Institute of Economic Development and Research
SCHOOL OF ECONOMICS
University of the Philippines

Discussion Paper No. 70-7 March 23, 1970

ON THE EXISTENCE OF MOMENTS OF THE ORDINARY LEAST SQUARES AND TWO-STAGE LEAST SQUARES ESTIMATORS

By

Roberto S. Mariano , 1944

NOTE: IEDR Discussion Papers are preliminary versions circulated privately to elicit critical comment. References in publications to Discussion Papers should be cleared with the author.

SUMMARY

This paper deals with two single-equation estimators in a set of simultaneous linear stochastic equations--namely, two-stage least squares (2SLS) and ordinary least squares (OLS).

Under the assumption that all predetermined variables in the model are exogenous, it is shown that for the general case with an arbitrary number of included endogenous variables, even moments of the 2SLS estimator are finite if the order is less than K_2-G_1+1. N is the sample size, G_1+1 the number of included endogenous variables, K_1 and K_2 respectively the number of included and excluded exogenous variables in the equation to be estimated.

The starting point for the proof of these results is the characterization of the OLS and 2SLS estimators as functions of non-central Wishart matrices. It is shown that the 2SLS estimator is given by $W_{22}^{-1}W_{21}$ where W is a non-central Wishart matrix partitioned as $W = \begin{pmatrix} w_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$, w_{11} being 1x1. The OLS estimator is similarly equal to $A_{22}^{-1}A_{21}$ where A is another non-central Wishart matrix which has the same covariance and means sigma matrices as W but different degrees of freedom.

The expressions for these two estimators are reduced to canonical form to set the stage for the proof of the main result. As a by-product of this reduction, the exact functions of the original parameters of the model which affect the probability distributions of the OLS and 2SLS estimators are determined.

ON THE EXISTENCE OF MOMENTS OF THE ORDINARY LEAST SQUARES AND TWO-STAGE LEAST SQUARES ESTIMATORS[1]

ROBERTO S. MARIANO[2]
UNIVERSITY OF THE PHILIPPINES

1. INTRODUCTION.

Basmann (4,5) has conjectured that the moments of the two-stage least squares (2SLS) estimator of an equation in a simultaneous system of linear stochastic equations exist if the order is less than K_2-G_1+1 where G_1+1 and K_2 are respectively the numbers of endogeneous variables included and predetermined variables excluded from the equation being estimated. Under the assumption that all predetermined variables are exogenous, Richardson (7) and Sawa (8) have confirmed this conjecture for the case of two included endogeneous variables with both the number of excluded exogenous variables and the number of equations in the model being arbitrary.

---

[1] Based on a dissertation submitted to the Department of Statistics, Stanford University for a Ph.D. degree in Statistics. The author is sincerely grateful for the suggestions and guidance given by his adviser, Professor T.W. Anderson, in the preparation of his dissertation.

[2] The author acknowledges the support of the Rockefeller Foundation and the National Science Foundation (GS-2635) for the research presented here.

Under the assumption that all predetermined variables in the model are exogenous, we prove in this paper that for the general case with an arbitrary number of included endogenous variables, even moments of the 2SLS estimator are finite if the order is less than K_2-G_1+1 and infinite otherwise, and that even moments of the ordinary least squares (OLS) estimator exist if the order is less than $N-K_1-G_1+1$, where N is the number of observations and K_1 the number of included exogenous variables.

The starting point for the proof of these results is the characterization of the OLS and 2SLS estimators as functions of non-central Wishart matrices. This is done in the next section. Section 3 deals with the reduction of the estimators to canonical form and in section 4, the proof of the main results is given.

In closing this section, we would like to indicate our notation for the Wishart and multivariate normal distributions.

We write $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu},\Sigma)$ to indicate that $\underset{\sim}{x}$ is a pxl normal random vector with mean $\underset{\sim}{\mu}$ and covariance matrix Σ.

If $W = Z'Z$, where Z is a nxp random matrix whose rows are mutually independent normal random vectors with a common covariance matrix Σ, then W has a Wishart distribution of order p, with n degrees of freedom, covariance matrix Σ and means sigma matrix $(EZ)'(EZ)$. We denote this by $W \sim W_p\left[n,\Sigma;(EZ)'(EZ)\right]$. The distribution of W is said to be central if $(EZ)'(EZ) = \underset{\sim}{0}$.

2. THE ESTIMATORS AS FUNCTIONS OF NON-CENTRAL WISHART MATRICES.

In a simultaneous system of G linear stochastic equation relating G endogenous and K predetermined variables, the single equation to be estimated may be written as

$$\underset{\sim}{y} = Y_1\underset{\sim}{\beta} + Z_1\underset{\sim}{\gamma} + \underset{\sim}{u} \qquad (2.1)$$

where $(\underset{\sim}{y}\ Y_1)$ is the $Nx(G_1+1)$ matrix of included endogenous variables, Z_1 the NxK_1 matrix of included predetermined variables, $\underset{\sim}{u}$ the $Nx1$ vector of disturbance terms and $\underset{\sim}{\beta}$ and $\underset{\sim}{\gamma}$ are vectors of unknown coefficients.

The reduced form equations for the G_1+1 endogenous variables included in (2.1) are

$$Y = Z\Pi' + V \qquad (2.2)$$

$$= Z_1\Pi_1' + Z_2\Pi_2' + V \qquad (2.3)$$

where $Y = (\underset{\sim}{y}\ Y)$, $Z = (Z_1\ Z_2)$, Z_2 is the NxK_2 matrix of excluded predetermined variables, V is the $Nx(G_1+1)$ matrix of reduced form disturbance terms and Π is the $(G_1+1)xK$ matrix of reduced form coefficients partitioned as $(\Pi_1\ \Pi_2)$ where Π_1 is $(G_1+1)xK_1$ and Π_2 is $(G_1+1)xK_2$.

In this paper, we make the following assumptions about the model:

(1) All predetermined variables are exogenous.

(2) The equation to be estimated is identified by zero-restrictions on the structural coefficients in the model.

(3) The sample size is greater than or equal to the total number of variables in the system. $(N \geq G+K)$.

(4) Z is a matrix of constants and is of full rank.

(5) The rows of V are mutually independent and identically distributed as normal random vectors with zero mean vector and positive definite covariance matrix Σ.

It is well known that assumption (2) is equivalent to the assumption that the rank of Π_2 is G_1 , which in turn implies that $K_2 \geq G_1$. Also, assumption (5) implies that in (2.1) , $E\underset{\sim}{u} = \underset{\sim}{o}$.

Under our set-up, it can be shown that the expression for Theil's k-class estimators of β (for example, as given in Johnston (6), p. 260) simplifies to

$$(Y_1{}'P_\kappa Y_1)^{-1} Y_1{}'P_\kappa y \tag{2.4}$$

where

$$P_k = \left[I - Z_1(Z_1{}'Z_1)^{-1}Z_1{}'\right] - k\left[I - Z(Z'Z)^{-1}Z'\right]. \tag{2.5}$$

It is well known that the OLS and 2SLS estimates of β correspond to $k=0$ and $k=1$, respectively, and hence it follows that the OLS and 2SLS estimators of β are given by $A_{22}^{-1}A_{21}$ and $W_{22}^{-1}W_{21}$ respectively, where

$$A = Y'P_0Y = \begin{pmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{2.6}$$

and

$$W = Y'P_1Y = \begin{pmatrix} w_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}. \tag{2.7}$$

In (2.6) and (2.7), A_{22} and W_{22} are both $G_1 x G_1$ submatrices.[3]

---

[3] In terms of A and W, the limited - information maximum likelihood estimator of β is the right-side characteristic vector of $(A-W)^{-1}W$ (normalized such that its first component is equal to unity) corresponding to the smallest root of $|W - \lambda(A-W)| = 0$.

Since Z is of full rank (by assumption (4)), there exists a $K \times K$ non-singular upper triangular matrix ϕ such that $\phi' Z' Z \phi = I$. Partition ϕ as follows:

$$\phi = \begin{array}{cc} \begin{pmatrix} \phi_{11} & \phi_{12} \\ \underset{\sim}{0} & \phi_{22} \end{pmatrix} & \begin{array}{c} k_1 \\ k_2 \end{array} \\ \begin{array}{cc} k_1 & k_2 \end{array} & \end{array} \tag{2.8}$$

and let

$$X = Z\phi \tag{2.9}$$

$$= (Z_1\phi_{11} \quad Z_1\phi_{12}+Z_2\phi_{22}) \tag{2.10}$$

$$= (X_1 \quad X_2). \tag{2.11}$$

Note that $X'X = I$ and ϕ_{11} and ϕ_{22} are both upper triangular non-singular matrices.

It can be verified that

$$XX' = Z(Z'Z)^{-1}Z' = X_1X_1' + X_2X_2' \tag{2.12}$$

and

$$X_1X_1' = Z_1(Z_1'Z_1)^{-1}Z_1', \tag{2.13}$$

so that the matrices P_0 and P_1 as given by (2.5) may also be expressed as

$$\begin{aligned} P_0 &= I - X_1X_1' \\ P_1 &= X_2X_2'. \end{aligned} \tag{2.14}$$

Furthermore, by (2.3), $EY = Z_1\Pi_1' + Z_2\Pi_2'$ and thus, it follows from (2.10) and (2.11) that

$$\begin{aligned}(EY)'P_0(EY) &= (EY)'P_1(EY) \\ &= \Pi_2 Z_2' P_1 Z_2 \Pi_2'. \end{aligned} \tag{2.15}$$

Now, by (2.2) and assumption (5), the rows of Y are mutually independent normal random vectors with common covariance matrix Σ and by (2.14), P_0 and P_1 are symmetric and idempotent with ranks $N-K_1$ and K_2 respectively. Hence, the following proposition holds:

Proposition 2.1. The OLS and 2SLS estimates of β are respectively given by $A_{22}^{-1}A_{21}$ and $W_{22}^{-1}W_{21}$, where

$$A \sim \mathcal{W}_{G_1+1}(N-K_1, \Sigma;\ M),$$

$$W \sim \mathcal{W}_{G_1+1}(K_2, \Sigma;\ M),$$

and

$$M = \Pi_2 Z_2' P_1 Z_2 \Pi_2'.$$

The above proposition indicates that the OLS and 2SLS estimators have the same functional form in terms of non-central Wishart matrices. Furthermore the two Wishart matrices involved have the same parameters except for the degrees of freedom. Since our approach in this paper is to look at these two estimators as the above indicated functions of

Wishart matrices in order to achieve a reduction to canonical form and analyze the existence of moments, we can then derive our results for the 2SLS estimator and from these infer the results for the OLS estimator by simply making the proper changes in the degrees of freedom of the Wishart matrix involved.

3. REDUCTION TO CANONICAL FORM.

From now on, let $\hat{\underset{\sim}{\beta}}$ be the 2SLS estimator of $\underset{\sim}{\beta}$. The usual procedure applied to reach a reduction to canonical form is by transforming the non-central Wishart matrix W so as to diagonalize M and transform Σ to the identity matrix. For example, see Anderson and Girshick (3). Unfortunately, due to the nature of $\underset{\sim}{\beta}$ as a function of W, $\underset{\sim}{\beta}$ becomes a complicated and intractable function of the resulting Wishart matrix if the above procedure were to be applied.

In this section, we shall make use of transformations of W and to R, say, such that $\hat{\underset{\sim}{\beta}}$ is a linear function of the components of $R_{22}^{-1}R_{21}$. The trade-off is that either the means sigma matrix of R is not diagonal or its covariance matrix is not the identity. To arrive at such transformations, we make use of a special property of the means sigma matrix M which we indicate in the following lemma.

Lemma 3.1. *The first row of* M *is a linear combination of its remaining rows with the components of* $\underset{\sim}{\beta}$ *as weights.*

Proof: Since $X_2'Z_1 = 0$ and $E\underset{\sim}{u} = 0$, it follows from (2.1) that

$$E(X_2'\underset{\sim}{y}) = E(X_2'Y_1)\,\underset{\sim}{\beta} \tag{3.1}$$

which implies that the first column of $E(X_2'Y)$ is a linear combination of the remaining columns, with the components of $\underset{\sim}{\beta}$ as weights. Hence, the lemma holds, since $M = E(X_2'Y)'E(X_2'Y)$. *Q.E.D.*

Note that the above lemma implies that the rank of the means sigma matrix M is less than or equal to G_1. Another lemma which we will use in the sequel is

Lemma 3.2. *If* F, H *and* L *are non-singular square matrices of the same size, each partitioned as*

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix}$$

and if all the diagonal blocks are non-singular, $L_{21} = \underset{\sim}{0}$, *and* $F = LHL'$, *then* $F_{22}^{-1}F_{21} = L_{21}'^{-1}\left[L_{12}' + H_{22}^{-1}H_{21}L_{11}\right]$.

Proof: The conclusion follows from expanding LHL'

in partitioned form and simplifying the expression for $F_{22}^{-1}F_{21}$ by using the assumption that $L_{21} = \underset{\sim}{0}$. <u>Q.E.D</u>.

We now proceed to our reduction to canonical form. Partition Σ and M as follows:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{3.2}$$

$$M = \begin{pmatrix} m_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{3.3}$$

where Σ_{22} and M_{22} are both $G_1 x G_1$, and let θ be a $G_1 x G_1$ non-singular matrix such that

$$\theta\Sigma_{22}\theta' = I \tag{3.4}$$

and

$$\theta M_{22}\theta' = D \tag{3.5}$$

where D is a $G_1 x G_1$ diagonal matrix whose main-diagonal elements are the characteristic roots of $\Sigma_{22}^{-1}M_{22}$ arranged in increasing order. Furthermore, let

$$\omega^2 = \sigma_{11} - 2\underset{\sim}{\beta}'\Sigma_{21} + \underset{\sim}{\beta}'\Sigma_{22}\underset{\sim}{\beta} \tag{3.6}$$

and

$$\Psi = \begin{array}{c} \left(\begin{array}{cc} \frac{1}{\omega} & \frac{-\underset{\sim}{\beta}'}{\omega} \\ \underset{\sim}{0} & \theta \end{array}\right) \begin{array}{l} 1 \\ G_1 \end{array} \\ \begin{array}{cc} 1 & G_1 \end{array} \end{array} . \tag{3.7}$$

By making use of Lemma 3.1 and noting that

$$\Psi = \begin{pmatrix} \frac{1}{\omega} & \underset{\sim}{0} \\ \underset{\sim}{0} & \theta \end{pmatrix} \begin{pmatrix} 1 & -\underset{\sim}{\beta}' \\ \underset{\sim}{0} & I \end{pmatrix}, \qquad (3.8)$$

it can be verified that

$$\Psi M \Psi' = \begin{pmatrix} 0 & \underset{\sim}{0} \\ \underset{\sim}{0} & D \end{pmatrix} \qquad (3.9)$$

and

$$\Psi \Sigma \Psi' = \Omega = \begin{pmatrix} 1 & \underset{\sim}{\rho}' \\ \underset{\sim}{\rho} & I \end{pmatrix}, \qquad (3.10)$$

where

$$\underset{\sim}{\rho} = \frac{\theta}{\omega} (\Sigma_{21} - \Sigma_{22} \underset{\sim}{\beta}). \qquad (3.11)$$

If we now let

$$U = \Psi W \Psi' = \begin{array}{cc|c} u_{11} & U_{12} & 1 \\ U_{21} & U_{22} & G_1 \\ 1 & G_1 & \end{array}, \qquad (3.12)$$

then

$$U \sim W_{G_1+1} \left[K_2, \Omega; \begin{pmatrix} 0 & \underset{\sim}{0} \\ \underset{\sim}{0} & D \end{pmatrix} \right]. \qquad (3.13)$$

And from (3.8), Proposition 2.1 and Lemma 3.2., it follows that

$$\hat{\underset{\sim}{\beta}} = \underset{\sim}{\beta} + \omega \theta' U_{22}^{-1} U_{21}. \qquad (3.14)$$

(3.14) indicates that in general, the 2SLS (as well as the OLS) probability distribution depends on the following parameters: $\underset{\sim}{\beta}$, ω, θ, $(\Sigma_{12}- \underset{\sim}{\beta}'\Sigma_{22})$ and the characteristic roots of $\Sigma_{22}^{-1}M_{22}$. The reduction given by (3.14) leads to a Wishart matrix U whose covariance matrix Ω is close to being the identity matrix but not quite, see (3.10). Note that $\Omega=I$ if and only if $\underset{\sim}{\beta} = \Sigma_{22}^{-1}\Sigma_{21}$.

In the next section, we shall find it easier to deal with a Wishart matrix whose covariance matrix is the identity. Another transformation on U is needed to reach such a situation. However, as an unavoidable consequence of this additional transformation, the resulting means sigma matrix becomes non-diagonal.

For the additional reduction, define the $(G_1+1)x(G_1+1)$ matrix

$$V = \begin{pmatrix} \frac{1}{\sqrt{1-\underset{\sim}{\rho}'\underset{\sim}{\rho}}} & \underset{\sim}{\rho}' \\ \frac{-\rho}{\sqrt{1-\underset{\sim}{\rho}'\underset{\sim}{\rho}}} & I \end{pmatrix} \qquad (3.16)$$

and let

$$R = V'UV = \begin{pmatrix} r_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}. \qquad (3.17)$$

From (3.10), Ω is positive definite and $|\Omega| = 1-\underset{\sim}{\rho}'\underset{\sim}{\rho}$. Thus $\underset{\sim}{\rho}'\underset{\sim}{\rho}<1$ and ∇ is well defined. It can be easily verified from (3.10) and (3.16) that

$$\nabla'\Omega\nabla = I. \tag{3.18}$$

Thus

$$R \sim W_{G_1+1}\left[K_2,\ I;\ \nabla'\begin{pmatrix} 0 & \underset{\sim}{\rho} \\ \underset{\sim}{\rho} & D \end{pmatrix} \nabla\right]. \tag{3.19}$$

Finally, it follows from (3.17) and Lemma 3.2. that

$$U_{22}^{-1}U_{21} = \underset{\sim}{\rho} + \sqrt{1-\underset{\sim}{\rho}'\underset{\sim}{\rho}}R_{22}^{-1}R_{21}, \tag{3.20}$$

which implies that

$$\underset{\sim}{\hat{\beta}} = \underset{\sim}{\beta} + \omega\Theta'\left(\underset{\sim}{\rho} + \sqrt{1-\underset{\sim}{\rho}'\underset{\sim}{\rho}}\ R_{22}^{-1}R_{21}\right). \tag{3.21}$$

4. EXISTENCE OF 2SLS MOMENTS.

Denote the first components of $\underset{\sim}{\hat{\beta}}$ and $R_{22}^{-1}R_{21}$ by $\hat{\beta}_1$ and $\hat{\beta}_1{*}$ respectively. We shall now show that for arbitrary G_1 and K_2, the even moments of $\hat{\beta}_1$ are finite if the order is less than K_2-G_1+1 and infinite otherwise. By (3.21) in the previous section, this result holds if and only if even moments of $\hat{\beta}_1{*}$ are finite for order less than K_2-G_1+1 and infinite otherwise.

From (3.19),

$$R = S'S \tag{4.1}$$

where S is a $K_2 x(G_1+1)$ matrix whose elements are mutually independent normal random variables with unit variance and means such that

$$(ES)'(ES) = V' \begin{pmatrix} 0 & \underset{\sim}{0} \\ \underset{\sim}{0} & D \end{pmatrix} V \, .$$

Partition S as follows:

$$S = \begin{array}{cc} (\underset{\sim}{s}_1 & S_2) \; K_2 \\ 1 & G_1 \end{array} \tag{4.2}$$

$$= \begin{array}{ccc} (\underset{\sim}{s}_1 & \underset{\sim}{s}_2 & S_3) \; K_2 , \\ 1 & 1 & G_1-1 \end{array} \tag{4.3}$$

so that from (4.1), we get

$$R_{22}^{-1}R_{21} = (S_2'S_2)^{-1}S_2'\underset{\sim}{s}_1 , \tag{4.4}$$

and let

$$(S_2'S_2)^{-1} = \begin{array}{cc} \begin{pmatrix} v^{22} & V^{23} \\ V^{32} & V^{33} \end{pmatrix} & \begin{array}{c} 1 \\ G_1-1 \end{array} \\ \begin{array}{cc} 1 & \quad G_1-1 \end{array} & \end{array} . \tag{4.5}$$

To prove our result on the existence of moments of $\hat{\beta}_1$, we use the following lemmas concerning v^{22}.

Lemma 4.1. Given S_3, $(v^{22})^{-1}$ is conditionally distributed as a non-central chi-square variate with K_2-G_1+1 degrees of freedom and non-centrality parameter

$$\tau^2 = E\, \underset{\sim}{s}_2' \left[I-S_3(S_3'S_3)^{-1}S_3'\right] E\underset{\sim}{s}_2 .$$

Proof. From (4.5),

$$\begin{aligned}(v^{22})^{-1} &= v_{22}-V_{32}'V_{33}^{-1}V_{32} \\ &= \underset{\sim}{s}_2' \left[I-S_3(S_3'S_3)^{-1}S_3'\right]\underset{\sim}{s}_2 .\end{aligned}$$

The random vector $\underset{\sim}{s}_2$, being the second column of S, is a K_2-dimensional normal random vector with covariance matrix I and non-zero mean vector. Furthermore, $I-S_3(S_3'S_3)^{-1}S_3'$ is a $K_2 x K_2$ idempotent symmetric matrix with rank K_2-G_1+1. Hence the lemma holds. Q.E.D.

Lemma 4.2. For m an arbitrary real number,

$$\text{(i) } E(v^{22})^m < +\infty \text{ if } 2m < K_2-G_1+1.$$

$$\text{(ii) } E(v^{22})^m = +\infty \text{ if } 2m \geq K_2-G_1+1.$$

Proof. Let $n = K_2-G_1+1$ and let τ^2 be as defined in Lemma 4.1. By Lemma 4.1, we get $(v^{22})^{-1}$ given S_3 is conditionally distributed as a central chi-square variable with $(m+2J)$ degrees of freedom, where J is a Poisson variable with parameter $\frac{\tau^2}{2}$. Thus,

$$E\,(v^{22})^m|S_3 = e^{-\tau^2/2} \sum_{j=o}^{\infty} \frac{(\tau^2/2)^j}{j!} \int_o^{\infty} \frac{e^{-y/2}\, y^{\frac{n+2j}{2} - m-1}}{2^{\frac{n+2j}{2}}\, \Gamma(\frac{n+2j}{2})}\, dy \; .$$

Each term in the infinite series is positive and the first term (for j=0) is $+\infty$ if $n \leq 2m$. Thus, $E(v^{22}) = +\infty$, if $n \leq 2m$.

Now, suppose $n > 2m$. Then

$$E\,(v^{22})^m|S_3 = e^{-\tau^2/2} \sum_{j=o}^{\infty} \frac{(\tau^2/2)^j}{j!} \frac{\Gamma(\frac{n+2j}{2} - m)}{\Gamma(\frac{n+2j}{2})}\, 2^{-m}$$

$$\leq e^{-\tau^2/2} \sum_{j=o}^{\infty} \frac{(\tau^2/2)^j}{j} \max\left[\frac{\Gamma(\frac{n}{2} - m)2^{-m}}{\Gamma(\frac{n}{2})}, 2^{-m}\right]$$

$$= \max\left[\frac{\Gamma(\frac{n}{2} - m)2^{-m}}{\Gamma(\frac{n}{2})}, 2^{-m}\right] .$$

This implies that

$$E(v^{22})^m \leq \max\left[\frac{\Gamma(\frac{n}{2} - m)2^{-m}}{\Gamma(\frac{n}{2})}, 2^{-m}\right] . \quad \underline{\text{Q.E.D.}}$$

Theorem 4.1. Let $\hat{\beta}_1$ *be the first component of* $\hat{\beta}$.

Then for m an arbitrary integer, $E(\hat{\beta}_1)^{2m}$ is finite if $2m < K_2-G_1+1$ and infinite if $2m \geq K_2-G_1+1$.

Proof. It suffices to prove that the conclusion of the theorem holds for $\hat{\beta}_1{*}$. Now let α_1 be the first component of $\underset{\sim}{\alpha} = (S_2'S_2)^{-1}S_2'(\underset{\sim}{s}_1 - E\underset{\sim}{s}_1)$. S_2 and s_1 are independent and hence

$$\underset{\sim}{\alpha}|S_2 \sim N_{G_1}\left[\underset{\sim}{0},(S_2'S_2)^{-1}\right],$$

which further implies that

$$\alpha_1|S_2 \sim N(0, v^{22}).$$

Therefore

$$E\alpha_1{}^{2m} = c_{2m}\; E(v^{22})^{2m}$$

where $c_{2j} = 1.3....(2j-1)$ for $j \geq 1$. By Lemma 4.2., it follows then that

$$\begin{aligned} E\alpha_1{}^{2m} &< +\infty \text{ if } 2m < K_2-G_1+1 \\ &= +\infty \text{ if } 2m \geq K_2-G_1+1. \end{aligned} \qquad (4.6)$$

We now show that a necessary and sufficient condition for the finiteness of $E(\beta_1{*})^{2m}$ is that $E\alpha_1{}^{2m}$ is finite. Then the theorem follows from (4.6) above.

Let $\underset{\sim}{\xi}' = (\xi_1,\ldots,\xi_{K_2})$ be the first row of $(S_2'S_2)^{-1}S_2'$ so that

$$\alpha_1 = \xi'(\underset{=}{s}_1 - E\underset{=}{s}_1)$$

$$= \sum_{i=i}^{K_2} \xi_i(s_{i1} - Es_{i1}), \qquad (4.7)$$

and

$$\beta_1 = \xi' s_1 = \sum_{i-1}^{K_2} \xi_i \, s_{i1}. \qquad (4.8)$$

Note that $\underset{\approx}{\xi}$ and $\underset{=}{s}_1$ are independent of each other and

$$\sum_{i=1}^{K_2} \xi_i^2 = v^{22} \qquad (4.9)$$

since the conditional variance of α_1 is v^{22}, and also $\sum_{i=1}^{K_2} \xi_i^2$ by (4.7).

Now, suppose $E\alpha_1^{2m} < +\infty$. Then by (4.6), $2m < K_2 - G_1 + 1$ and hence by Lemma 4.2, $E(v^{22})^m < +\infty$, or equivalently, $E\left(\sum_{i=1}^{K_2} \xi_i^2\right)^m < +\infty$. This implies that $E\,\xi_i^{2m} < +\infty$ for $i=1,2,\ldots,K$ and hence $E(\xi_i s_{i1})^{2m} < +\infty$ for $i = 1,2,\ldots,K_2$

since ξ_i and s_{i1} are independent and all moments of s_{i1} exist. This finally implies that $E\left(\sum_{i=1}^{K_2} \xi_i \, s_{i1}\right)^{2m}$ is finite; or equivalently, $E(\hat{\beta}_1{*})^{2m}$ is finite.

To show that the other direction of the implication also holds, we write

$$E\,(\hat{\beta}_1{*})^{2m}|S_2 = \sum_{h=o}^{2m} \binom{2m}{h} \left[E(\hat{\beta}_1{*}|S_2)\right]^{2m-h} E\left[\alpha_1^h\,|S_2\right]$$

$$= \sum_{j=o}^{m} c_{2j}\binom{2m}{2j}\left[E(\hat{\beta}_1{*}|S_2)\right]^{2(m-j)} (v^{22})^j$$

$$= c_{2m}(v^{22})^m + \sum_{j=o}^{m-1} c_{2j}\binom{2m}{2j}(v^{22})^j\left[E(\hat{\beta}_1{*}|S_2)\right]^{2(m-j)} \tag{4.10}$$

where $c_0=1$ and c_{2j}, $j=1,2,\ldots, m$ are as defined previously. If $E(\alpha_1)^{2m} = +\infty$, then $2m \geq K_2-G_1+1$ and hence $E(v^{22})^m = +\infty$, by Lemma 4.2. This implies that $E(\beta_1{*})^{2m} = +\infty$ since all the terms in (4.10) are non-negative. This completes the proof of the theorem. Q.E.D.

By Proposition 2.1., we also have the following result concerning the OLS estimator as an immediate corollary to Theorem 4.1.

Corollary 4.1. *Let* $\tilde{\beta}_1$ *be the first component of the OLS estimator of* β. *Then for* m *an arbitrary integer,* $E(\tilde{\beta}_1)^{2m}$ *is finite if* $2m < N-K_1-G_1+1$ *and infinite if* $2m \geq N-K_1-G_1+1$.

REFERENCES

(1) Anderson, T.W. (1946). The non-central Wishart distribution and certain problems of multivariate statistics, Ann. Math. Stat., 17, 409-431.

(2) Anderson, T.W. (1964). Corrections to "Some extensions of the Wishart distribution" and "The non-central Wishart distribution and certain problems of multivariate statistics", Ann. Math. Stat., 35, 923-924.

(3) Anderson, T.W. and M.A. Girshick (1944). Some extensions of the Wishart distribution, Ann. Math. Stat., 15, 345-357.

(4) Basmann, R.L. (1961). A note on the exact finite sample frequency functions of generalized classical linear estimators in two leading over-identified cases, J. Amer. Stat. Assoc., 56, 619-636.

(5) Basmann, R.L. (1963). A note on the exact finite sample frequency functions of generalized classical linear estimators in a leading three equation case, J. Amer. Stat. Assoc., 58, 161-171.

(6) Johnston, J. (1963). Econometric Methods, McGraw - Hill Book Co., New York.

(7) Richardson, D.H. (1968). The exact distribution of a structural coefficient estimator, J. Amer. Stat. Assoc., 63, 1214-1226.

(8) Sawa, T. (1969). The exact sampling distribution of ordinary least squares and two-stage least squares estimators, J. Amer. Stat. Assoc., 64, 923-937.