{N} \right)$

is finite.

Lemma: If Assumptions 1, 2, 3, and 5 are satisfied, then the limiting distribution of $\sqrt{N}(\tilde{\delta}_1 - \delta_1)$ is normal with mean vector 0 and variance-covariance matrix $\text{Plim}_{N \rightarrow \infty} N(A_1' R A_1)^{-1}$.

Proof: Write

$$(3.5) \quad \sqrt{N} (\tilde{\delta}_1 - \delta_1) = \left[\frac{A_1' \Omega^{-1} X}{N} \left(\frac{X' \Omega^{-1} X}{N} \right)^{-1} \frac{X' \Omega^{-1} A_1}{N} \right]^{-1} \frac{A_1' \Omega^{-1} X}{N} \left(\frac{X' \Omega^{-1} X}{N} \right)^{-1} \frac{X' \Omega^{-1} w_1}{\sqrt{N}}$$

By Assumption 5(iii)

$$(3.6) \quad \text{Plim}_{N \rightarrow \infty} \left[\frac{A_1' \Omega^{-1} X}{N} \left(\frac{X' \Omega^{-1} X}{N} \right)^{-1} \frac{X' \Omega^{-1} A_1}{N} \right]^{-1} \frac{A_1' \Omega^{-1} X}{N} \left(\frac{X' \Omega^{-1} X}{N} \right)^{-1} = \bar{G},$$

a finite matrix.

It follows from Anderson's [1971, pp. 23-25, 585] Theorem 2.6.1 that under

Assumptions 5(i) and (ii), $\frac{X' \Omega^{-1} w_1}{\sqrt{N}}$ converges in distribution to normal

with mean vector 0 and variance-covariance matrix $\text{Plim}_{N \rightarrow \infty} \frac{X' \Omega^{-1} X}{N}$.

Now applying the limit theorem (x)b in Rao [1973, 2c.4, p. 122], we have the result of the Lemma.

Theorem: Suppose that Assumptions 1, 2, 3, and 4 are true. Then the limiting distribution of $\sqrt{N} (\tilde{\delta}_1^* - \delta_1)$ is the same as that of $\sqrt{N} (\tilde{\delta}_1 - \delta_1)$.

Proof: Write

$$(3.7) \quad \tilde{\delta}_1^* - \delta_1 = (A_1' R(\hat{\theta}^*) A_1)^{-1} A_1' R(\hat{\theta}^*) w_1$$

where $R(\hat{\theta}^*) = \hat{\Omega}^{*-1} X (X' \hat{\Omega}^{*-1} X)^{-1} X' \hat{\Omega}^{*-1}$.

By a Taylor's expansion with remainder about the true parameter θ^0 , we obtain,

$$\begin{aligned} (3.8) \quad & (A_1' R(\hat{\theta}^*) A_1)^{-1} A_1' R(\hat{\theta}^*) w_1 = \left(\frac{1}{N} A_1' R(\theta^0) A_1 \right)^{-1} \left(\frac{1}{N} A_1' R(\theta^0) w_1 \right) \\ & + \sum_{i=1}^k \left\{ \left(\frac{1}{N} A_1' R(\theta^1) A_1 \right)^{-1} \left(\frac{1}{N} A_1' \left(\frac{\partial}{\partial \theta_i} R(\theta) \right)_{\theta=\theta^1} w_1 \right) \right. \\ & - \left(\frac{1}{N} A_1' R(\theta^1) A_1 \right)^{-1} \left(\frac{1}{N} A_1' \left(\frac{\partial}{\partial \theta_i} R(\theta) \right)_{\theta=\theta^1} A_1 \right) \\ & \left. \cdot \left(\frac{1}{N} A_1' R(\theta^1) A_1 \right)^{-1} \left(\frac{1}{N} A_1' R(\theta^1) w_1 \right) \right\} (\hat{\theta}_i^* - \theta_i^0) \end{aligned}$$

where θ^1 is between θ^0 and θ . By Assumption 4 and limit theorem (x)a in Rao [1973, 2c.4, p. 122], it follows that

$$(3.9) \quad (\tilde{\delta}_1^* - \delta_1) = (\tilde{\delta}_1 - \delta_1) + o_p[N^{-(\frac{1}{2} + g)}]$$

or

$$(3.10) \quad \sqrt{N} (\tilde{\delta}_1^* - \delta_1) = \sqrt{N} (\tilde{\delta}_1 - \delta_1) + o_p(N^{-g})$$

which, in view of the theorem (x)d in Rao [1973, 2c.4, p. 122], gives the result of the Theorem.

4. AN EXAMPLE: KLEIN MODEL I

To illustrate the GTSLSR and GTSLSRM estimators developed in the previous section, we utilize Klein's Model I (see Klein [1950]). This model is dynamic in the sense that it contains some lagged values. It consists of three behavioral equations and three identities expressed in terms of six jointly dependent variables and eight predetermined variables. The random coefficients version of the model is briefly described below:

$$(4.1) \quad C_t = \beta_1 P_t + \beta_2 (W_t + W'_t) + \gamma_{1(t)} + \gamma_{2(t)} P_{t-1} \quad (\text{consumption equation})$$

$$(4.2) \quad I_t = \beta'_1 P_t + \gamma'_{1(t)} + \gamma'_{2(t)} P_{t-1} + \gamma'_{3(t)} K_{t-1} \quad (\text{investment equation})$$

$$(4.3) \quad W_t = \beta''_1 X_t + \gamma''_{1(t)} + \gamma''_{2(t)} X_{t-1} + \gamma''_{3(t)} (t - 1931) \quad (\text{labour equation})$$

$$(4.4) \quad X_t = C_t + I_t + G_t \quad (\text{National Accounts: Definition 1})$$

$$(4.5) \quad P_t = X_t - W_t - T_t \quad (\text{Profit-and-Loss Account: Definition 2})$$

$$(4.6) \quad K_t = K_{t-1} + I_t \quad (\Delta \text{ in capital stock} = \text{Investment: Definition 3})$$

where

C_t = aggregate consumption in year t

P_t = total profits in year t

W_t = total wages paid by private industry in year t

I_t = net investment in year t

K_t = capital stock in year t

X_t = total production of private industry in year t

T_t = taxes in year t

G_t = government expenditures in year t .

W_t = government wage bill in year t

P_{t-1} = total profits lagged one period

K_{t-1} = capital stock lagged one period

X_{t-1} = total production of private industry lagged one period

t = time trend.

Our objective is to estimate the β 's and the means of γ 's on the basis of the U.S. data for the period 1921-41. It may be noted that in the above model a time trend and lagged endogenous variables appear. Thus, the exogenous variables contain time trends and our assumptions 4(ii) and (iii), and 5(ii) and (iii) are not satisfied. However, these assumptions can be easily modified and the Lemma and the Theorem in Section 3 can be reformulated as in Anderson [1971, pp. 23-25, 572 and 585] so that they are true even when the time trends are present among the exogenous variables. Further, the results of the Lemma and the Theorem will continue to hold when lagged endogenous variables are included in the model provided the lagged endogenous variables are independent of the disturbances and random coefficients,

The fixed coefficients version of the model was estimated by Rothenberg and Leenders [1964] and others. We present the GTSLSR and GTSLSRM estimates for the above random coefficients version of Klein's Model I in Table 1, along with the usual TSLS estimates and also with standard errors of parameter estimates.

A comparison of the standard errors in columns 3 and 4 of Table 1 reveals that the conventional estimator of the variance of TSLS estimator may be an underestimator if the model (4.1)-(4.6) with random coefficients is true. Similarly, comparing the standard errors of TSLS estimates in column 4 with those of GTSLSR estimates in column 5 shows that the GTSLSR estimator can be more efficient than the TSLS estimator. The GTSLSR and GTSLSRM estimates

and their standard errors in columns 5 and 6 are identical for the consumption and labor equations because, in these cases, all estimates of the variances of random coefficients are positive. However, the GTSLSR and GTSLSRM estimates differ for the investment equation because the estimates of the variances of some random coefficients are negative. For this equation the GTSLSRM estimator appears to be less efficient than the GTSLSR estimator. This result is somewhat surprising because in the Monte Carlo results published by Froehlich [1973] the truncated estimators of coefficient variances are found to be preferable in terms of mean square error to untruncated estimators.

In order to provide a further quantitative comparison of the efficiency of alternative estimators considered we have estimated the asymptotic variance-covariance matrix of each of these estimators. The determinant and sum of all the elements of each of the estimated variance-covariance matrices for each equation are given in Table 2. Comparing these two measures of efficiency, given in columns 4 and 5 of Table 2, suggests that the TSLS estimator can be less efficient than the GTSLSR estimator. A similar comparison of efficiencies between the GTSLSR and GTSLSRM estimators in columns 5 and 6 suggests that the GTSLSR estimator can be more efficient than the GTSLSRM estimator.

TABLE 1

ALTERNATIVE ESTIMATES OF THE FIXED COEFFICIENTS AND THE MEANS OF RANDOM COEFFICIENTS IN KLEIN'S MODEL 1: 1921-41

Equations	Variables	Coefficient or mean coefficient estimates (Standard Errors)			
		TSLS ^a	TSLS ^b	GTSLSR	GTSLSRM
Consumption	P	0.0173 (0.1180)	0.0173 (0.1305)	0.0206 (0.1300)	0.0206 (0.1300)
	(W + W')	0.8102 (0.0402)	0.8102 (0.0447)	0.8125 (0.0447)	0.8125 (0.0447)
	Intercept	16.5548 (1.3208)	16.5548 (1.4563)	16.4780 (1.4561)	16.4780 (1.4561)
	P ₋₁	0.2162 (0.1073)	0.2162 (0.1192)	0.2115 (0.1184)	0.2115 (0.1184)
Investment	P	0.1502 (0.1732)	0.1502 (0.2707)	0.1251 (0.2233)	0.1403 (0.2764)
	Intercept	20.2782 (7.5427)	20.2782 (11.3435)	20.1388 (8.7280)	19.9946 (11.2188)
	P ₋₁	0.6159 (0.1628)	0.6159 (0.2555)	0.6422 (0.2104)	0.6246 (0.2590)
	K ₋₁	-0.1578 (0.0361)	-0.1578 (0.0546)	-0.1571 (0.0421)	-0.1562 (0.0539)
Labour	X	0.4389 (0.0356)	0.4389 (0.0371)	0.4288 (0.0317)	0.4288 (0.0317)
	Intercept	1.5003 (1.1478)	1.5003 (1.1527)	1.8343 (1.0413)	1.8343 (1.0413)
	X ₋₁	0.1467 (0.0388)	0.1467 (0.0380)	1.1528 (0.0334)	1.1528 (0.0334)
	(t-1931)	0.1304 (0.0291)	0.1304 (0.0323)	0.1324 (0.0314)	0.1324 (0.0314)

^a The standard errors are computed from the matrix $\hat{\sigma}_{w_1}^2 [A_1' P A_1]^{-1}$
 where $\hat{\sigma}_{w_1}^2 = \frac{1}{N} (y_1 - A_1 \hat{\delta}_1)' (y_1 - A_1 \hat{\delta}_1)$

^b The standard errors are obtained from the matrix $[A_1' P A_1]^{-1}$.
 $A_1' P \hat{\Omega}_M^* P A_1 [A_1' P A_1]^{-1}$ where $\hat{\Omega}_M^*$ is a modified estimator of Ω
 used in the GTSLSRM.

TABLE 2

VARIABILITY MEASURES FOR KLEIN'S MODEL I: 1921-41

Equations	Scalar measures for the Variance- Covariance Matrix	Methods of Estimation			
		TSLSa	TSLS ^b	GTSLSR	GTSLSRM
Consumption	Sum of All the Elements Determinant	1.6426 3.2389×10^{-9}	1.9990 7.2702×10^{-9}	2.0003 7.2090×10^{-9}	2.0003 7.2090×10^{-9}
Investment	Sum of All the Elements Determinant	56.0668 6.0667×10^{-9}	126.8102 2.1554×10^{-7}	75.0810 3.6886×10^{-8}	124.0147 2.2126×10^{-7}
Labour	Sum of All the Elements Determinant	1.3051 6.6761×10^{-12}	1.2932 1.1142×10^{-11}	1.0552 6.1459×10^{-12}	1.0552 6.1459×10^{-12}

^a See footnote a in Table 1.^b See footnote b in Table 1.

5. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

In this paper we have studied the problem of estimation of simultaneous equation models with fixed and random coefficients and developed some useful estimators. More specifically, we have developed an efficient single equation method of estimating the structural coefficients of endogenous variables and the means of structural coefficients of exogenous variables and established its asymptotic properties. Further, Klein's Model I has been reformulated, making the coefficients of predetermined variables random. Our estimates for this reformulated Model are found to be better than the conventional TSLS estimates.

Several suggestions for further work could be made. First, there is need to study further the problems of identification and estimation of simultaneous equations models when the coefficients of endogenous variables are also random. Second, it will be interesting to examine the familiar problems of undersized samples in the context of simultaneous equations models with random coefficients.

Wilfrid Laurier University, Waterloo, Canada
Lucknow University, Lucknow, India
University of Western Ontario, London, Canada.

References

- Anderson, T. W., The Statistical Analysis of Time Series (New York: Wiley, 1971).
- Fuller, W. A., and G. E. Battese, "Transformations for Estimation of Linear Models with Nested-Error Structure," Journal of the American Statistical Association, 68 (September 1973), 626-632.
- Froehlich, B. R., "Small Sample Properties of Various Consistent Variance Estimators Proposed for a Random Coefficient Regression Model," Journal of the American Statistical Association, 68 (June 1973), 329-35.
- Hildreth, C., and J. Houck, "Some Estimates for a Linear Model with Random Coefficients," Journal of the American Statistical Association, 63 (June 1968), 584-595.
- Kadane, J. B., "Comparison of K-Class Estimators when the Disturbances are Small," Econometrica, 39 (July 1971), 723-737.
- Kelejian, H. H., "Random Parameters in a Simultaneous Equation Framework: Identification and Estimation," Econometrica, 42 (May 1974) 517-527.
- Klein, L., Economic Fluctuations in the United States, 1921-1941 (New York: Wiley, 1950).
- Mehta, J. S., and P. A. V. B. Swamy, "The Existence of Moments of Some Simple Bayes Estimator of Coefficients in Simultaneous Equation Model," Federal Reserve Board, Washington, No. 71 (December 1975).
- Nagar, A. L. "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," Econometrica, 27 (October 1959), 575-595.
- Nerlove, M., Estimation and Identification of Cobb-Douglas Production Functions (Chicago: Rand McNally, 1965, 34-35, 61-82).

- Raj, Baldev, "Linear Regression with Random Coefficients: The Finite Sample and Convergence Properties," Journal of the American Statistical Association, 70 (March 1975), 127-137.
- Rao, C. R., Linear Statistical Inference and Its Applications, 2nd edition, (New York: Wiley, 1973).
- Rothenberg, T., and C. Leenders, "Efficient Estimation of Simultaneous Equation Systems," Econometrica, 32 (January-April 1964), 57-76.
- Rubin, H., "Note on Random Coefficients," in T. C. Koopmans, ed., Statistical Inference in Dynamic Economic Models (Cowles Commission Monograph No. 10 (1950), 419-421).
- Sargent, T. J., "The Observational Equivalence of Natural and Unnatural Rate Theories of Macro-economics," Journal of Political Economy, 84 (June 1976), 631-640.
- Sawa, T., "The Exact Sampling Distribution of Ordinary Least Squares and Two Stage Least Squares Estimators," Journal of American Statistical Association, 64 (September 1969), 923-937.
- Swamy, P. A. V. B., "Efficient Inference in a Random Coefficient Regression Model," Econometrica, 38 (March 1970), 311-323.
- Swamy, P. A. V. B., and J. Holmes, "The Use of Undersized Samples in the Estimation of Simultaneous Equation Systems," Econometrica, 39 (May 1971), 455-459.
- Swamy, P. A. V. B., and J. S. Mehta, "Bayesian and Non-Bayesian Analysis of Switching Regressions and of Random Coefficient Regression Models," Journal of the American Statistical Association, 70 (September 1975), 593-602.
- Swamy, P. A. V. B., and J. S. Mehta, "Estimation of Linear Models with Time and Cross-Varying Coefficients," Journal of the American Statistical Association (forthcoming).

Theil, H., and L. B. Mennes, "Conception Stochastique de Coefficients Multiplicateurs dans l'Adjustement Linéaire des Séries Temporelles," Publications de l'Institut Statistique de l'Université de Paris, 8 (June 1959), 211-227.

Theil, H. (assisted by J. S. Cramer, H. Moerman, A. Russchen), Economic Forecasts and Policy (Amsterdam: North-Holland, 1961).

Zellner, A., "On the Aggregation Problem: A New Approach to a Troublesome Problem," Center for Mathematical Studies in Business and Economics, University of Chicago, No. 6628 (July 1966). Published in K. Fox et al. (eds.), Econometric Models, Estimation and Risk Programming: Essays in Honour of Gerhard Tintner, (New York: Springer-Verlag, 1969).